Comparative Cytogenetics

Content of volume **17**, 2023

Chromosome complements of *Channa lucius* and *C. striata* from Phu Quoc Island and karyotypic evolution in snakehead fishes (Actinopterygii, Channidae)

Denis V. Prazdnikov[1]

1 *Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninsky pr. 33, Moscow, 119071, Russia*

Corresponding author: Denis V. Prazdnikov (pdvfish3409@rambler.ru)

Citation: Prazdnikov DV (2023) Chromosome complements of *Channa lucius* and *C. striata* from Phu Quoc Island and karyotypic evolution in snakehead fishes (Actinopterygii, Channidae). Comparative Cytogenetics 17(1): 1–12. https://doi.org/10.3897/compcytogen.v17.i1.94943

Abstract

Snakehead fishes of the family Channidae are obligatory air-breathers freshwater predators, the vast majority of which belong to the genus *Channa* Scopoli, 1777. *Channa* species are characterized by high karyotypic diversity due to various types of chromosomal rearrangements. It is assumed that, in addition to the lifestyle, fragmentation and isolation of snakehead populations contribute to an increase in karyotypic diversity. However, the chromosome complements of many isolated populations of widespread *Channa* species remain unknown, and the direction of karyotype transformations is poorly understood. This paper describes the previously unstudied karyotypes of *Channa lucius* (Cuvier, 1831) and *C. striata* (Bloch, 1793) from Phu Quoc Island and analyzes the trends of karyotypic evolution in the genus *Channa*. In *C. lucius*, the karyotypes are differed in the number of chromosome arms (2n = 48, NF = 50 and 51), while in *C. striata*, the karyotypes are differed in the diploid chromosome number (2n = 44 and 43, NF = 48). A comparative cytogenetic analysis showed that the main trend of karyotypic evolution of *Channa* species is associated with a decrease in the number of chromosomes and an increase in the number of chromosome arms, mainly due to fusions and pericentric inversions. The data obtained support the assumption that fragmentation and isolation of populations, especially of continental islands, contribute to the karyotypic diversification of snakeheads and are of interest for further cytogenetic studies of Channidae.

Keywords

Introduction

The family Channidae includes two genera of freshwater snakehead fishes (*Parachanna* Teugels et Daget, 1984 and *Channa* Scopoli, 1777) with a disjunct range (Courtenay and Williams 2004; Rüber et al. 2020). The genus *Parachanna* is restricted to tropical Africa and contains three species. The genus *Channa* is more numerous in terms of the number of species (more than 40 species described to date) distributed mainly in Southern Asia (Fricke et al. 2022; Froese and Pauly 2022). The estimated number of species in this genus varies as the group is subject to frequent taxonomic revisions and the currently accepted nominal species may constitute species complexes (Adamson et al. 2010; Cioffi et al. 2015; Conte-Grand et al. 2017).

Appearing in the Eocene, snakehead fishes have undergone a long evolution with multiple range expansions and repeated contacts with lineages that had diverged in isolation (Adamson et al. 2010; Rüber et al. 2020), reflected in their karyotypic diversity. Among the cytogenetically studied *Channa* species, the number of chromosomes varies from 2n = 32 to 2n = 112, while the number of chromosome arms from NF = 46 to NF = 116 (Kumar et al. 2019). Given such high karyotypic diversity, it is obvious that the evolutionary dynamism in the genus *Channa* is a result of various types of chromosomal rearrangements, the main of which are pericentric inversions, fusions, and polyploidization (Dhar and Chatterjee 1984; Rishi and Haobam 1990; Tanomtong et al. 2014; Cioffi et al. 2015). At the same time, the trends of karyotypic evolution remain poorly understood.

In the course of evolution, snakehead fishes developed accessory air-breathing organs, which allows them to do without water for a long time and migrate over land to colonize new habitats (Sayer 2005; Bressman et al. 2019). These features have contributed to the distribution of snakeheads in a suitable climate zone, in particular species, such as *Channa lucius* and *C. striata*, which have relatively large ranges in South Asia, including many freshwater habitats on both mainland and islands (Adamson et al. 2010; Tan et al. 2012). The wide geographical distribution associated with the lifestyle of *Channa* species, together with complex hydrographic and geological events in their habitats, led to the fragmentation and isolation of populations (Adamson et al. 2012; Tan et al. 2012; Robert et al. 2019), which in turn could have contributed to karyotypic changes. For example, the geographical separation of South Asian populations of *C. punctata* contributed to the fixation of various types of chromosomal rearrangements in different parts of the range, which led to karyotypic variability from 2n = 32 (NF = 58–64) to 2n = 34 (NF = 64) (Ruma et al. 2006; Kumar et al. 2013; Rakshit et al. 2015). Interpopulation chromosomal variability found among the cytogenetically studied snakehead species has led to the assumption that lifestyle, fragmentation, and isolation of populations contribute to an increase in karyotypic diversity (Cioffi et al. 2015). In this regard, it is of interest to study karyotypes in previously unexplored small and/or isolated populations of widespread *Channa* species.

This study presents chromosome complements of *C. lucius* and *C. striata* from Phu Quoc Island and comparative cytogenetic analysis (chromosome number and karyotype composition) of the genus *Channa*. The trends in the karyotypic evolution of snakeheads and chromosomal diversification due to the isolation of island populations are discussed.

Material and methods

Individuals of *Channa* species were collected from Phu Quoc Island (Gulf of Thailand, Vietnam) (Fig. 1) in December of 2011. Four individuals (two females and two males) of *C. lucius* (Cuvier, 1831) were karyotyped from the Bai Dai River and Duong Dong River basins (Fig. 1). Six individuals (two females and four males) of *C. striata* (Bloch, 1793) were karyotyped from the Bai Dai River basin (Fig. 1). Snakehead vouchers were deposited in the Southern Department of the Vietnam-Russian Tropical Center (Ho Chi Minh City). The total number of metaphase plates studied for each species was 65 and 82, respectively.

Chromosome preparations were obtained from the anterior part of the kidney according to previously published methods (Ojima and Kurishita 1980; Blanco et al. 2012) with the initial treatment of live fish with colchicine (injection of 0.08% solution into the spinal muscle). The anterior kidney tissue was incubated in 0.075 M KCl (hypotonic solution) for 24 min at 28 °C and fixed in 96% ethanol mixed with glacial acetic acid (3:1 ratio). Chromosome preparations made using standard air-drying techniques were stained with 5% Giemsa solution in phosphate buffer at pH 6.8 for 7 min. Mitotic chromosomes were analyzed under a microscope Leica DM 1000 with DFC 295 camera and LAS EZ software. Chromosomes were classified as metacentric (m), submetacentric (sm), subtelocentric (st), and acrocentric (a) according to their arm ratios (Levan et al. 1964). To determine the number of chromosome arms (NF), chromosomes of the m and sm groups were considered biarmed and those of the st/a group uniarmed. For statistical analysis of the results and data visualization, I used Excel 2021 software. The regression between the proportion of biarmed chromosomes and diploid chromosome number, and the Spearman correlation were calculated.
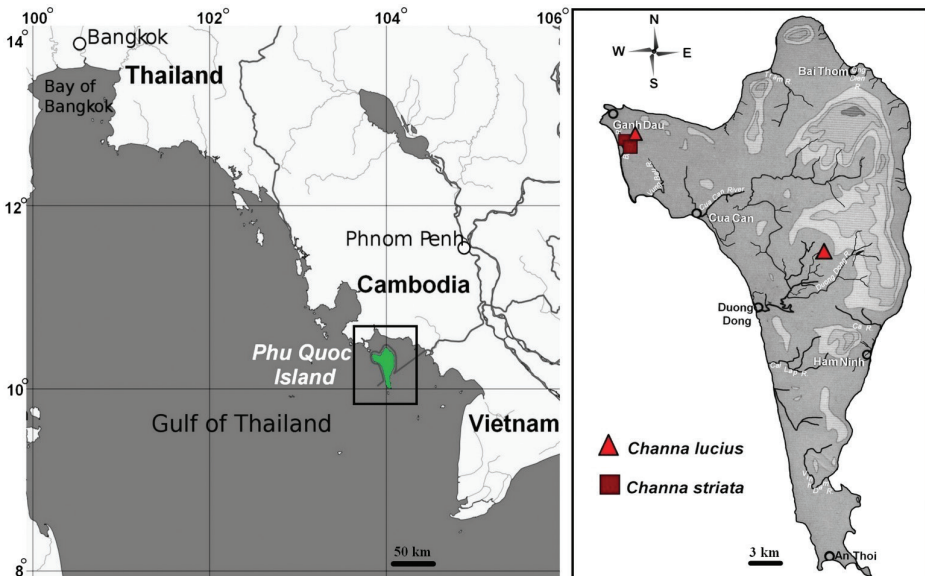
Figure 1. Map showing the location of Phu Quoc Island (left) and island details with *Channa* species collection sites (right).

Results and discussion

Karyotypic diversity in *Channa lucius* and *C. striata*

For *C. lucius* from both studied localities of Phu Quoc Island, the same diploid number of 2n = 48 was characteristic, but a different karyotype composition. In individuals from the Bai Dai River basin, the karyotype consisted of 2 metacentric chromosomes (m) and 46 subtelocentric and acrocentric chromosomes (st/a), NF = 50 (Fig. 2A). The karyotype of individuals from the Duong Dong River basin consisted of 3m and 45 st/a, NF = 51 (Fig. 2B). *C. striata* from the Bai Dai River basin had karyotypes differing in the number of biarmed chromosomes with 2n = 44 (2m+2sm+40st/a) and 2n = 43 (3m+2sm+38st/a), NF = 48 (Fig. 2C, D). In the two studied species, no differences were observed between male and female karyotypes.

Comparative analysis of island and mainland populations of *C. lucius* showed interpopulation chromosomal variability. Populations from Thailand and Phu Quoc Island differed in the number of m-chromosomes (Table 1), which is probably due to pericentric inversion. Populations of *C. striata* were characterized by different levels of chromosomal polymorphism. Previous studies of mainland populations of *C. striata* have shown marked karyotypic variability ranging from 2n = 40 to 2n = 44 (Table 1). For a population from Northeastern Thailand, an atypical karyotype with 2n = 43 was found containing an unpaired large m-chromosome (Cioffi et al. 2015). It is assumed that individuals with 2n = 43 could have arisen both as a result of hybridization of two parental karyotypes with 2n = 44 and 2n = 42, and as a result of centric fusion of chromosomes in *C. striata* with 2n = 44 (Cioffi et al. 2015). In the polymorphic population of *C. striata* from Phu Quoc Island, the karyomorph with 2n = 43 was heterozygous for centric fusion. The maintenance and preservation of such a heterozygous state with an unpaired m-chromosome in different populations of *C. striata* may be evidence in favor of the fact that individuals with 2n = 43 produce viable gametes. Interestingly, on Phu Quoc Island, heterozygous karyotypes were also found in goby fish (Prazdnikov 2018). Previous studies have revealed the important role of heterozygous chromosomal rearrangements in maintaining karyotypic diversity in different groups of animals (Guerrero and Kirkpatrick 2014; Dobigny et al. 2017; Llaurens et al. 2017; Wellenreuther and Bernatchez 2018).

The probable maximum age of isolation of Phu Quoc Island from the Cambodian mainland is about ten thousand years when sea levels rose after the end of the last glacial period (Kuznetsov and Kuznetsova 2011). The short-term isolation of the island populations of *C. lucius* and *C. striata* probably contributed either to the appearance of chromosomal polymorphism or its maintenance due to at least two types of chromosomal rearrangements. Further cytogenetic studies of these two snakehead species from different river basins of the island and an increase in the sample size will most likely reveal an even greater range of variability in the number of 2n and NF.
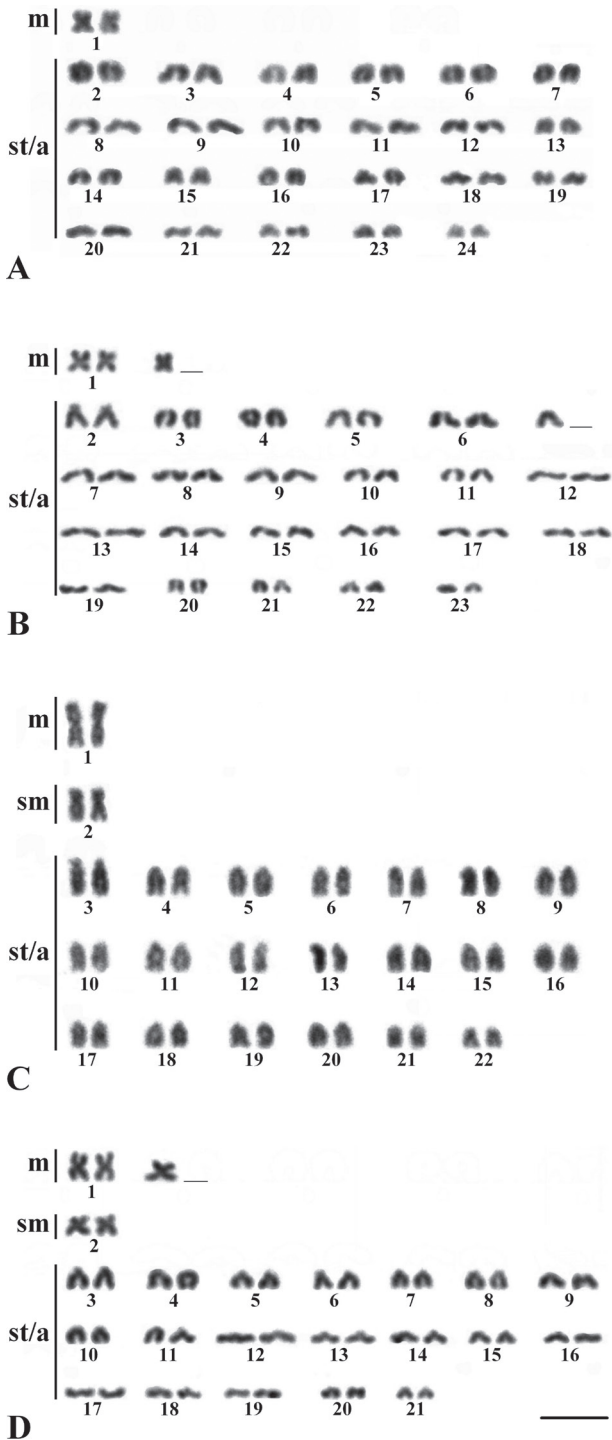
Figure 2. Karyotypes of *Channa lucius* (**A, B**) and *C. striata* (**C, D**) from Phu Quoc Island. Numerals indicate the paired chromosomes. Scale bar: 5 μm.

Table 1. Diploid chromosome number (2n), chromosome arm number (NF), karyotype structure, and collection site of *Channa lucius* and *C. striata*.

Species	2n	NF	Karyotype structure	Locality	References
C. lucius	48	50	2m+46st/a	Northeastern Thailand (Bung Klua reservoir in the Roi-Et)	Khakhong et al. 2014
	48	52	2m+2sm+2st+42a	Thailand	Donsakul and Magtoon 1991
	48	52	4m/sm/st+44a	Southern Thailand (Tapi Basin)	Cioffi et al. 2015
	48	50	2m+46st/a	Vietnam (Phu Quoc Island, Bai Dai River Basin)	This study
	48	51	3m+45st/a	Vietnam (Phu Quoc Island, Duong Dong River Basin)	This study
C. striata	40	50	8m+2sm+2st+28a	India	Banerjee et al. 1988
	40	48	8m+6st+26a	India (Assam, Meghalaya)	Dhar and Chatterjee 1984
	40	48	8m+2st+30a	India (Imphal)	Rishi and Haobam 1990
	40	50	8m+2sm+30st/a	India (WB)	Manna and Prasad 1973
	40	58	8m+10sm+22a	India (Manipur)	Sobita and Bhagirath 2006
	40	48	6m+2sm+10st+22a	Northeastern India	Kumar et al. 2013; 2019
	42	48	6m+36st/a	Northeast Thailand (Khon Kaen, Mahasakam)	Supiwong et al. 2009
	44	46	2m+42a	China	Wu et al. 1994
	43	50	7m/sm/st+36a	Northeastern Thailand (Chi Basin)	Cioffi et al. 2015
	44	50	6m/sm/st+38a	Central and Southern Thailand (Chao Phraya Basin, Tapi Basin)	Cioffi et al. 2015
	44	48	2m+2sm+40a	Thailand	Donsakul and Magtoon 1991
	44	48	2m+2sm+40st/a	Vietnam (Phu Quoc Island, Bai Dai River Basin)	This study
	43	48	3m+2sm+38st/a		This study

Karyotypic evolution in *Channa* species

An analysis of cytogenetic data (Dhar and Chatterjee 1984; Banerjee et al. 1988; Rishi and Haobam 1990; Sobita and Bhagirath 2006; Tanomtong et al. 2014; Cioffi et al. 2015; Kumar et al. 2019) indicated that the karyotypic evolution in *Channa* species occurred in different directions and at different rates, which led to a wide chromosomal diversity from 2n = 32 to 2n = 112 (Fig. 3). The proportion of biarmed chromosomes in the karyotype varies widely from 0% to 100%. The regression between the proportion of biarmed chromosomes in the karyotype and the diploid number is y = -0.0046x + 0.551 (R^2 = 0.102), and the Spearman correlation is R_s = -0.28 (Fig. 4). The weak correlation between the two variables (2n and proportion of m/sm chromosomes) is apparently due to chromosomal rearrangements that affected the trends of karyotypic evolution in the genus *Channa*.

The probable ancestral karyotype of snakeheads consisted of 48 uniarmed chromosomes, which would require a minimum number of chromosome rearrangements during the karyotype transformation of the number of *Channa* species. Among the cytogenetically studied species, *C. argus* and *C. lucius* have a karyotype with 2n = 48; the latter is also characterized by plesiomorphic features, such as the gular scales, which is absent in most species of Asian snakeheads (Li et al. 2006). The main trend of karyotypic evolution of *Channa* species is associated with a decrease in the number of chromosomes due to centric fusions (Robertsonian translocations) and an increase in the number of chromosome arms due to pericentric inversions (Fig. 5). As a result, in some populations

of *C. punctata*, symmetrical karyotypes with 2n = 32 appeared, consisting exclusively of biarmed chromosomes (Dhar and Chatterjee 1984; Rakshit et al. 2015). Another direction of karyotype transformation is associated with an increase in the number of chromosomes as a result of centric fission and polyploidization, followed by an increase
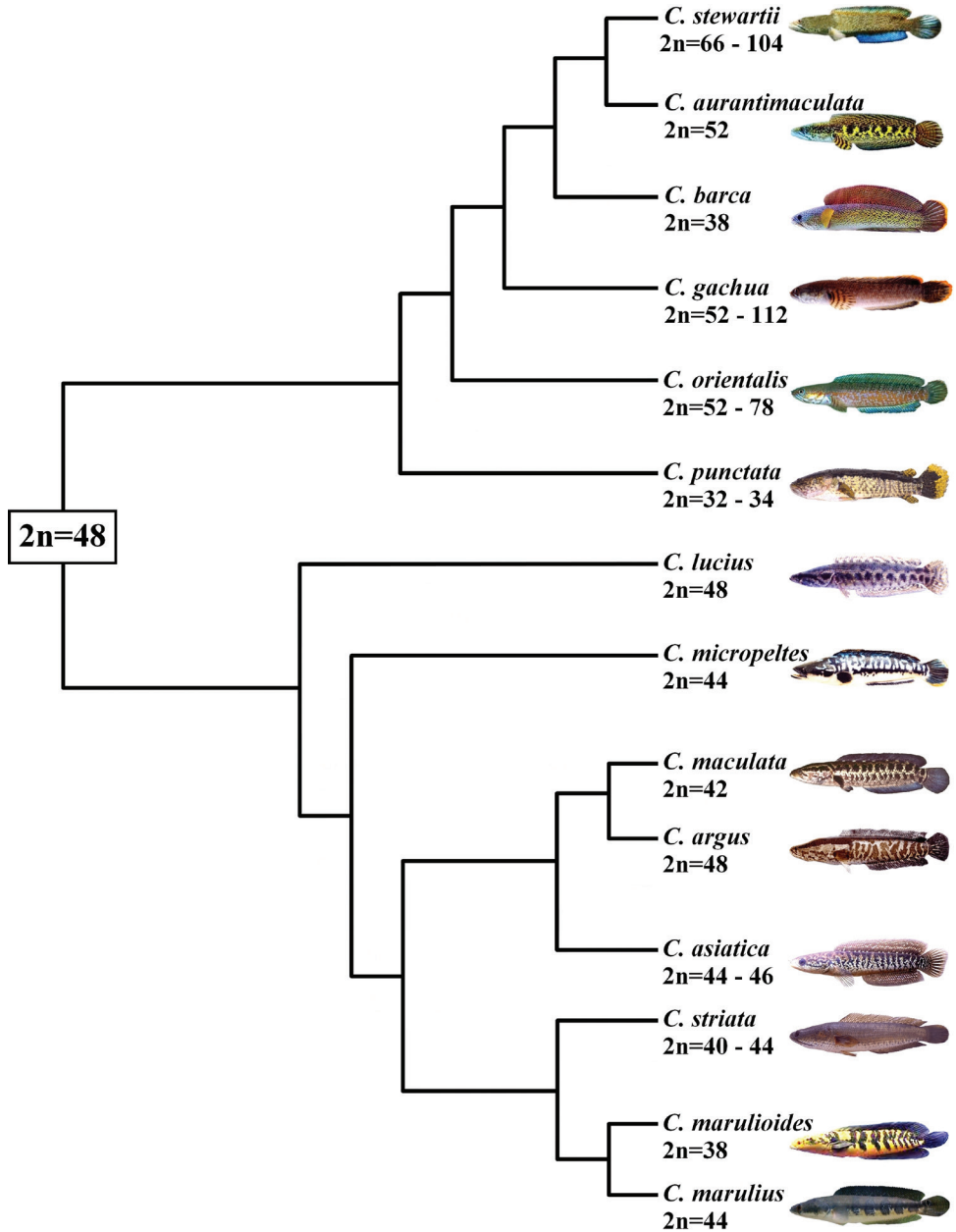
Figure 3. Phylogenetic tree of the cytogenetically studied *Channa* species (based on Kumar et al. 2019 with modifications and additions) indicating the putative ancestral karyotype (in a rectangle) and range of variability of diploid chromosome numbers.

in the proportion of m/sm chromosomes as a result of centric fusions and pericentric inversions (Fig. 5). The huge variability in 2n of *C. gachua* (2n = 52–112), *C. orientalis* (2n = 52–78), and *C. stewartii* (2n = 66–104) indicates both the possibility of ploidy change in different populations (Cioffi et al. 2015; Kumar et al. 2019) and their co-evolution (Tanomtong et al. 2014). An additional direction of karyotype transformation in *Channa* is associated with an increase in the proportion of biarmed chromosomes without changing 2n (mainly due to pericentric inversions) (Fig. 5). It is likely that another mechanism, such as centromere repositioning, could also be involved in alterations in chromosome morphology (Sobita and Bhagirath 2006; Rakshit et al. 2015).

Chromosomal rearrangements, which involve karyotypic structural changes such as inversions and fusions, may play an important role in the adaptive evolution of fish (Wellband et al. 2019; Cayuela et al. 2020). Rearrangements disturb homologous chromosome pairing during meiosis, resulting in tight linkage among genes encoding adaptations (for example, to salinity gradient and temperature) within rearranged regions (Barth et al. 2017; Wellenreuther and Bernatchez 2018). Such chromosomal rearrangements suppress recombination, and important functional genes are inherited together, which may contribute to adaptive population divergence (Barth et al. 2017). *C. gachua* is known to be well adapted to survive in a variety of habitats, in higher mountain areas with fluctuating climates, and has more resistance than other *Channa* species (Courtenay and Williams 2004; Tanomtong et al. 2014), which may be due to the high level of karyotypic variability.
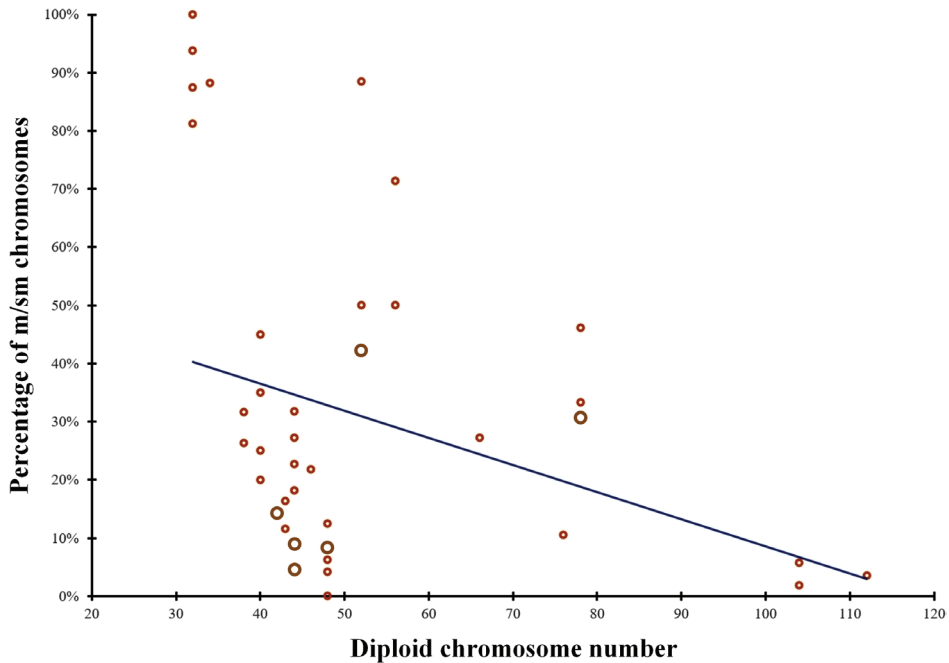
Figure 4. Scatter plot of a diploid chromosome number and proportion of metacentric/submetacentric chromosomes (m/sm) with overall regression line for the genus *Channa*. The diameter and color of a circle indicate the number of species from 1 to 2.
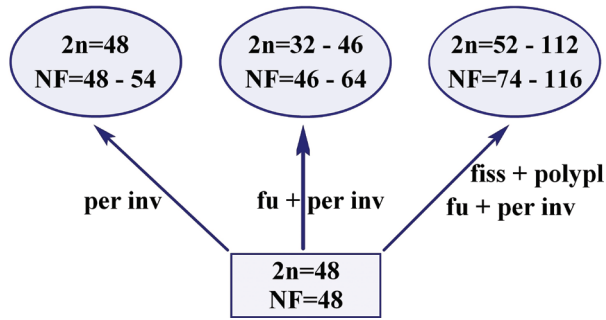
Figure 5. The trends of karyotypic evolution for the genus *Channa*. The thick blue arrow shows the most likely main trend in karyotypic evolution. The lower rectangle shows the ancestral karyotype. fu – centric fusions, fiss – centric fissions, per inv – pericentric inversions, polypl – polyploidization.

The proposed expansion of the ranges of modern taxa of Asian snakeheads at the Miocene/Pliocene boundary, combined with climatic fluctuations, led to repeated isolations of populations, especially continental islands, and secondary contacts between them (Adamson et al. 2010; Tan et al. 2012; Wang et al. 2021), which probably influenced the chromosome diversification in the genus *Channa* with the formation of intrapopulation and interpopulation chromosomal variability. The karyotypic diversity of snakeheads can also increase as a result of hybridization, which is possible even between species that differ in the number of 2n and NF (Ou et al. 2018). Obviously, further studies will make it possible to reveal even greater karyotypic diversity associated with the appearance of biarmed chromosomes within the framework of the main trend in the karyotypic evolution of *Channa* species.

Acknowledgements

References

Adamson EA, Hurwood DA, Mather PB (2010) A reappraisal of the evolution of Asian snakehead fishes (Pisces, Channidae) using molecular data from multiple genes and fossil calibration. Molecular Phylogenetics and Evolution 56(2): 707–717. https://doi.org/10.1016/j.ympev.2010.03.027

Adamson EA, Hurwood DA, Mather PB (2012) Insights into historical drainage evolution based on the phylogeography of the chevron snakehead fish (*Channa striata*) in the Mekong Basin. Freshwater Biology 57(11): 2211–2229. https://doi.org/10.1111/j.1365-2427.2012.02864.x

Banerjee SK, Misra KK, Banerjee S, Ray-Chaudhuri SP (1988) Chromosome numbers, genome sizes, cell volumes and evolution of snake-head fish (family Channidae). Journal of Fish Biology 33(5): 781–789. https://doi.org/10.1111/j.1095-8649.1988.tb05523.x

Barth JM, Berg PR, Jonsson PR, Bonanomi S, Corell H, Hemmer-Hansen J, Jakobsen KS, Johannesson K, Jorde PE, Knutsen H, Moksnes PO, Star B, Stenseth NC, Svedang H, Jentoft S, Andre C (2017) Genome architecture enables local adaptation of Atlantic cod despite high connectivity. Molecular Ecology 26(17): 4452–4466. https://doi.org/10.1111/mec.14207

Blanco DR, Bertollo LAC, Lui RL, Vicari MR, Margarido VP, Artoni RF, Moreira-Filho O (2012). A new technique for obtaining mitotic chromosome spreads from fishes in the field. Journal of Fish Biology 81(1): 351–357. https://doi.org/10.1111/j.1095-8649.2012.03325.x

Bressman NR, Love JW, King TW, Horne CG, Ashley-Ross MA (2019) Emersion and terrestrial locomotion of the northern snakehead (*Channa argus*) on multiple substrates. Integrative Organismal Biology 1(1): obz026. https://doi.org/10.1093/iob/obz026

Cayuela H, Rougemont Q, Laporte M, Mérot C, Normandeau E, Dorant Y, Tørresen OK, Hoff SNK, Jentoft S, Sirois P, Castonguay M, Jansen T, Praebel K, Clément M, Bernatchez L (2020) Shared ancestral polymorphisms and chromosomal rearrangements as potential drivers of local adaptation in a marine fish. Molecular Ecology 29(13): 2379–2398. https://doi.org/10.1111/mec.15499

Cioffi MB, Bertollo LAC, Villa MA, de Oliveira EA, Tanomtong A, Yano CF (2015) Genomic organization of repetitive DNA elements and its implications for the chromosomal evolution of channid fishes (Actinopterygii, Perciformes). PLoS ONE 10(6): e0130199. https://doi.org/10.1371/journal.pone.0130199

Conte-Grand C, Britz R, Dahanukar N, Raghavan R, Pethiyagoda R, Tan HH, Hadiaty RK, Yaakob NS, Rüber L (2017) Barcoding snakeheads (Teleostei, Channidae) revisited: Discovering greater species diversity and resolving perpetuated taxonomic confusions. PLoS ONE 12(9): e0184017. https://doi.org/10.1371/journal.pone.0184017

Courtenay WR, Williams JD (2004) Snakeheads (Pisces, Channidae): A Biological Synopsis and Risk Assessment, US Geological Survey Circular. US Geological Survey. Denver, 143 pp. https://doi.org/10.3133/cir1251

Dhar NJ, Chatterjee K (1984) Chromosomal evolution in Indian murrels (Channiformes: Channidae). Caryologia 37(4): 359–371. https://doi.org/10.1080/00087114.1984.10797714

Dobigny G, Britton-Davidian J, Robinson TJ (2017) Chromosomal polymorphism in mammals: an evolutionary perspective. Biological Reviews 92(1): 1–21. https://doi.org/10.1111/brv.12213

Donsakul T, Magtoon W (1991) A chromosome study on five species of fishes (*Channa*, family Channidae), from Thailand. Proceedings of 29th Kasetsart University Annual Conference (Fisheries section). Bangkok, 561–574.

Froese R, Pauly D (2022) FishBase. World Wide Web electronic publication. www.fishbase.org [version (06/2022). Electronic version accessed 28 Aug 2022]

Fricke R, Eschmeyer WN, Van der Laan R (2022) Eschmeyer's Catalog of Fishes: Genera, Species, References. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [Electronic version accessed 8 Aug 2022]

Guerrero R., Kirkpatrick M (2014) Local adaptation and the evolution of chromosome fusions. Evolution 68(10): 2747–2756. https://doi.org/10.1111/evo.12481

Khakhong S, Supiwong W, Tanomtong A, Sriuttha M, Jearranaiprepame P, Soemphol W, Ji-wyam W (2014) A First Chromosomal Characterization of NORs in Splendid Snakehead Fish, *Channa lucius* (Perciformes, Channidae). Cytologia 79(2): 133–139. https://doi.org/10.1508/cytologia.79.133

Kumar R, Kushwaha B, Nagpure N S, Behera BK, Lakra WS (2013) Karyological and molecular diversity in three freshwater species of the genus *Channa* (Teleostei, Perciformes) from India. Caryologia 66(2): 109–119. https://doi.org/10.1080/00087114.2013.821829

Kumar R, Baisvar VS, Kushwaha B, Waikhom G, Singh M (2019) Evolutionary analysis of genus *Channa* based on karyological and 16S rRNA sequence data. Journal of Genetics 98(5): 1–14. https://doi.org/10.1007/s12041-019-1156-4

Kuznetsov AN, Kuznetsova SP (2011) Forest vegetation of Phu Quoc Island. In: Kalyakin MV (Ed.) Materials of zoological and botanical studies in Phu Quoc Island, South of Vietnam. KMK Scientific Press, Moscow–Hanoi, 7–52. [In Russian with English summary]

Levan A, Fredg A, Sandberg A (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201–220. https://doi.org/10.1111/j.1601-5223.1964.tb01953.x

Li X, Musikasinthorn P, Kumazawa Y (2006) Molecular phylogenetic analyses of snakeheads (Perciformes: Channidae) using mitochondrial DNA sequences. Ichthyological Research 53: 148–159. https://doi.org/10.1007/s10228-005-0321-3

Llaurens V, Whibley A, Joron M (2017) Genetic architecture and balancing selection: the life and death of differentiated variants. Molecular Ecology 26(9): 2430–2448. https://doi.org/10.1111/mec.14051

Manna GK, Prasad R (1973) Chromosomes in three species of fish (*Channa*). Nucleus 16: 150–157.

Ojima Y, Kurishita A (1980) A new method to increase the number of mitotic cells in the kidney tissue for fish chromosome studies. Proceedings of the Japan Academy, Series B 56(10): 610–615. https://doi.org/10.2183/pjab.56.610

Ou M, Zhao J, Luo Q, Hong X, Zhu X, Liu H, Chen K (2018) Characteristics of hybrids derived from *Channa argus*♀× *Channa maculata*♂. Aquaculture 492: 349–356. https://doi.org/10.1016/j.aquaculture.2018.04.038

Prazdnikov DV (2018) Chromosomal variability and karyotypic evolution in island populations of channid and gobioid fishes (Actinopterygii, Perciformes). International Conference Chromosome 2018. Novosibirsk, August 20–24, 2018. Novosibirsk, 61–62.

Rakshit A, Paul A, Bhattacharjee S, Banik T, Saran R, Mandal B, Poddar D, Gangopadhyay, K. (2015). Cytogenetic and molecular profiling of spotted snake head fish *Channa punctatus* (Bloch, 1793) from three districts (Nadia, Hooghly and north 24 Parganas) of west Bengal, India. International Journal of Fisheries and Aquatic Studies 3(1): 312–319.

Rishi KK, Haobam MS (1990) A chromosomal study on four species of snake-heads (Ophiocephalidae: Pisces) with comments on their karyotypic evolution. Caryologia 43(2): 163–167. https://doi.org/10.1080/00087114.1990.10796995

Robert R, Amit NH, Sukarno NM, Majapun RJ, Kumar SV (2019) Population genetic structure of Asian snakehead fish (*Channa striata*) in North Borneo: Implications for conservation of local freshwater biodiversity. Ecological Research 34(1): 55–67. https://doi.org/10.1111/1440-1703.1008

Rüber L, Tan HH, Britz R (2020) Snakehead (Teleostei: Channidae) diversity and the Eastern Himalaya biodiversity hotspot. Journal of Zoological Systematics and Evolutionary Research 58(1): 356–386. https://doi.org/10.1111/jzs.12324

Ruma F, Ahmed ATA, Alam SS (2006) Karyotype analysis of *Channa punctata* Bloch and *Channa orientalis* Schneider with Giemsa, CMA and DAPI. Cytologia 71(4): 425–430. https://doi.org/10.1508/cytologia.71.425

Sayer MD (2005) Adaptations of amphibious fish for surviving life out of water. Fish and Fisheries 6(3): 186–211. https://doi.org/10.1111/j.1467-2979.2005.00193.x

Sobita N, Bhagirath T (2006) Chromosomal differentiations in the evolution of channid fishes—molecular genetic perspective. Caryologia 59(3): 235–240. https://doi.org/10.1080/00087114.2006.10797920

Supiwong W, Jearranaiprepame P, Tanomtong A (2009) A new report of karyotype in the chevron snakehead fish, *Channa striata* (Channidae, Pisces) from Northeast of Thailand. Cytologia 74: 317–322. https://doi.org/10.1508/cytologia.74.317

Tan MP, Jamsari AFJ, Siti Azizah MN (2012) Phylogeographic pattern of the striped snakehead, *Channa striata* in Sundaland: ancient river connectivity, geographical and anthropogenic singnatures. PLoS ONE 7(12): e52089. https://doi.org/10.1371/journal.pone.0052089

Tanomtong A, Supiwong W, Jearranaiprepame P, Khakhong S, Kongpironchuen C, Getlekha N (2014) A new natural autotetraploid and chromosomal characteristics of dwarf snakehead fish, *Channa gachua* (Perciformes, Channidae) in Thailand. Cytologia 79(1): 15–27. https://doi.org/10.1508/cytologia.79.15

Wang J, Li C, Chen J, Wang J, Jin J, Jiang S, Yan L, Lin H, Zhao J (2021) Phylogeographic structure of the dwarf snakehead (*Channa gachua*) around Gulf of Tonkin: Historical biogeography and pronounced effects of sea-level changes. Ecology and Evolution 11(18): 12583–12595. https://doi.org/10.1002/ece3.8003

Wellband K, Mérot C, Linnansaari T, Elliott JAK, Curry RA, Bernatchez L (2019) Chromosomal fusion and life history-associated genomic variation contribute to within-river local adaptation of Atlantic salmon. Molecular Ecology 28(6): 1439–1459. https://doi.org/10.1111/mec.14965

Wellenreuther M, Bernatchez L (2018) Eco-evolutionary genomics of chromosomal inversions. Trends in Ecology & Evolution 33(6): 427–440. https://doi.org/10.1016/j.tree.2018.04.002

Wu G, Ma J, Hu H, Lou J, Chen K, Lin G 1994 The karyotype of *Channa striatus* and *Channa micropeltes*. Freshwater Fisheries Danshui Yuye 24(4): 3–5.

ORCID

Denis V. Prazdnikov https://orcid.org/0000-0003-0447-7431

COMPARATIVE
A peer-reviewed open-access journal
Cytogenetics
International Journal of Plant & Animal Cytogenetics,
Karyosystematics, and Molecular Systematics

Karyotype differentiation in the *Nothobranchius ugandensis* species group (Teleostei, Cyprinodontiformes), seasonal fishes from the east African inland plateau, in the context of phylogeny and biogeography

Eugene Yu. Krysanov[1], Béla Nagy[2], Brian R. Watters[3],
Alexandr Sember[4], Sergey A. Simanovsky[1]

1 *Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninsky Prospect 33, 119071, Moscow, Russia* **2** *15, voie de la Liberté, 77870, Vulaines sur Seine, France* **3** *6141 Parkwood Drive, Nanaimo, British Columbia V9T 6A2, Nanaimo, Canada* **4** *Laboratory of Fish Genetics, Institute of Animal Physiology and Genetics, Czech Academy of Sciences, Rumburská 89, 27721, Liběchov, Czech Republic*

Corresponding author: Sergey A. Simanovsky (sergey.a.simanovsky@gmail.com)

Abstract

The karyotype differentiation of the twelve known members of the *Nothobranchius ugandensis* Wildekamp, 1994 species group is reviewed and the karyotype composition of seven of its species is described herein for the first time using a conventional cytogenetic protocol. Changes in the architecture of eukaryotic genomes often have a major impact on processes underlying reproductive isolation, adaptation and diversification. African annual killifishes of the genus *Nothobranchius* Peters, 1868 (Teleostei: Nothobranchiidae), which are adapted to an extreme environment of ephemeral wetland pools in African savannahs, feature extensive karyotype evolution in small, isolated populations and thus are suitable models for studying the interplay between karyotype change and species evolution. The present investigation reveals a highly conserved diploid chromosome number (2n = 36) but a variable number of chromosomal arms (46–64) among members of the *N. ugandensis* species group, implying a significant role of pericentric inversions and/or other types of centromeric shift in the karyotype evolution of the group. When superimposed onto a phylogenetic tree based on molecular analyses of two mitochondrial genes the cytogenetic characteristics did not show any correlation with the phylogenetic relationships within the lineage. While karyotypes of many

other *Nothobranchius* spp. studied to date diversified mainly via chromosome fusions and fissions, the *N. ugandensis* species group maintains stable 2n and the karyotype differentiation seems to be constrained to intrachromosomal rearrangements. Possible reasons for this difference in the trajectory of karyotype differentiation are discussed. While genetic drift seems to be a major factor in the fixation of chromosome rearrangements in *Nothobranchius*, future studies are needed to assess the impact of predicted multiple inversions on the genome evolution and species diversification within the *N. ugandensis* species group.

Keywords

2n uniformity, chromosomes, chromosome evolution, chromosome inversion, cytogenetics, karyotype variability

Introduction

The cyprinodontiform fish genus *Nothobranchius* Peters, 1868 currently comprises 96 valid species, occurring mainly in seasonal wetlands of river drainages in northeastern, eastern and south-eastern Africa that are subject to seasonal rainfall (Nagy and Watters 2021). All known species feature an annual or semi-annual life cycle as a key adaptation to reproduce in an unpredictable biome of temporary freshwater pools that appear during monsoons, and which become desiccated during the dry season (Vanderplank 1940; Watters 2009; Nagy 2015). Because of their life cycle, annual killifishes form small populations with non-overlapping generations that are biogeographically isolated. Their low dispersal ability leads to strong spatial genetic structure of *Nothobranchius* spp. (Bartáková et al. 2013, 2015; Dorn et al. 2014) with a strong effect of genetic drift, including bottlenecks and founder effects, on their genome evolution (Bartáková et al. 2013; Cui et al. 2019; van der Merwe et al. 2021).

Nothobranchius spp. are small fishes, mostly reaching 30–70 mm in standard length, with only a few species achieving 100 mm or more. They show marked sexual dimorphism and dichromatism; the typically robust and colourful males contrast with the slightly smaller and dull-coloured females (Jubb 1981; Wildekamp 2004). Representative male phenotypes of the *Nothobranchius ugandensis* Wildekamp, 1994 species group are shown in Fig. 1. The male colour pattern is species-specific and thus provides an important diagnostic character for species discrimination (e.g., Jubb 1981; Nagy 2018; Nagy et al. 2020). The genus includes *N. furzeri* Jubb, 1971, the vertebrate species with the shortest lifespan recorded in captivity (less than 12 weeks), and which has emerged as a model organism for biological and molecular studies of ageing (e.g. Cellerino et al. 2016). Another species, *N. rachovii* Ahl, 1926, exhibits the lowest recorded diploid chromosome number (2n = 16) within the genus and one of the lowest diploid chromosome numbers among all karyotyped fishes (Arai 2011). With its remarkably large chromosomes, it is a convenient model for laboratory chromosome studies of fish genotoxicity (e.g. van der Hoeven et al. 1982; Krysanov 1992; Krysanov et al. 2018).

Phylogenetic analysis revealed that the genus *Nothobranchius* comprises a monophyletic lineage that includes seven subgenera in geographically segregated clades (van der Merwe et al. 2021). The *N. ugandensis* species group (sensu Nagy et al.
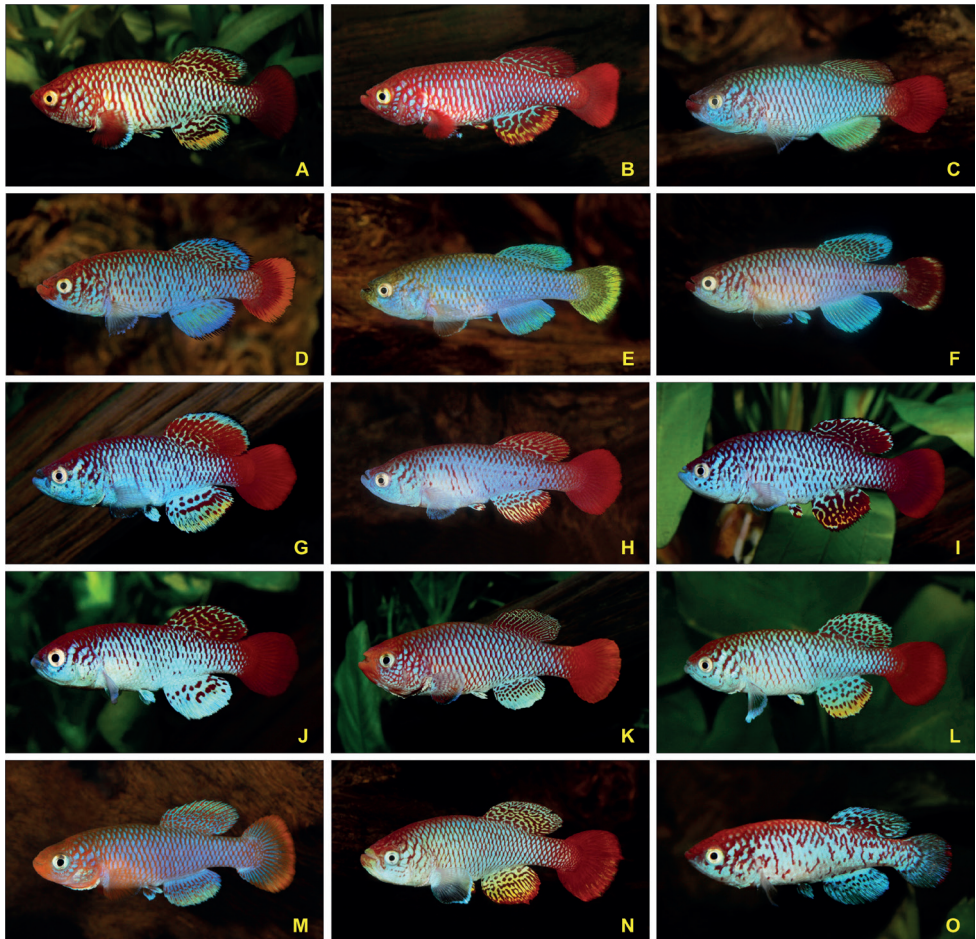
Figure 1. Selected male specimens of representatives of the *Nothobranchius ugandensis* species group (*denotes populations from which karyotype data was determined) **A** *N. nubaensis* Wadi Al Ghallah SD 10-5, southern Sudan **B** *N. nubaensis* Fugnido EHKS 09-01*, western Ethiopia **C** *N. albertinensis* Olobodagi UG 99-23, northwestern Uganda **D** *N. ugandensis* Busesa UG 99-5 (red phenotype), southeastern Uganda **E** *N. ugandensis* Busesa UG 99-5 (blue/yellow phenotype), southeastern Uganda **F** *N. ugandensis* Namasagali UG 99-3* (red phenotype with submarginal band in caudal fin), south-central Uganda **G** *N. derhami* Ahero KEN 19-16*, western Kenya **H** *N. attenboroughi* Nata TAN 93-3, north-central Tanzania **I** *N. venustus* Chato TZN 19-5*, north-central Tanzania **J** *N. moameensis* Mabuki TZN 19-8*, north-central Tanzania **K** *N. hoermanni* Bumburi TZHK 2018-03*, central Tanzania **L** *N. torgashevi* TNT 2014-04*, south-central Tanzania **M** *N. streltsovi* TSTS 10-05, south-central Tanzania **N** *N. itigiensis* Itigi TAN 03-8*, central Tanzania **O** *N. kardashevi* Mpanda K 2011-25*, southwestern Tanzania. The fishes on the photos have a size of 45–50 mm SL (standard length). Photographs by Béla Nagy (**A, B, G, I, J–L, O**) and Brian Watters (**C–F, H, M, N**).

2020) belongs to the subgenus *Zononothobranchius* Radda, 1969. The species group currently comprises 12 members, known from the inland plateau of eastern Africa (see Table 1 and Fig. 2).

Table 1. Listing of all known species of the *Nothobranchius ugandensis* species group with indication of associated drainage and region of occurrence.

Species	Drainage	Region of occurrence
N. albertinensis Nagy, Watters et Bellstedt, 2020	Lake Albert basin and Albert Nile drainage	North-western Uganda
N. attenboroughi Nagy, Watters et Bellstedt, 2020	Grumeti and other small systems draining into eastern shore of Lake Victoria	Northern Tanzania
N. moameensis Nagy, Watters et Bellstedt, 2020	Moame and other smaller river systems draining into southern shore of Lake Victoria	
N. derhami Valdesalici et Amato, 2019	Nyando system northeast of Lake Victoria	South-western Kenya
N. hoermanni Nagy, Watters et Bellstedt, 2020	Mhwala system in the upper Wembere drainage, and the Wala system, in the Malagarasi drainage	Central Tanzania
N. itigiensis Nagy, Watters et Bellstedt, 2020	Upper Ruaha drainage and the Bahi Swamp	
N. streltsovi Valdesalici, 2016	Nkululu, tributary of the Ugalla in the Malagarasi drainage	
N. torgashevi Valdesalici, 2015	Wembere drainage in the endorheic Lake Eyasi basin	
N. kardashevi Valdesalici, 2012	Katuma system	South-western Tanzania
N. nubaensis Valdesalici, Bellemans, Kardashev et Golubtsov, 2009	Wadi Al Ghallah system and Khor Abu Habl system in the White Nile drainage, and the Sobat system in the Blue Nile drainage	Southern Sudan and south-western Ethiopia
N. ugandensis Wildekamp, 1994	Lake Victoria and Lake Kyoga basins, and Victoria Nile and Achwa drainages	Central and northern Uganda, and south-western Kenya
N. venustus Nagy, Watters et Bellstedt, 2020	Small stream systems as part of southwestern shore of Lake Victoria basin, and Kongwa system in the southern part of the lake	North-western Tanzania

Cytogenetic data, available for 65 *Nothobranchius* species and a taxonomically undetermined *Nothobranchius* sp. Kasenga, indicate remarkable karyotype dynamics with chromosome counts ranging from 16 to 50 (Scheel 1990; Krysanov et al. 2016; Krysanov and Demidova 2018). Sex chromosomes of the XY type have been found in two closely related species, *N. furzeri* and *N. kadleci* Reichard, 2010 (Reichwald et al. 2015; Štundlová et al. 2022), while six other representatives with scattered positions across the phylogeny possess an $X_1X_1X_2X_2/X_1X_2Y$ multiple sex chromosome system (Ewulonu et al. 1985; Krysanov et al. 2016; Krysanov and Demidova 2018; Simanovsky et al. 2019). Consequently, the genus *Nothobranchius* represents an excellent model for studying processes that shape karyotype differentiation and their relevance to species diversification and reproductive isolation.

In the present study, we examined the karyotype differentiation of seven members of the *N. ugandensis* species group by conventional karyotyping. The karyotypes of the remaining five species of this group have been previously reported (Krysanov and Demidova 2018). Aiming to interpret all known cytogenetic patterns in the phylogenetic context, we constructed a phylogenetic tree based on two mitochondrial genes.

Materials and methods

In total, we analysed thirty-three individuals belonging to seven species from the *N. ugandensis* species group (details provided in Table 2). The experiments were carried
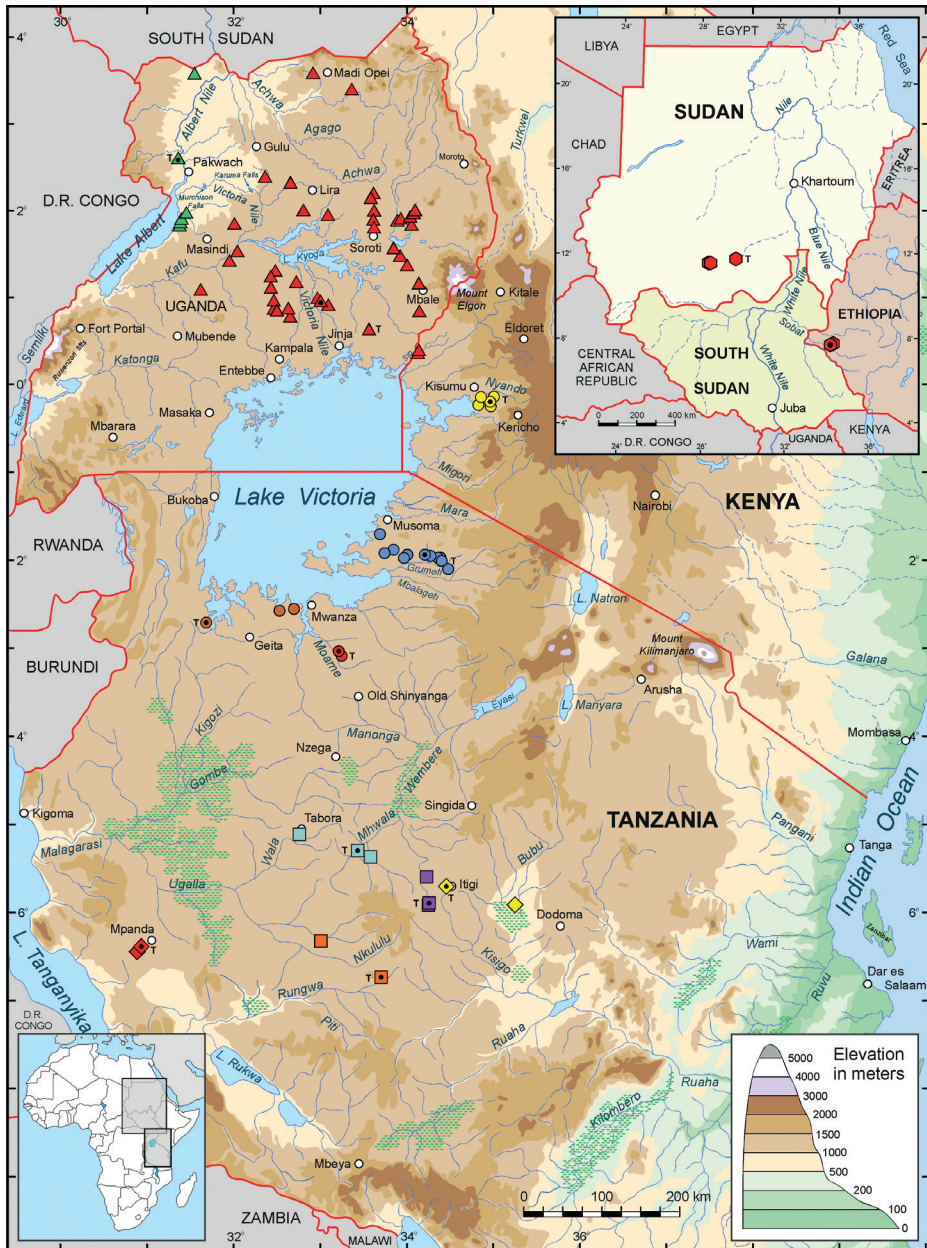
Figure 2. Distribution of species in eastern and northeastern Africa belonging to the *Nothobranchius ugandensis* species group: *N. albertinensis* (green triangle), *N. ugandensis* (red triangle), *N. derhami* (yellow-filled circle), *N. attenboroughi* (blue-filled circle), *N. venustus* (orange-brown-filled circle), *N. moameensis* (red-filled circle), *N. hoermanni* (blue-green square), *N. torgashevi* (purple square), *N. itigiensis* (yellow diamond), *N. streltsovi* (orange square), *N. kardashevi* (red diamond), and *N. nubaensis* (red hexagon; on inset map). T, type localities. Symbols with a black dot indicate sites of individuals used for karyotype analyses. Note that the presently known entire ranges of the respective species are shown, and individual symbols may in some cases represent multiple sites in close proximity to one another.

Table 2. Number of individuals karyotyped (N), population codes and geographic coordinates for studied members of the *Nothobranchius ugandensis* species group.

Species	N	Population code	GPS coordinates
N. albertinensis	2 larvae	Packwach UGN 17-16	02°36.31'N, 31°23.07'E
N. attenboroughi	4 larvae	Mugeta TAN 17-13	01°56.77'S, 34°14.25'E
N. derhami	2♀/2♂	Ahero KEN 19-16	00°12.85'S, 34°57.44'E
N. hoermanni	4♀/2♂	Bumburi TZHK 2018-03	05°18.23'S, 33°26.07'E
N. itigiensis	2♀/4♂	Itigi TAN 03-8	05°41.93'S, 34°28.80'E
N. kardashevi *	2♀/2♂	Mpanda K 2011-25	06°22.06'S, 30°56.16'E
N. moameensis	2♀/2♂	Mabuki TZN 19-8	03°01.46'S, 33°12.25'E
N. nubaensis *	2♀/2♂	Fugnido EHKS 09-01	07°44.48'N, 34°15.03'E
N. streltsovi *	2♀/2♂	TNT 2014-07	06°40.87'S, 33°41.00'E
N. torgashevi *	3♀/4♂	TNT 2014-04	05°53.09'S, 34°17.12'E
N. ugandensis *	2♀/3♂	Namasagali UG 99-3	00°57.41'N, 33°01.67'E
N. venustus	3♀/4♂	Chato TZN 19-5	02°42.59'S, 31°43.69'E

* Data from Krysanov and Demidova (2018).

out in accordance with the rules of the Severtsov Institute of Ecology and Evolution (IEE) and approved by IEE's Ethics Committee (orders No. 27 of November 9, 2018 and No. 55 of December 12, 2021).

Cytogenetic analysis

Chromosome preparations from adult individuals were obtained following Kligerman and Bloom (1977), with modifications described in Krysanov and Demidova (2018). For larvae a modified technique was used. The 1–2-week-old larvae were held in a 0.1% colchicine solution in aquarium water for 3–5 hours, then they were euthanized with an overdose of tricaine methanesulfonate (MS-222) and dissected under a Stemi 2000-C stereomicroscope (Carl Zeiss, Germany). All abdominal organs were taken for chromosome preparations. The organs were incubated with a 0.075M KCl hypotonic solution for 20 minutes and fixed in three changes of a 3:1 methanol: acetic acid solution for 20 minutes each. Finally, the fixed organs were incubated in 50–100 μL of 50% glacial acetic acid, suspended, and dropped onto hot slides (45 °C).

The chromosome spreads were air-dried, stained with 4% Giemsa solution in a phosphate buffer solution (pH 6.8) for 8 minutes and then analysed using an Axioplan 2 imaging microscope (Carl Zeiss, Germany) equipped with a CV-M4⁺CL camera (JAI, Japan) and Ikaros software (MetaSystems, Germany). At least 10 complete metaphases per individual were analysed. Final images were processed using Photoshop software (Adobe, USA). Karyotypes were arranged according to the centromere position following the nomenclature of Levan et al. (1964), but modified as metacentric (m), submetacentric (sm) and subtelocentric/acrocentric (st/a). Chromosome pairs were arranged according to their size in each chromosome category. To determine the chromosomal arm number per karyotype (nombre fondamental, NF), metacentrics and submetacentrics were considered as biarmed, and subtelocentrics/acrocentrics as monoarmed.

Phylogenetic analyses

We constructed the phylogenetic tree for the purpose of cytogenetic data interpretation. The sequences used for the phylogenetic analysis were from Nagy et al. (2020). However, only one representative per species was chosen for this study. The phylogenetic hypothesis was based on the analysis of two mitochondrial genes *Cytochrome oxidase subunit I (COI)* and *NADH dehydrogenase 2 (ND2)*. Multiple sequence alignment was performed with Clustal Omega (Sievers et al. 2011), and the alignments of the two genes were concatenated into a single dataset of 2511 bp in length. *Nothobranchius taeniopygus* Hilgendorf, 1891 and *N. rubroreticulatus* Blache et Miton, 1960 were selected as outgroup as representatives of closely related species groups. Phylogenetic analysis of the dataset was performed using Bayesian inference in MrBayes 3.2.7 (Ronquist and Huelsenbeck 2003). The analysis was set to Markov chain Monte Carlo simulation (mcmc) with default heating conditions. The evolutionary model for the GTR substitution model was set with gamma-distributed rate variation across sites and a proportion of invariable sites (GTR + I + I'), searching the tree space for 2 million generations starting with random trees and a sampling frequency of each 500 generations. The tree file was imported into Figtree 1.4.4. (Rambaut 2009) for tree drawing.

Results and discussion

Cytogenetic characteristics (2n, NF and karyotype structure) of the analysed representatives of the *N. ugandensis* species group are shown in Fig. 3 and Table 3. Known cytogenetic data for the *N. ugandensis* species group (Krysanov and Demidova 2018 and this study) are arranged in the context of phylogenetic tree analysis in Fig. 4. All twelve species share the same 2n = 36 and the largest pair of metacentric chromosomes (pairs No. 1; Fig. 3). At the same time, the species varied considerably regarding the ratio of monoarmed (subtelocentric, acrocentric) vs. biarmed (metacentric, submetacentric) chromosomes. Accordingly, NF ranged from 54 to 64 within our seven analysed species and from 46 to 64 when considering also the species studied by Krysanov and Demidova (2018). Within our sampling, we recorded the lowest number of biarmed chromosomes (18) in *N. moameensis*, while *N. derhami* had the highest number of such chromosomes (28). All species exhibited different karyotype structures except for *N. attenboroughi* and *N. ugandensis.* Notably, these two species are widely separated geographically and belong to different clades in the molecular phylogeny (Figs 2, 4). Lastly, we did not observe consistently any type of chromosome polymorphism within our sampling and thus we also did not detect any heteromorphic sex chromosomes or the presence of multiple sex chromosome systems.

According to data previously available for 66 representatives (including *N.* sp. Kasenga) (Krysanov et al. 2016; Krysanov and Demidova 2018), *Nothobranchius* killifishes display high karyotype variability. Here, we studied the karyotypes of seven members of the *N. ugandensis* species group and thus increased the number
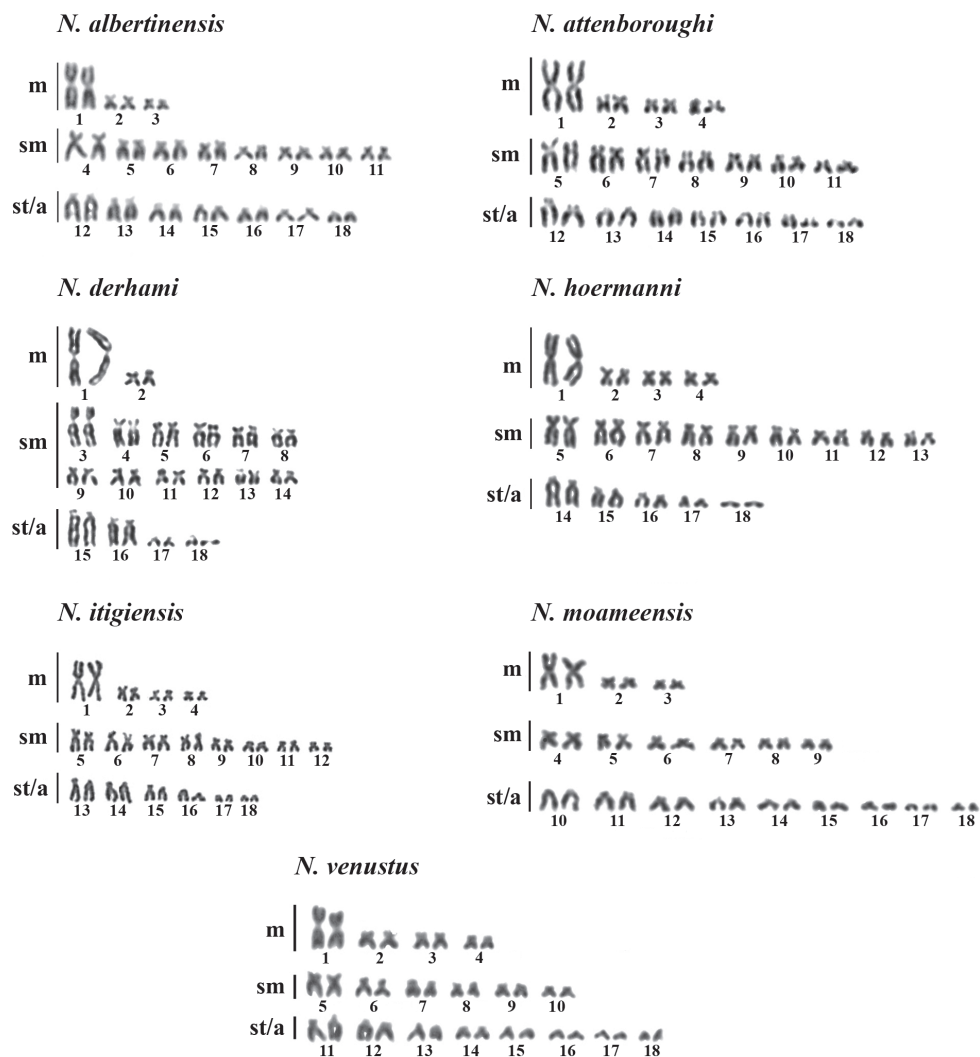
Figure 3. Karyotypes of seven studied members of the *Nothobranchius ugandensis* species group. Scale bar: 10 μm.

of chromosomally characterized representatives of the genus *Nothobranchius* to 73. A major finding of our survey is that all 12 studied species from the *N. ugandensis* species group maintain a stable 2n = 36 (Krysanov and Demidova 2018; this study) which contrasts with the generally extensive karyotype dynamics known for the *Nothobranchius* genus as a whole. Nevertheless, the karyotypes of the 12 species vary considerably in the proportion of monoarmed and biarmed chromosomes which is further reflected in the wide range of their NF values (46–64). Therefore, while the karyotypes of many other studied *Nothobranchius* spp. underwent frequent

Table 3. Diploid chromosome numbers (2n), numbers of chromosome arms (NF) and karyotype structure of all members of *Nothobranchius ugandensis* species group.

Species	2n	NF	Karyotype structure	References
N. albertinensis	36	58	6m + 16sm + 14st/a	This study
N. attenboroughi	36	58	8m + 14sm + 14st/a	This study
N. derhami	36	64	4m + 24sm + 8st/a	This study
N. hoermanni	36	62	8m + 18sm + 10st/a	This study
N. itigiensis	36	60	8m + 16sm + 12st/a	This study
N. kardashevi	36	62	6m + 20sm + 10st/a	Krysanov and Demidova 2018
N. moameensis	36	54	6m + 12sm + 18st/a	This study
N. nubaensis	36	62	14m + 12sm + 10st/a	Krysanov and Demidova 2018
N. streltsovi	36	48	6m + 6sm + 24st/a	Krysanov and Demidova 2018
N. torgashevi	36	46	6m + 4sm + 26st/a	Krysanov and Demidova 2018
N. ugandensis	36	58	8m + 14sm + 14st/a	Krysanov and Demidova 2018
N. venustus	36	56	8m + 12sm + 16st/a	This study
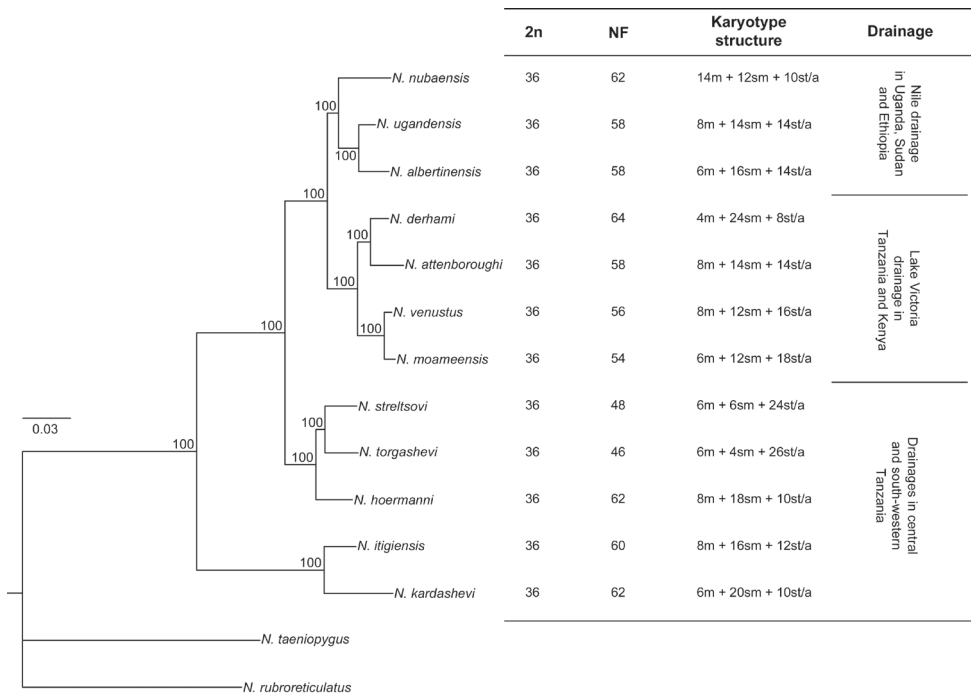
Figure 4. Karyotype characteristics and phylogenetic relationships, as well as associated drainage system information, for members of the *Nothobranchius ugandensis* species group. Karyotype characteristics are plotted onto the phylogenetic tree which is based on analysis of the mitochondrial molecular markers *Cytochrome oxidase subunit I (COI)* and *NADH dehydrogenase 2 (ND2)*, using Bayesian inference.

interchromosomal rearrangements (typically fusions and fissions) (Krysanov and Demidova 2018), karyotype differentiation in the *N. ugandensis* species group seems to be restricted to intrachromosomal structural changes that have led to shifts

in the centromere positions without changes of 2n. The most probable responsible mechanisms might be pericentric inversions (i.e., two-break rearrangements where the segment between the two breaks, which is then inverted by 180° and re-inserted to the chromosome, contains the centromere) and possibly also centromere repositioning (i.e., the replacement of the old centromere by the new one located elsewhere on the chromosome; Schubert 2018). Finally, we cannot exclude the possible contribution of other relevant rearrangements such as reciprocal and non-reciprocal translocations.

The stable 2n = 36 is also shared by all but four studied representatives belonging to the subgenus *Zononothobranchius* (Krysanov et al. 2016; Krysanov and Demidova 2018; van der Merwe et al. 2021) which encompasses the *N. ugandensis* species group and further the *N. brieni* species group (sensu Nagy 2018), the *N. neumanni* species group (sensu Wildekamp et al. 2014), the *N. rubroreticulatus* species group (sensu van der Merwe et al. 2021), and the *N. taeniopygus* species group (sensu Watters et al. 2019). Interestingly, all representatives of the subgenus with 2n other than 36 belong to the *N. brieni* species group. Since only the *N. ugandensis* species group has been fully cytogenetically characterised (Krysanov and Demidova 2018; this study), we cannot make any general conclusions about the karyotype stability/variability in the subgenus as a whole.

Nothobranchius genomes are known to harbour a high amount of repetitive DNA (about 60–80 %; Reichwald et al. 2009, 2015; Cui et al. 2019; Štundlová et al. 2022) that is capable of facilitating chromosome rearrangements (Redi et al. 1990; King 1993; Li et al. 2017; Brown and Freudenreich 2021) as has recently been documented for *N. furzeri* and *N. kadleci* (Štundlová et al. 2022). As the amount and distribution of repetitive DNA may vary considerably among *Nothobranchius* species (Voleníková et al. in prep.), a hypothesis worth testing experimentally would be to determine if the species from the *N. ugandensis* species group exhibit a low proportion of clustered repeats/heterochromatin blocks in their genomes, which would correspond to a limited rate of karyotype dynamics at the interchromosomal level. A striking example of positive correlation between the stable karyotypes and low amount of repeats/constitutive heterochromatin was described in haemulid fishes (Motta-Neto et al. 2019).

It is noteworthy that the *N. ugandensis* species group, forming part of the Inland Clade, diverged approximately 4 million years ago (MYA) according to van der Merwe et al. (2021). When compared to species groups in other *Nothobranchius* clades, the *N. ugandensis* species group had sufficient time for the establishment of at least some interchromosomal rearrangements which are otherwise frequent especially in the Southern and Coastal clade (Krysanov and Demidova 2018). Nevertheless, the karyotype changes in the *N. ugandensis* species group are relatively frequent given the NF range but restricted to intrachromosomal changes only. Therefore, another hypothesis worthy of future experimental testing is whether or not any constraints related to the 3D nuclear genome architecture in species belonging to this lineage are responsible for a dramatic decrease in the probability of emergence and fixation of interchromosomal rearrangements. Intriguing examples of the interplay between chromosome rearrangements and nuclear architecture have been recently reported (Vara et al. 2021; Sidiropoulos et al. 2022; Wang et al. 2022).

Chromosome inversions are known to suppress recombination in the rearranged region but only in the heterozygous constitution (Sturtevant and Beadle 1936; Stevinson et al. 2011). While inversion heterozygotes might represent a transient populational polymorphism (King 1993), they may also be maintained by balancing selection between populations with gene flow as inversion can lock together a set of alleles of adjacent genes which may confer selective advantage for local adaptation or the evolution of complex life-history traits (Hoffmann and Rieseberg 2008; Wellenreuther and Bernathez 2018). Such cases have been reported in an increasing number of teleost species (Kirubakaran et al. 2016; Arostegui et al. 2019; Pearse et al. 2019; Wilder et al. 2020; Petrou et al. 2021). Neither our sampling, nor that of Krysanov and Demidova (2018), included individuals polymorphic for cytologically detectable inversion(s); therefore, we do not suspect that inversions might have adaptive effects in our studied system. Moreover, we found no correspondence between the karyotype variation and phylogenetic relationships (Fig. 4), nor with biogeographic distribution (Fig. 2) (discussed further below). Our data suggests independent parallel processes of karyotype differentiation within the *N. ugandensis* species group where the inversions were fixed mainly by other (e.g., neutral) processes than by natural selection. Therefore, given the structuring into small, isolated populations, the most reasonable explanation for the fixation of inversions in members of the *N. ugandensis* species group might be via random genetic drift including bottlenecks and founder effects (King 1993; Hoffmann and Rieseberg 2008; Connallon et al. 2018). The latter is consistent with the ability of killifishes to disperse and colonize new sites during major floods during the rainy season (van der Merwe et al. 2021). While the possible contribution of natural selection needs to be tested, inversions could contribute to reproductive isolation between conspecific populations and closely related species by various mechanisms (King 1993; Said et al. 2018; Villoutreix et al. 2020). The reproductive isolation might be triggered also by centromere repositioning (Lu and He 2019) which is another possible mechanism that could contribute to centromeric shifts observed in our studied species.

The *N. ugandensis* species group was recovered as monophyletic in Nagy et al. (2020) and van der Merwe et al. (2021). The topology of our phylogeny presented herein, for the purpose of comparing phylogenetic relationships with karyotype differentiation (Fig. 4), is congruent with previous results in the above-mentioned analyses. Within this group, well-defined clades, comprising the following species assemblages, exhibit strong branch support: *N. nubaensis* from the northern part of the distribution of the species group in Sudan and Ethiopia with *N. albertinensis* and *N. ugandensis* from Uganda, from the upper Nile drainage; *N. attenboroughi*, *N. derhami*, *N. moameensis* and *N. venustus* from systems associated with the near-shore zones of the Lake Victoria basin in south-western Kenya and north-western Tanzania; *N. hoermanni*, *N. torgashevi* and *N. streltsovi* from central Tanzania; and *N. itigiensis* and *N. kardashevi* from central and south-western Tanzania (Nagy et al. 2020; this study).

The biogeographic relationships among members of the *N. ugandensis* species group in central Tanzania can be explained by Palaeo-Lake Manonga, when rifting at the end of the Miocene led to ponding of the east-west rivers in northern Tanzania,

forming the shallow lake basin (Nagy et al. 2020). Palaeo-Lake Manonga would have provided a connection with the Malagarasi system in western Tanzania and the currently endorheic lesser systems in central Tanzania (Harrison and Mbago 1997; Van Damme and Pickford 2003). Further, members of three species groups are distributed along a south-north axis from southern and central Tanzania to the Lake Victoria drainage in northern Tanzania. Within the *N. ugandensis* species group, the presence of *N. ugandensis* in Uganda, with *N. nubaensis* in a basal phylogenetic position in that clade, suggests an ancestral dispersal northward through Uganda and further along the Nile drainage, as the latter species is currently known from southern Sudan and southwestern Ethiopia. Other species groups on the inland plateau of eastern Africa show striking similarities in distribution patterns and phylogenetic relationships, namely around an ancient Lake Manonga basin, along an east-west axis in central Tanzania, as well as northwards dispersal.

In previous studies of *Nothobranchius* (Nagy et al. 2016, 2017, 2020; Watters et al. 2019, 2020; van der Merwe et al. 2021) it was proposed that geomorphological changes separated drainages and thereby isolated populations that then speciated in peripatry and allopatry, evolving into distinct species, resulting in numerous local endemics. Rapid generation turnover in relatively small populations of these strictly seasonal fishes would have accelerated the effect of genetic drift, while during the episodes of aridity the collapse of populations may have led to population bottlenecks (Nagy et al. 2020).

Acknowledgements

References

Arai R (2011) Fish karyotypes – a Check List. Springer, 340 pp. https://doi.org/10.1007/978-4-431-53877-6

Arostegui MC, Quinn TP, Seeb LW, Seeb JE, McKinney GJ (2019) Retention of a chromosomal inversion from an anadromous ancestor provides the genetic basis for alternative freshwater ecotypes in rainbow trout. Molecular Ecology 28: 1412–1427. https://doi.org/10.1111/mec.15037

Bartáková V, Reichard M, Janko K, Polačik M, Blažek R, Reichwald K, Cellerino A, Bryja J (2013) Strong population genetic structuring in an annual fish, *Nothobranchius furzeri*, suggests multiple savannah refugia in southern Mozambique. BMC Ecology and Evolution 13: e196. https://doi.org/10.1186/1471-2148-13-196

Bartáková V, Reichard M, Blažek R, Polačik M, Bryja J (2015) Terrestrial fishes: rivers are barriers to gene flow in annual fishes from the African savanna. Journal of Biogeography 42: 1832–1844. https://doi.org/10.1111/jbi.12567

Brown RE, Freudenreich CH (2021) Structure-forming repeats and their impact on genome stability. Current Opinion in Genetics & Development 67: 41–51. https://doi.org/10.1016/j.gde.2020.10.006

Cellerino A, Valenzano DR, Reichard M (2016) From the bush to the bench: the annual *Nothobranchius* fishes as a new model system in biology. Biological Reviews 91: 511–533. https://doi.org/10.1111/brv.12183

Connallon T, Olito C, Dutoit L, Papoli H, Ruzicka F, Yong L (2018) Local adaptation and the evolution of inversions on sex chromosomes and autosomes. Philosophical Transactions of the Royal Society B 373: e20170423. https://doi.org/10.1098/rstb.2017.0423

Cui R, Medeiros T, Willemsen D, Iasi LNM, Collier GE, Graef M, Reichard M, Valenzano DR (2019) Relaxed selection limits lifespan by increasing mutation load. Cell 178: 385–399. [e20] https://doi.org/10.1016/j.cell.2019.06.004

Dorn A, Musilova Z, Platzer M, Reichwald K, Cellerino A (2014) The strange case of East African annual fish: aridification correlates with diversification for a savannah aquatic group? BMC Evolutionary Biology 14: e210. https://doi.org/10.1186/s12862-014-0210-3

Ewulonu UV, Haas R, Turner BJ (1985) A multiple sex chromosome system in the annual killifish, *Nothobranchius guentheri*. Copeia 2: 503–508. https://doi.org/10.2307/1444868

Harrison T, Mbago ML (1997) Introduction: paleontological and geological research in the Manonga Valley, Tanzania. In: Harrison T (Ed.) Neogene Paleontology of the Manonga Valley, Tanzania: a window into the evolutionary history of East Africa. Springer, Boston, 32 pp. https://doi.org/10.1007/978-1-4757-2683-1_1

Hoffmann AA, Rieseberg LH (2008) Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? Annual Review of Ecology, Evolution, and Systematics 39: 21–42. https://doi.org/10.1146/annurev.ecolsys.39.110707.173532

Jubb RA (1981) *Nothobranchius*. T.F.H., Neptune City, 61 pp.

King M (1993) Species Evolution: The Role of Chromosomal Change. Cambridge University Press, Cambridge.

Kirubakaran TG, Grove H, Kent MP, Sandve SR, Baranski M, Nome T, De Rosa MC, Righino B, Johansen T, Otterå H, Sonesson A, Lien S, Andersen Ø (2016) Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. Molecular Ecology 25: 2130–2143. https://doi.org/10.1111/mec.13592

Kligerman AD, Bloom SE (1977) Rapid chromosome preparations from solid tissues of fishes. Journal of the Fisheries Research Board of Canada 34: 266–269. https://doi.org/10.1139/f77-039

Krysanov (1992) Aneuploidy in postnatal ontogenesis of fishes. Acta Zoologica Fennica 191: 177–182.

Krysanov E, Demidova T (2018) Extensive karyotype variability of African fish genus *Nothobranchius* (Cyprinodontiformes). Comparative Cytogenetics 12(3): 387–402. https://doi.org/10.3897/CompCytogen.v12i3.25092

Krysanov E, Demidova T, Nagy B (2016) Divergent karyotypes of the annual killifish genus *Nothobranchius* (Cyprinodontiformes, Nothobranchiidae). Comparative Cytogenetics 10(3): 439–445. https://doi.org/10.3897/CompCytogen.v10i3.9863

Krysanov EYu, Ordzhonikidze KG, Simanovsky SA (2018) Cytogenetic indicators in estimation of environmental state. Russian Journal of Developmental Biology 49: 36–41. https://doi.org/10.1134/S1062360418010034

Levan A, Fredga K, Sandberg A (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201–220. https://doi.org/10.1111/j.1601-5223.1964.tb01953.x

Li S-F, Su T, Cheng G-Q, Wang B-X, Li X, Deng C-L, Gao W-J (2017) Chromosome evolution in connection with repetitive sequences and epigenetics in plants. Genes 8: e290. https://doi.org/10.3390/genes8100290

Lu M, He X (2019) Centromere repositioning causes inversion of meiosis and generates a reproductive barrier. Proceedings of the National Academy of Sciences 116(43): 21580–21591. https://doi.org/10.1073/pnas.1911745116

Motta-Neto CCd, Cioffi MdB, Costa GWWFd, Amorim KDJ, Bertollo LAC, Artoni RF, Molina WF (2019) Overview on karyotype stasis in Atlantic grunts (Eupercaria, Haemulidae) and the evolutionary extensions for other marine fish groups. Frontiers in Marine Science 6: e628. https://doi.org/10.3389/fmars.2019.00628

Nagy B (2015) Life history and reproduction of *Nothobranchius* fishes. Journal of the American Killifish Association 47(4–6): 182–192. https://doi.org/10.5281/zenodo.4393325

Nagy B (2018) *Nothobranchius ditte*, a new species of annual killifish from the Lake Mweru basin in the Democratic Republic of the Congo (Teleostei: Nothobranchiidae). Ichthyological Exploration of Freshwaters 28(2): 115–134.

Nagy B, Cotterill FPD, Bellstedt DU (2016) *Nothobranchius sainthousei*, a new species of annual killifish from the Luapula River drainage in northern Zambia (Teleostei: Cyprinodontiformes). Ichthyological Exploration of Freshwaters 27(3): 233–254. https://doi.org/10.2989/16085914.2017.1372270

Nagy B, Watters BR (2021) A review of the conservation status of seasonal *Nothobranchius* fishes (Teleostei: Cyprinodontiformes), a genus with a high level of threat, inhabiting ephemeral wetland habitats in Africa. Aquatic Conservation: Marine and Freshwater Ecosystems 32(1): 199–216. https://doi.org/10.1002/aqc.3741

Nagy B, Watters BR, van der Merwe PDW, Cotterill FPD, Bellstedt DU (2017) *Nothobranchius cooperi* (Teleostei: Cyprinodontiformes): a new species of annual killifish from the Luapula River drainage, northern Zambia. African Journal of Aquatic Science 42(3): 201–218. https://doi.org/10.2989/16085914.2017.1372270

Nagy B, Watters BR, van der Merwe PDW, Cotterill FPD, Bellstedt DU (2020) Review of the *Nothobranchius ugandensis* species group from the inland plateau of eastern Africa with descriptions of six new species (Teleostei: Nothobranchiidae). Ichthyological Exploration of Freshwaters, 30(1): 21–73. https://doi.org/10.23788/IEF-1129

Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadía-Cardoso A, Anderson EC, Rundio DE, Williams TH, Naish KN, Moen T, Liu S, Kent M, Moser M, Minkley DR, Rondeau EB, Brieuc MSO, Sandve SR, Miller MR, Cedillo L, Baruch K, Hernandez AG, Ben-Zvi G, Shem-Tov D, Barad O, Kuzishchin K, Garza JC, Lindley ST, Koop BF,

Thorgaard GH, Palti Y (2019) Sex-dependent dominance maintains migration supergene in rainbow trout. Nature Ecology & Evolution 3: 1731–1742. https://doi.org/10.1038/s41559-019-1044-6

Petrou EL, Fuentes-Pardo AP, Rogers LA, Orobko M, Tarpey C, Jiménez-Hidalgo I, Moss ML, Yang D, Pitcher TJ, Sandell T, Lowry D, Ruzzante DE, Hauser L (2021) Functional genetic diversity in an exploited marine species and its relevance to fisheries management. Proceedings of the Royal Society B: Biological Sciences 288: e20202398. https://doi.org/10.1098/rspb.2020.2398

Rambaut A, Drummond AJ (2012) FigTree version 1.4. https://github.com/rambaut/figtree

Redi CA, Garagna S, Della Valle G, Bottiroli G, Dell´Orto P, Viale G, Peverali FA, Raimondi E, Forejt J (1990) Differences in the organization and chromosomal allocation of satellite DNA between the European long-tailed house mice *Mus domesticus* and *Mus musculus*. Chromosoma 99: 11–17. https://doi.org/10.1007/BF01737284

Reichwald K, Lauber C, Nanda I, Kirschner J, Hartmann N, Schories S, Gausmann U, Taudien S, Schilhabel MB, Szafranski K, Glöckner G, Schmid M, Cellerino A, Schartl M, Englert C, Platzer M (2009) High tandem repeat content in the genome of the short-lived annual fish *Nothobranchius furzeri*: a new vertebrate model for aging research. Genome Biology 10: R16. https://doi.org/10.1186/gb-2009-10-2-r16

Reichwald K, Petzold A, Koch P, Downie BR, Hartmann N, Pietsch S, Baumgart M, Chalopin D, Felder M, Bens M, Sahm A, Szafranski K, Taudien S, Groth M, Arisi I, Weise A, Bhatt SS, Sharma V, Kraus JM, Schmid F, Priebe S, Liehr T, Görlach M, Than ME, Hiller M, Kestler HA, Volff JN, Schartl M, Cellerino A, Englert C, Platzer M (2015) Insights into sex chromosome evolution and aging from the genome of a short-lived fish. Cell 163: 1527–1538. https://doi.org/10.1016/j.cell.2015.10.071

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19(12): 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Said I, Byrne A, Serrano V, Cardeno C, Vollmers C, Corbett-Detig R (2018) Linked genetic variation and not genome structure causes widespread differential expression associated with chromosomal inversions. Proceedings of the National Academy of Sciences 115(21): 5492–5497. https://doi.org/10.1073/pnas.1721275115

Scheel JJ (1990) Atlas of killifishes of the Old World. T.F.H. Publications, Neptune, 448 pp.

Schubert I (2018) What is behind "centromere repositioning"? Chromosoma 127: 229–234. https://doi.org/10.1007/s00412-018-0672-y

Sidiropoulos N, Mardin BR, Rodríguez-González FG, Bochkov ID, Garg S, Stütz AM, Korbel JO, Aiden EL, Weischenfeldt J (2022) Somatic structural variant formation is guided by and influences genome architecture. Genome Research 32(4): 643–655. https://doi.org/10.1101/gr.275790.121

Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology 7: e539. https://doi.org/10.1038/msb.2011.75

Simanovsky SA, Demidova TB, Spangenberg VE, Matveevsky SN, Ordzhonikidze KG, Kolomiets OL, Krysanov EY (2019) Multiple sex chromosome system X_1X_2Y in African

killifish genera *Nothobranchius* and *Fundulosoma* (Cyprinodontiformes). Frontiers in Marine Science. Conference Abstract: XVI European Congress of Ichthyology. https://doi.org/10.3389/conf.fmars.2019.07.00158

Stevison LS, Hoehn KB, Noor MAF (2011) Effects of inversions on within- and between-species recombination and divergence. Genome Biology and Evolution 3: 830–841. https://doi.org/10.1093/gbe/evr081

Štundlová J, Hospodářská M, Lukšíková K, Voleníková A, Pavlica T, Altmanová M, Richter A, Reichard M, Dalíková M, Pelikánová Š, Marta A, Simanovsky SA, Hiřman M, Jankásek M, Dvořák T, Bohlen J, Ráb P, Englert C, Nguyen P, Sember A (2022) Sex chromosome differentiation via changes in the Y chromosome repeat landscape in African annual killifishes *Nothobranchius furzeri* and *N. kadleci*. Chromosome Research 30: 309–333. https://doi.org/10.1007/s10577-022-09707-3

Sturtevant AH, Beadle GW (1936) The relations of inversions in the X chromosome of *Drosophila melanogaster* to crossing over and disjunction. Genetics 21: 554–604. https://doi.org/10.1093/genetics/21.5.554

Van Damme D, Pickford M (2003) The late Cenozoic Thiaridae (Mollusca, Gastropoda, Cerithioidea) of the Albertine Rift valley (Uganda-Congo) and their bearing on the origin and evolution of the Tanganyikan thalassoid malacofauna. Hydrobiologia 498: 1–83. https://doi.org/10.1023/A:1026298512117

van der Hoeven JC, Bruggeman IM, Alink GM, Koeman JH (1982) The killifish *Nothobranchius rachowi*, a new animal in genetic toxicology. Mutation Research 97: 35–42. ttps://doi.org/10.1016/0165-1161(82)90017-6

van der Merwe PDW, Cotterill FPD, Kandziora M, Watters BR, Nagy B, Genade T, Flügel TJ, Svendsen DS, Bellstedt DU (2021) Genomic fingerprints of palaeogeographic history: the tempo and mode of rift tectonics across tropical Africa has shaped the diversification of the killifish genus *Nothobranchius* (Teleostei: Cyprinodontiformes), Molecular Phylogenetics and Evolution 158: e106988. https://doi.org/10.1016/j.ympev.2020.106988

Vanderplank FL (1940) A study of *Nothobranchius taeniopygus* in its natural habitat. The Aquarist 10: 247–249.

Vara C, Paytuví-Gallart A, Cuartero Y, Álvarez-González L, Marín-Gual L, Garcia F, Florit-Sabater B, Capilla L, Sanchéz-Guillén RA, Sarrate Z, Aiese Cigliano R, Sanseverino W, Searle JB, Ventura J, Marti-Renom MA, Le Dily F, Ruiz-Herrera A (2021) The impact of chromosomal fusions on 3D genome folding and recombination in the germ line. Nature Communications 12(1): 1–17. https://doi.org/10.1038/s41467-021-23270-1

Villoutreix R, Ayala D, Joron M, Gompert Z, Feder JL, Nosil P (2021) Inversion breakpoints and the evolution of supergenes. Molecular Ecology 30: 2738–2755. https://doi.org/10.1111/mec.15907

Wang LB, Li ZK, Wang LY, Xu K, Ji TT, Mao YH, Ma SN, Liu T, Tu CF, Zhao Q, Fan XN, Liu C, Wang LY, Shu YJ, Yang N, Zhou Q, Li W (2022) A sustainable mouse karyotype created by programmed chromosome fusion. Science 377(6609): 967–975. https://doi.org/10.1126/science.abm1964

Watters BR (2009) The ecology and distribution of *Nothobranchius* fishes. Journal of the American Killifish Association 42(2): 37–76.

Watters BR, Nagy B, van der Merwe PDW, Cotterill FPD, Bellstedt DU (2019) Review of the *Nothobranchius taeniopygus* species group from central and western Tanzania with descriptions of five new species and redescription of *Nothobranchius taeniopygus* (Teleostei: Nothobranchiidae). Ichthyological Exploration of Freshwaters 29(3): 239–278. https://doi.org/10.23788/IEF-1110

Watters BR, Nagy B, van der Merwe PDW, Cotterill FPD, Bellstedt DU (2020) Redescription of the seasonal killifish species *Nothobranchius ocellatus* and description of a related new species *Nothobranchius matanduensis*, from eastern Tanzania (Teleostei: Nothobranchiidae). Ichthyological Exploration of Freshwaters 30(2): 151–178. https://doi.org/10.23788/IEF-1149

Wellenreuther M, Bernatchez L (2018) Eco-evolutionary genomics of chromosomal inversions. Trends in Ecology and Evolution 33(6): 427–440. https://doi.org/10.1016/j.tree.2018.04.002

Wildekamp RH (2004) A world of killies. Atlas of the oviparous cyprinodontiform fishes of the world (Vol. IV). American Killifish Association, Elyria, 398 pp.

Wildekamp RH, Watters BR, Shidlovskiy KM (2014) Review of the *Nothobranchius neumanni* species group with descriptions of three new species from Tanzania (Cyprinodontiformes: Nothobranchiidae). Journal of the American Killifish Association 47(1): 2–30.

Wilder AP, Palumbi SR, Conover DO, Therkildsen NO (2020) Footprints of local adaptation span hundreds of linked genes in the Atlantic silverside genome. Evolution Letters 4: 430–443. https://doi.org/10.1002/evl3.189

ORCID

Eugene Yu. Krysanov https://orcid.org/0000-0001-7916-4195
Béla Nagy https://orcid.org/0000-0003-4718-0822
Brian R. Watters https://orcid.org/0000-0002-7651-6500
Alexandr Sember https://orcid.org/0000-0003-4441-9615
Sergey A. Simanovsky https://orcid.org/0000-0002-0830-7977

RESEARCH ARTICLE

Comparative karyotype analysis of eight Cucurbitaceae crops using fluorochrome banding and 45S rDNA-FISH

Chao-Wen She[1,2], Xiang-Hui Jiang[1], Chun-Ping He[2]

1 *Key Laboratory of Research and Utilization of Ethnomedicinal Plant Resources of Hunan Province, Huaihua University, Huaihua, Hunan, 418008, China* **2** *College of Life Sciences and Chemistry, Hunan University of Technology, Zhuzhou, Hunan, 412007, China*

Corresponding author: Chao-Wen She (shechaowen@aliyun.com)

Citation: She C-W, Jiang X-H, He C-P (2023) Comparative karyotype analysis of eight Cucurbitaceae crops using fluorochrome banding and 45S rDNA-FISH. Comparative Cytogenetics 17(1): 31–58. https://doi.org/10.3897/compcytogen.v17.i1.99236

Abstract

To have an insight into the karyotype variation of eight Cucurbitaceae crops including *Cucumis sativus* Linnaeus, 1753, *Cucumis melo* Linnaeus, 1753, *Citrullus lanatus* (Thunberg, 1794) Matsumura et Nakai, 1916, *Benincasa hispida* (Thunberg, 1784) Cogniaux, 1881, *Momordica charantia* Linnaeus, 1753, *Luffa cylindrica* (Linnaeus, 1753) Roemer, 1846, *Lagenaria siceraria* var. *hispida* (Thunberg, 1783) Hara, 1948 and *Cucurbita moschata* Duchesne ex Poiret, 1819, well morphologically differentiated mitotic metaphase chromosomes were prepared using the enzymatic maceration and flame-drying method, and the chromosomal distribution of heterochromatin and 18S-5.8S-26S rRNA genes (45S rDNA) was investigated using sequential combined PI and DAPI (CPD) staining and fluorescence *in situ* hybridization (FISH) with 45S rDNA probe. Detailed karyotypes were established using the dataset of chromosome measurements, fluorochrome bands and rDNA FISH signals. Four karyotype asymmetry indices, CV_{CI}, CV_{CL}, M_{CA} and Stebbins' category, were measured to elucidate the karyological relationships among species. All the species studied had symmetrical karyotypes composed of metacentric and submetacentric or only metacentric chromosomes, but their karyotype structure can be discriminated by the scatter plot of M_{CA} vs. CV_{CL}. The karyological relationships among these species revealed by PCoA based on *x*, *2n*, TCL, M_{CA}, CV_{CL} and CV_{CI} was basically in agreement with the phylogenetic relationships revealed by DNA sequences. CPD staining revealed all 45S rDNA sites in all species, (peri)centromeric GC-rich heterochromatin in *C. sativus*, *C. melo*, *C. lanatus*, *M. charantia* and *L. cylindrica*, terminal GC-rich heterochromatin in *C. sativus*. DAPI counterstaining after FISH revealed pericentromeric DAPI⁺ heterochromatin in *C. moschata*. rDNA FISH detected two 45S loci in five species and five 45S loci in three species. Among

these 45S loci, most were located at the terminals of chromosome arms, and a few in the proximal regions. In *C. sativus*, individual chromosomes can be precisely distinguished by the CPD band and 45S rDNA signal patterns, providing an easy method for chromosome identification of cucumber. The genome differentiation among these species was discussed in terms of genome size, heterochromatin, 45S rDNA site, and karyotype asymmetry based on the data of this study and previous reports.

Keywords

Cucurbitaceae, cytotaxonomy, fluorescence *in situ* hybridization, fluorochrome banding, karyotype, karyotype asymmetry, ribosomal RNA genes (rDNA)

Introduction

Cucurbitaceae, which is among the economically most important plant families, consists of about 123 genera with over 800 species distributed most in tropical and subtropical areas and very rare in temperate regions (Jeffrey 2005). Cucurbitaceous species (cucurbits) have a large range of fruit characteristics, and are cultivated worldwide in a variety of environmental conditions (Bisognin 2002). Among the cultivars of this family, cucumber (*Cucumis sativus* Linnaeus, 1753), melon (*Cucumis melo* Linnaeus, 1753), watermelon (*Citrullus lanatus* (Thunberg, 1794) Matsumura et Naka, 1916), wax gourd (*Benincasa hispida* (Thunberg, 1784) Cogniaux, 1881), bitter gourd (*Momordica charantia* Linnaeus, 1753), sponge gourd (*Luffa cylindrica* (Linnaeus, 1753) Roemer, 1846), bottle gourd (*Lagenaria siceraria* (Molina) Standley, 1930), squash and pumpkin (*Cucurbita* Linnaeus, 1753), all of which belong to subfamily Cucurbitoideae (Jeffrey 2005), are grown as vegetable crops with global or local economic importance, providing human with edible and medicinal fruits (Bisognin 2002).

In higher plants, karyotype analysis has been used to characterize the genome at chromosome level, to elucidate cytotaxonomic relationships among taxa, to reveal the genetic aberrations, to understand the trends in chromosome evolution, to integrate genetic and physical maps (Moscone et al. 1999; Peruzzi et al. 2009, 2017; Han et al. 2011; Guerra 2012; Siljak-Yakovlev and Peruzzi 2012; She and Jiang 2015; She et al. 2015, 2017, 2020; Astuti et al. 2017; Kadluczka and Grzebelus 2021). In general, a description of the karyotype includes the chromosome number, the absolute and relative length of chromosomes, the position of primary and secondary constrictions, the distribution of heterochromatic segments, the number and position of rDNA sites and other DNA sequences, and the degree of asymmetry (Li and Chen 1985; Levin 2002; She and Jiang 2015; She et al. 2015, 2017, 2020). Among the karyotypic parameters, the karyotype asymmetry, which is determined by the variation in chromosome length (interchromosomal asymmetry) and the variation in centromere position (intrachromosomal asymmetry) in a chromosome complement, is an important karyotype character reflecting the general morphology of chromosomes, and is thus widely used in plant cytotaxonomy (Stebbins 1971; Paszko 2006; Peruzzi et al. 2009, 2017; Peruzzi and Eroglu 2013; Astuti et al. 2017; Dehery et al. 2020; Kadluczka and Grzebelus 2021).

In most cases, karyotyping is hampered by the lack of chromosome markers, which limits the identification of individual chromosomes. To overcome this obstacle, Giemsa and fluorochrome banding techniques as well as fluorescence *in situ* hybridization (FISH) technologies were successively applied in plant chromosome analysis. Double fluorochrome staining, such as CMA (chromomycin A3)/ DAPI (4,6-diamidino-2-phenylindole) staining, and PI (propidium iodide)/ DAPI staining (called CPD staining) were employed to reveal simultaneously GC-rich and AT-rich heterochromatic regions on chromosomes (Schweizer 1976; She et al. 2006, 2015). FISH with repetitive DNA sequences as well as large-insert genomic DNA clones on mitotic metaphase or pachytene chromosomes can generate specific signal pattern in a plant species (Moscone et al. 1999; Hasterok et al. 2001; Koo et al. 2010; Liu et al. 2010; She and Jiang 2015; She et al. 2015, 2017, 2020). Both fluorochrome bands and FISH signals are effective markers for chromosome identification. Using the combined data of chromosome measurements, fluorochrome bands and FISH signals, we can construct detailed molecular cytogenetic karyotype of a plant species that displays morphological characteristics of chromosomes, distribution of heterochromatin and locations of DNA sequences (de Moraes et al. 2007; She and Jiang 2015; She et al. 2015, 2017, 2020). Comparison of karyotypes taking advantage of molecular cytogenetics can provide valuable information on the phylogenetic relationships and chromosome evolution among related species (Moscone et al. 1999; de Moraes et al. 2007; Weiss-Schneeweiss et al. 2008; Siljak-Yakovlev and Peruzzi 2012; She et al. 2015, 2017, 2020).

Cytogenetic studies of cucurbits started in 1920s. Earlier cytogenetic studies restricted to chromosome counting to determine the basic chromosome numbers of this family, as well as karyomorphological descriptions of some species, mainly focused on *Cucumis* Linnaeus, 1753 and *Citrullus* Schrader, 1836 (Bhaduri and Bose 1947; Trivedi and Roy 1970; Singh and Roy 1974; Turkov et al. 1975; Dane and Tsuchiya 1976; Ramachandran and Seshadri 1986; Li 1989; Beevy and Kuriachan 1996). The family was found to have several basic numbers such as x = 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, and 20, of which x = 11 is the ancestral number (Carta et al. 2020). Cytogenetic observations also revealed that, except for a species of *Benincasa* Savi, 1818, the mitotic chromosomes of all other cucurbits investigated so far were rather small in size and similar in morphology, resulting in the difficulty of chromosome identification using conventional cytological procedures (Bhaduri and Bose 1947; Trivedi and Roy 1970; Singh and Roy 1974; Li 1989). C-banding technique and CMA/DAPI staining were employed for the characterization of cucumber chromosomes, revealing that individual chromosomes could be distinguished by the C- or fluorochrome banding patterns (Chen et al. 1998; Hoshi et al. 1998; Plader et al. 1998). However, the C- and fluorochrome banding techniques have rarely been successfully applied in other cucurbits till now. In recent two decades, FISH technologies have been employed in the chromosome analysis of more than 60 Cucurbitaceous species. FISH with repetitive DNA sequences, fosmid or BAC (artificial bacterial chromosome) clones, and bulked oligonucleotides probes on mitotic metaphase or pachytene chromosomes were used for karyotyping (Koo et al. 2002, 2005; Li et al. 2007; Xu et al. 2007; Han et al. 2008; Liu et al. 2010; Waminal et al. 2011; Waminal and Kim 2012, 2015; Zhang et al. 2015a, 2015b; Pellerin et al. 2018a, 2018b; Xie et al. 2019b), comparative cytogenetic analysis (Han et al. 2009; Koo et al. 2010; Zhao et al. 2011; Yagi et al. 2015;

Zhang et al. 2015a; Li et al. 2016; Zhang et al. 2016), construction of cytogenetic map (Ren et al. 2009, 2012; Han et al. 2011; Sun et al. 2013), and chromosome-specific painting (Han et al. 2015). FISH experiments with 45S rDNA alone or both 5S and 45S rDNA as probes have been performed in a lot of cultivated and wild cucurbits including the eight cultivated species investigated herein (Chen et al. 1999; Hoshi et al. 1999; Koo et al. 2002; Li et al. 2007; Xu et al. 2007; Han et al. 2008; Liu et al. 2010; Waminal et al. 2011; Zhao et al. 2011; Waminal and Kim 2012, 2015; Guo et al. 2013; Reddy et al. 2013; Yagi et al. 2015; Zhang et al. 2015b; Li et al. 2016; Zhang et al. 2016; Pellerin et al. 2018a, 2018b; Xie et al. 2019b). In cucumber, the FISH signals of both 45S rDNA and centromeric satellite Type III allow for unequivocal identification of all mitotic metaphase chromosomes (Han et al. 2008). Also, the 45S rDNA FISH and self-GISH signal patterns enabled individual chromosomes of cucumber to be characterized (Zhang et al. 2015b). However, in the other seven cultivated species involved in this study, the rDNA sites can only mark a minority of the chromosomes in the complement (Chen et al. 1999; Li et al. 2007; Xu et al. 2007; Waminal et al. 2011; Ren et al. 2012; Waminal and Kim 2012; Guo et al. 2013; Li et al. 2016; Reddy et al. 2013; Xie et al. 2019b). In melon, a combination of CentM, 45S rDNA, and 5S rDNA with 21 fosmids of cucumber enabled each of the 12 chromosome pairs to be identified (Liu et al. 2010). As a whole, the karyotypes of cucumber and melon have been adequately investigated using molecular cytogenetic methods, while those of the other six species involved in this study have not been well molecular-cytogenetically studied. The karyotype of cucumber has been standardized (Han et al. 2008), but the karyotype data of the other seven species were incomplete, and showed inconsistency among the previous reports (Li 1989; Li et al. 2007; Xu et al. 2007; Liu et al. 2010; Waminal et al. 2011; Waminal and Kim 2012; Guo et al. 2013). Further cytogenetic investigations are needed for establishment of detailed karyotypes of the eight Cucurbitaceae crops to elucidate the genome differentiation at chromosome-level.

In the current study, using the enzymatic maceration and flame-drying (EMF) method, well morphologically differentiated mitotic metaphase chromosomes of the eight Cucurbitaceae crops were prepared. The chromosomes were characterized by sequential CPD staining and FISH with 45S rDNA probe. Detailed karyotypes of these species were quantitatively constructed using the combined data of chromosome measurements, fluorochrome bands and 45S rDNA FISH signals. Four different karyotype asymmetry indices of each species were calculated for evaluating the karyological relationships among these species. The molecular cytogenetic karyotypic data were assessed to gain insights into the genome differentiation and evolutionary relationships among the eight species.

Material and methods

Plant material

The seeds of *Cucumis sativus* Linnaeus, 1753, *Cucumis melo* Linnaeus, 1753, *Citrullus lanatus* (Thunberg, 1794) Matsumura et Nakai, 1916, *Benincasa hispida* (Thunberg, 1784) Cogniaux, 1881, *Momordica charantia* Linnaeus, 1753, *Luffa cylindrica*

(Linnaeus, 1753) Roemer, 1846, *Lagenaria siceraria* var. *hispida* (Thunberg, 1783) Hara, 1948 and *Cucurbita moschata* Duchesne ex Poiret, 1819 were obtained from commercial seed companies in China. Cultivar accessions used in this study are described in Suppl. material 1.

Chromosome preparation

The seeds were germinated on moist filter paper in Petri dishes at 28 °C in the dark. Actively growing root tips were excised and treated in saturated α-bromonaphthalene at 28 °C for 1.0 h, and then fixed in a freshly prepared mixture of methanol and glacial acetic acid (3:1, v/v) at 4 °C, overnight. Mitotic metaphase chromosome spreads were prepared from meristem root tip cells according to She et al. (2006). The fixed root tips (2–3 mm) were thoroughly washed in double distilled water and digested in an enzyme mixture of 1% cellulase RS (Yakult Pharmaceutical Industry Co., Ltd. Tokyo, Japan) and 1% pectolyase Y-23 (Yakult Pharmaceutical Industry Co., Ltd. Tokyo, Japan) in citric buffer (0.01 mM citric acid-sodium citrate, pH 4.5) at 28 °C for 1.0–1.5 h. The digested root tips were washed by double distilled water and transferred to a glass slide, and then mashed thoroughly with the fixative by using fine-pointed forceps. Then, the slides were dried over the flame of an alcohol lamp. The slides with abundant division cells and well-spread metaphase chromosomes were selected using an Olympus BX51 phase contrast microscope, and then stored at -20 °C until use.

CPD staining

CPD staining was performed following the procedure indicated by She et al. (2006). In brief, the chromosome preparations were sequentially treated with RNase A and pepsin and then stained with a mixture of 0.6 $\mu g \cdot mL^{-1}$ PI and 3 $\mu g \cdot mL^{-1}$ DAPI in a 30% (v/v) solution of Vectashield H100 (Vector Laboratories, Burlingame, US) for more than 30 min. Chromosome spreads were observed using an Olympus BX60 epifluorescence microscope with UV and green excitation filters. Images were captured and merged using a cooled CCD camera (CoolSNAP EZ; Photometrics, Tucson, US) controlled by METAMORPH software (Molecular Devices, California, US).

Fluorescence *in situ* hybridization

The probe that was used to detect the 26S-5.8S-18S rRNA gene was a 9.04-kb 45S rDNA insert from tomato (see She et al. 2006), which was labeled with biotin-16-dUTP using Nick Translation Kit (Roche Diagnostics, Mannheim, Germany).

FISH with the 45S rDNA probe was conducted on the slides previously stained by CPD. The stained slides were washed in 2× SSC, twice for 15 min each, dehydrated in a graded ethanol series (70%, 90%, and 100%), air-dried at room temperature. Hybridization was performed as described by She et al. (2006). The biotin-labeled probe was detected by Fluorescein Avidin D (Vector Laboratories, Burlingame, USA). The chromosomes were counterstained and mounted with 3 $\mu g\ mL^{-1}$ DAPI in 30% (v/v)

solution of Vectashield H-1000, and observed using the epifluorescence microscope mentioned above. Images were captured digitally using METAMORPH software with UV and blue excitation filters for DAPI and fluorescein, respectively.

Karyotype analysis

Karyotype analysis followed the methodology as described by She et al. (2015). For each species, five metaphase cells whose chromosomes dispersed and condensed moderately (not reaching the maximum degree of condensation but not having decondensed terminals) were selected for measuring the length of long arm (L) and short arm (S) of each chromosome and the length of each fluorochrome band in a chromosome complement. Five metaphase cells with the maximum degree of condensation were used for measuring the absolute length of each chromosome. For the numeric characterization of the karyotypes the following parameters were calculated: (1) chromosome relative length (RL, % of haploid complement); (2) arm ratio (AR = L/S); (3) total chromosome length of the haploid complement (TCL; i.e. the karyotype length); (4) mean chromosome length (C); (5) size of the fluorochrome band (expressed as percentage of the karyotype length); (6) percent distance from the centromere to the rDNA site; (7) mean centromeric index (CI); (8) Four different karyotype asymmetry indices including coefficient of variation (CV) of centromeric index (CV_{CI}), coefficient of variation (CV) of chromosome length (CV_{CL}), mean centromeric asymmetry (M_{CA}) and Stebbins' asymmetry category. The meaning and calculation formulae of these asymmetry indices were given in Paszko (2006) and Peruzzi and Eroglu (2013). The arm ratio was used to classify chromosomes following the Levan's system (Li and Chen 1985). The chromosomes were arranged in order of decreasing length except those of *C. sativus* which were organized according to the chromosome nomenclature as described by Han et al. (2008). Idiograms were drawn quantitatively based on the dataset of chromosome measurements as well as the position and size of fluorochrome bands and rDNA-FISH signals.

To visualize karyotype asymmetry relationships among the eight species, bidimensional scatter diagrams for these species with M_{CA} vs. CV_{CL} were plotted. To determine the karyological relationships among the eight species, a principal coordinate analysis (PCoA) using Gower's similarity coefficient were performed based on six quantitative parameters (*x*, 2*n*, TCL, M_{CA}, CV_{CL}, CV_{CI}) according to the proposal by Peruzzi and Altınordu (2014). Statistical analyses were performed with Statistica for Windows 10.0, and PCoA scatter plot was generated.

Results

General karyotype features

Using the EMF method, dispersed and morphologically well differentiated mitotic metaphase chromosomes were obtained and used for karyotyping (Fig. 1). The metaphase chromosomes with the maximum condensation degree were not very
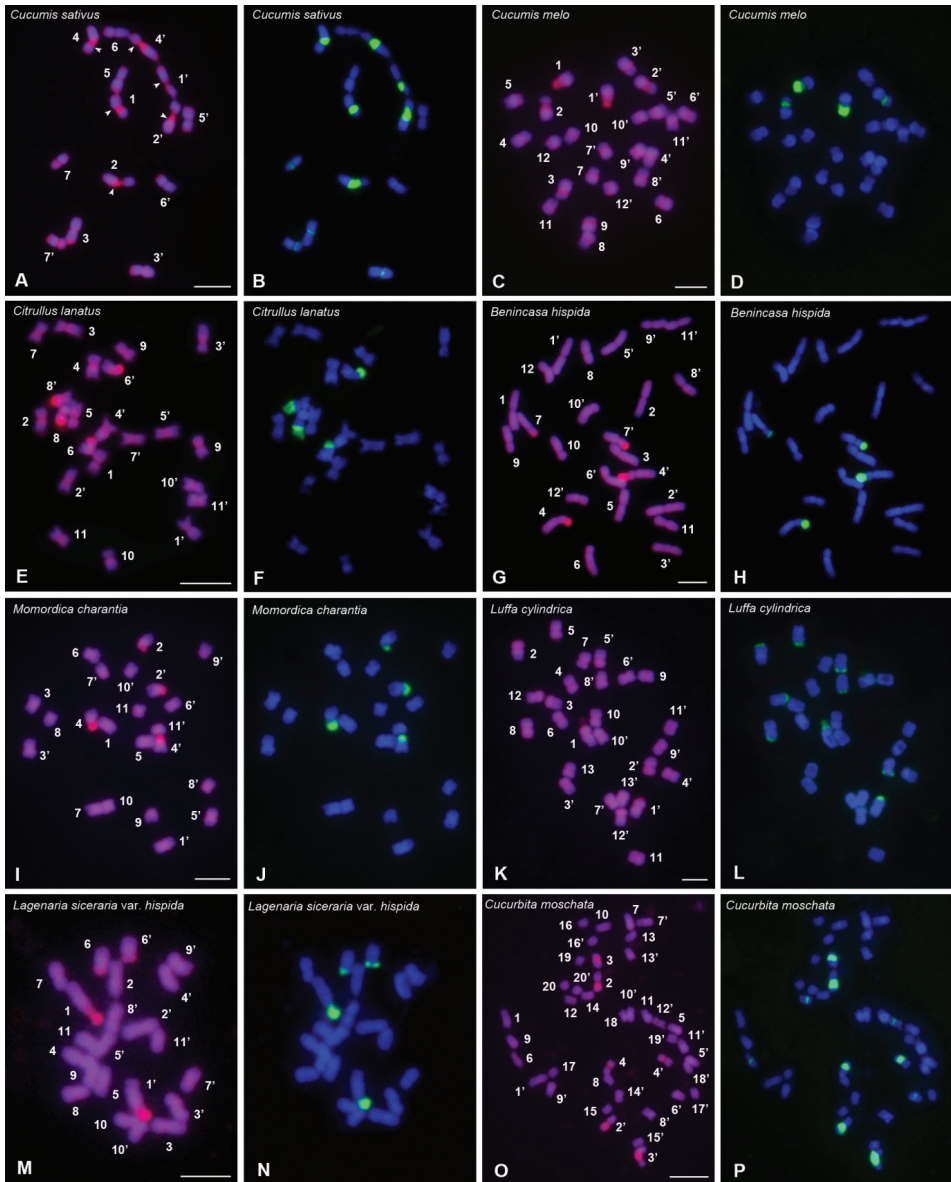
Figure 1. Mitotic chromosomes from *C. sativus* (**A,B**), *C. melo* (**C,D**), *C. lanatus* (**E,F**), *B. hispida* (**G,H**), *M. charantia* (**I,J**), *L. cylindrica* (**K,L**), *L. siceraria* var. *hispida* (**M,N**) and *C. moschata* (**O,P**) stained using CPD staining and sequential FISH with biotin-labelled 45S rDNA probe. **A, C, E, G, I, K, M** and **O** are the chromosomes stained using CPD. The chromosome numbers are designated by karyotyping. **B, D, F, H, J, L, N** and **P** are the chromosomes displaying 45S rDNA signals (green). The total DNA was counterstained using DAPI (blue). Scale bars: 10 μm.

appropriate for karyotyping because of the reduction of morphological discrimination, but were suitable for the measurement of TCLs because of the comparability of TCLs between species (Suppl. material 2). The detailed karyotype features and the nuclear

DNA contents of the eight species are summarized in Table 1. The measurement data of the chromosomes of each species are given in Suppl. material 3. The distribution of fluorochrome bands and 45S rDNA sites are presented in Table 2. Idiograms displaying the chromosome measurements, as well as the position and size of fluorochrome bands and 45S rDNA FISH signals are illustrated in Figure 2.

The diploid chromosome numbers are $2n = 2x = 14$ for *C. sativus*, $2n = 2x = 22$ for *C. lanatus*, *M. charantia* and *L. siceraria* var. *hispida*, $2n = 2x = 24$ for *C. melo* and *B. hispida*, $2n = 2x = 26$ for *L. cylindrica*, and $2n = 2x = 40$ for *C. moschata* (Table 1). According to the classification of Lima-de-Faria (1980), the metaphase chromosomes of *B. hispida* are of medium size with a mean chromosome length of 4.66 μm and a TCL of 55.93 μm, while those of the other seven species are of small size with a mean chromosome length between 1.91 μm (*C. moschata*) and 3.48 μm (*C. sativus*) and a TCL between 21.31 μm (*M. charantia*) to 38.15 μm (*C. moschata*). The TCLs of the eight species are basically in proportion to the nuclear DNA contents reported (Table 1). The smallest RRL (range of relative length) is observed in *C. sativus* (11.88~16.52), while the largest RRL is showed in *L. siceraria* var. *hispida* (6.67~14.52). That is, *C. sativus* and *L. siceraria* var. *hispida* exhibited the smallest and the largest variation in chromosome length, respectively. The mean centromeric index (CI) of the chromosome complements varied between 45.35 ± 2.73 (*L. cylindrica*) and 39.87 ± 6.37 (*C. melo*). That is, *L. cylindrica* and *C. melo* are characterized by the smallest and the largest level of variation in the centromeric index, respectively.

The karyotypes are composed of only metacentric (m) chromosomes (*L. cylindrica*) or metacentric and submetacentric (sm) chromosomes (the other seven species)

Table 1. Karyotype parameters of the eight Cucurbitaceae crops.

Species	KF	Genome size	TCL ± SE (μm)	C (μm)	RRL	CI±SE	CV_{CI}	M_{CA}	CV_{CL}	St
Cucumis sativus	2n = 14 = 12m (4SAT) + 2sm (2SAT)	367 Mb (Huang et al. 2009)	24.36 ± 1.47	3.48	11.88–16.52	44.56 ± 4.94	11.09	10.88	12.62	1A
Cucumis melo	2n = 24 = 16m (4SAT) + 8sm	450 Mb (Garcia-Mas et al. 2012)	33.34 ± 3.21	2.78	6.87–10.72	39.87 ± 6.37	15.99	20.27	14.51	2A
Citrullus lanatus	2n = 22 = 20m + 2sm	425 Mb (Guo et al. 2013)	24.94 ± 1.94	2.27	7.78–10.44	42.46 ± 3.41	8.03	15.08	10.92	1A
Benincasa hispida	2n = 24 = 16m (2SAT) + 8sm	913 Mb (Xie et al. 2019a)	55.93 ± 4.06	4.66	6.78–10.44	41.33 ± 6.20	14.99	17.33	12.67	2A
Momordica charantia	2n = 22 = 20m(4SAT) + 2sm	339 Mb (Urasaki et al. 2016)	21.31 ± 0.85	1.94	6.64–11.63	42.47 ± 4.60	10.83	15.06	17.76	2A
Luffa cylindrica	2n = 26 = 26m	656 Mb (Wu et al. 2020)	43.75 ± 2.16	3.36	7.03–10.41	45.35 ± 2.73	6.01	8.86	9.66	1A
Lagenaria siceraria var. *hispida*	2n = 22 = 20m(2SAT) + 2sm	334 Mb (Achigan-Dako et al. 2008)	28.73 ± 1.69	2.61	6.67–14.52	41.94 ± 3.35	8.00	15.54	24.98	1B
Cucurbita moschata	2n = 40 = 38m + 2sm	372 Mb (Sun et al. 2017)	38.15 ± 2.55	1.91	3.40–6.63	43.82 ± 3.11	7.11	12.37	18.61	2A

Notes: KF, karyotype formula; Genome size, nuclear DNA content of haploid (Values taken from previous reports and the genotypes used in DNA measurements are not necessarily identical to those in this study); TCL, total chromosome length of the haploid complement (i.e. karyotype length); C, mean chromosome length; RRL, ranges of chromosome relative length; CI, mean centromeric index; CV_{CI}, CV_{CL}, Coefficient of variation of the centromeric index and chromosome length, respectively; M_{CA}, Mean centromeric asymmetry; St, the karyotype asymmetry category of Stebbins.
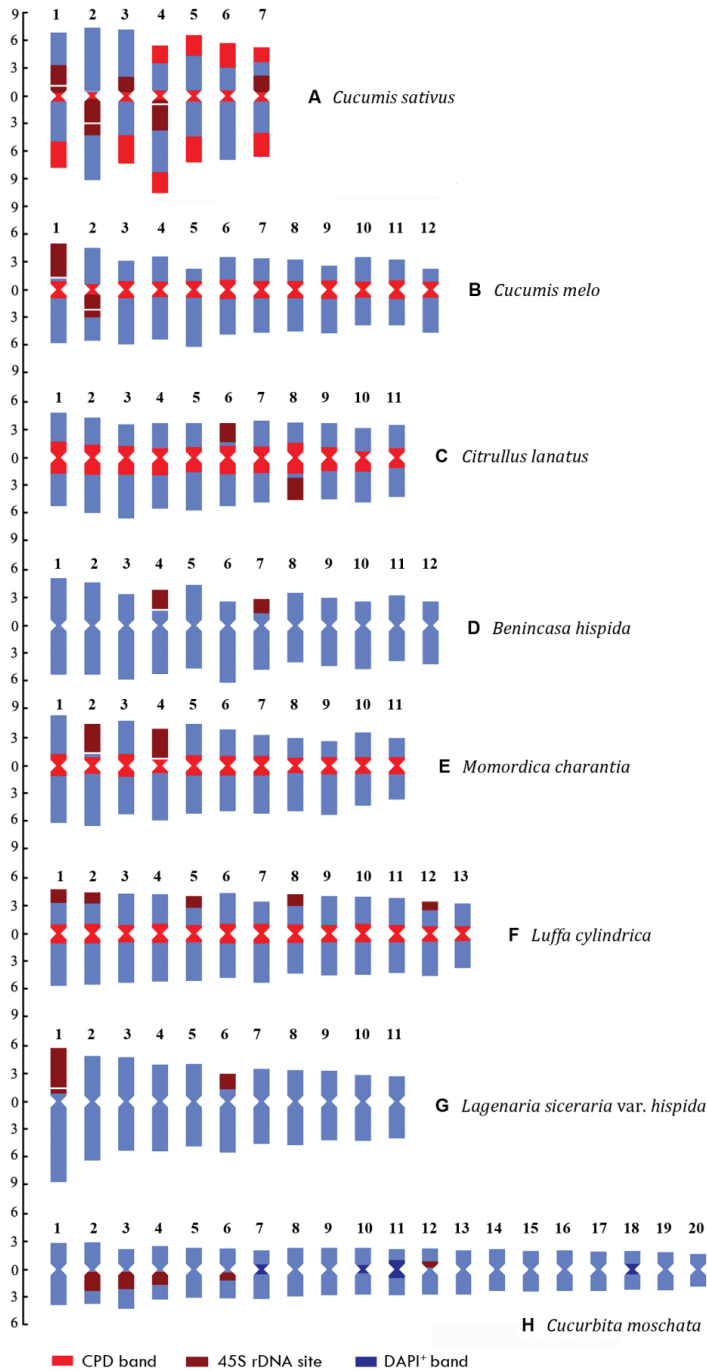
Figure 2. Idiograms of the eight species that display the chromosome measurements, and the position and size of the fluorochrome bands and 45S rDNA FISH signals. **A, B, C, D, E, F, G** and **H** indicate *C. sativus*, *C. melo*, *C. lanatus*, *B. hispida*, *M. charantia*, *L. cylindrica*, *L. siceraria* var. *hispida* and *C. moschata*, respectively. The ordinate scale on the left indicates the relative length of the chromosomes (i.e. % of haploid complement). The numbers at the top indicate the serial number of chromosomes.

Table 2. The distribution of fluorochrome bands and rDNA sites in the eight Cucurbitaceae crops.

Species	Fluorochrome bands				Number (pairs) and location of 45S rDNA sites[†#]
	Type	Distribution[†]	Amount (%)[‡]	Band size (mean)[§]	
Cucumis sativus	CPD	All 45S sites	9.86	1.74–3.93 (2.78)	Five: 1, 3, 7S-PROX (22.06%, 12.04%, 17.41%), 2, 4L-PROX(21.53%, 15.98%)
		All CENs	13.89	0.74–1.74 (1.41)	
		1, 3, 4, 5, 7L-TERs; 4, 5, 6, 7S-TERs	21.98	1.64–3.00 (2.49)	
Cucumis melo	CPD	All 45S sites	6.00	2.42–3.59 (3.00)	Two: 1S-TER(27.03%), 2L-PROX(22.12%)
		All CENs, PCENs	21.62	1.18–2.04 (1.80)	
Citrullus lanatus	CPD	All 45S sites	4.52	2.15–2.37 (2.26)	Two: 6S-TER(41.44%), 8L-TER(48.18%)
		All CENs, PCENs	31.25	2.11–3.50 (2.84)	
Benincasa hispida	CPD	All 45S sites	3.77	1.51–2.26 (1.89)	Two: 4, 7S-TER(41.38%, 46.43%)
Momordica charantia	CPD	All 45S sites	6.95	3.28–3.67 (3.48)	Two: 2S-TER(27.41%), 4S
		All CENs, PCENs	25.72	1.89–2.98 (2.34)	
Luffa cylindrica	CPD	All 45S sites	6.58	1.03–1.83 (1.32)	Five: 1, 2, 5, 8, 12S-TER(61.27%, 71.43%, 69.42%, 70.87%, 70.19%)
		All CENs, PCENs	25.89	1.70–2.21(1.99)	
Lagenaria siceraria var. *hispida*	CPD	All 45S sites	6.77	1.89–4.89 (3.39)	Two: 1S-TER(15.03%), 6S-TER(47.19%)
Cucurbita moschata	CPD	All 45S sites	6.60	0.68–1.99 (1.32)	Five: 2, 3, 4, 6L-PROX(27.03%, 21.26%, 20.10%, 13.44%), 12S-PROX(15.38%)
	DAPI[+]	7, 10, 11, 18-PCENs (post-FISH)	4.88	0.83–1.96 (1.22)	

[†] S and L represent short and long arms, respectively; CEN, PCEN, PROX and TER represent centromeric, pericentromeric, proximal, terminal position, respectively; figures ahead of the positions are the designations of the chromosome pair involved.

[‡] Amount of bands in the genome expressed as percentage of the karyotype length.

[§] The percentage of the size of the bands of each chromosome pair in relation to the karyotype length.

[#] The percentages in square brackets are the percentage distance from centromere to the rDNA site ($di = d \times 100/a$; d = distance of starting point of terminal sites judged by CPD bands or center of non-terminal sites judged by FISH signals from the centromere, a = length of the corresponding chromosome arm).

(Table 1; Suppl. material 3; Fig. 2). As a whole, metacentric chromosomes are the most common form of chromosomes in the complements of the eight species studied, representing 86.60% of all chromosomes. The chromosome pairs 1, 2 and 4 in *C. sativus*, pairs 1 and 2 in *C. melo*, pair 4 in *B. hispida*, pairs 2 and 4 in *M. charantia*, and pair 1 in *L. siceraria* var. *hispida* are satellite chromosomes (SATs) with secondary constrictions inside or in close proximity to the 45S rDNA sites (Figs 1A, C, G, I, M, 2A, B, D, E, G).

The four different karyotype asymmetry indices are given in Table 1. Among these indices, CV_{CI} is the measure of the heterogeneity of centromere position, M_{CA} characterizes the intrachromosomal asymmetry, and CV_{CL} measures the interchromosomal asymmetry (Peruzzi and Eroglu 2013; Astuti et al. 2017). The ranges of CV_{CI}, M_{CA} and CV_{CL} are as follow: CV_{CI} = 6.01 (*L. cylindrica*)-15.99 (*C. melo*), M_{CA} = 8.86 (*L. cylindrica*)-20.27 (*C. melo*), CV_{CL} = 9.66 (*L. cylindrica*)-24.98 (*L. siceraria* var. *hispida*). The M_{CA} values reveal that *L. cylindrica* and *C. melo* have the lowest and the highest intrachromosomal asymmetry, respectively. The CV_{CL} values reveal that *L. cylindrica* and *L. siceraria* var. *hispida* have the least and the most asymmetric karyotype among the eight species in terms of interchromosomal asymmetry. According to the classification of Stebbins (1971), these karyotypes fall into 1A, 1B or 2A categories. That is, the karyotypes of all species studied are rather symmetric.
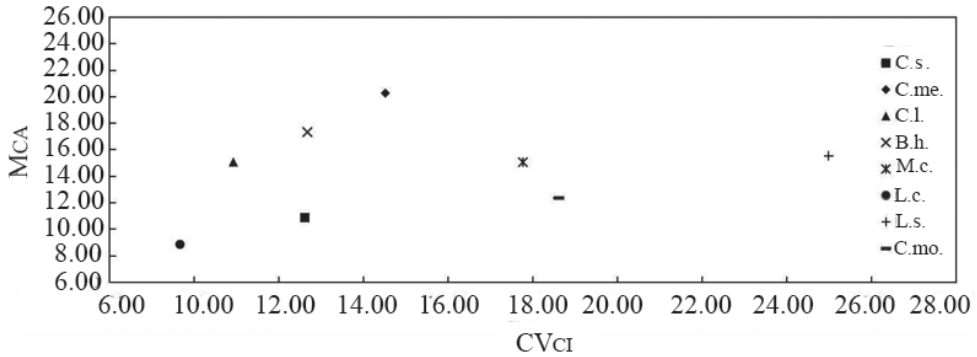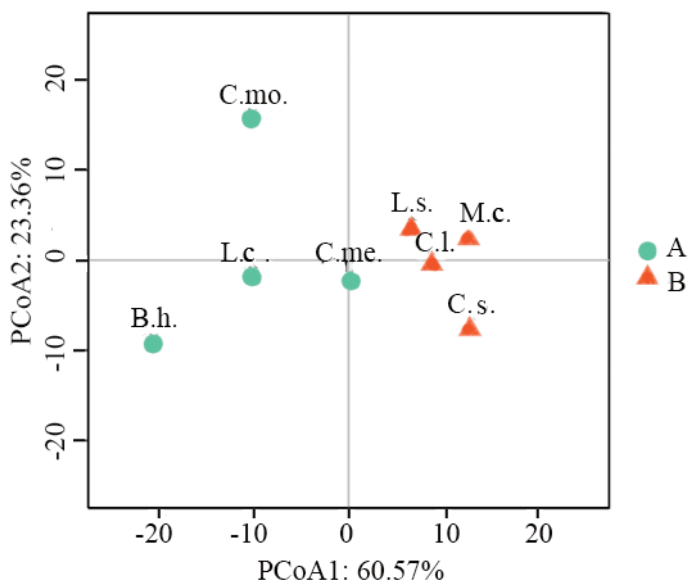
Figure 3. Bidimensional scatter plot of M_{CA} vs. CV_{CL} for the eight Cucurbitaceae species. C.s., C.me., C.l., B.h., M.c., L.c., L.s., and C.mo. represent *C. sativus, C. melo, C. lanatus, B. hispida, M. charantia, L. cylindrica, L. siceraria* var. *hispida* and *C. moschata*, respectively.

The karyotype asymmetry relationships among the eight species that are expressed by means of bidimensional scatter plot of M_{CA} vs. CV_{CL} are illustrated in Figure 3. It is evident that the karyotype structure of these species can be discriminated by this couple of parameters. As depicted in the scatter plot, *L. cylindrica* is the most symmetric karyotype in terms of both intra- and inter-chromosomal asymmetry, while *C. melo* and *L. siceraria* var. *hispida* are the most asymmetric karyotypes in terms of intra- and inter-chromosomal index, respectively (Fig. 3).

Karyological relationships among the studied species revealed by PCoA based on six karyological parameters are illustrated in Figure 4. The PCoA scatter plot shows that the eight species can be divided into two groups along the direction of PCoA1: *L. siceraria* var. *hispida, C. lanatus, M. charantia* and *C. sativus* in one group with the former three species closely clustering together, *C. melo, L. cylindrica, B. hispida* and *C. moschata* in another group in which *C. melo* occupies the middle position of the two groups and *C. moschata* occupies the most isolated position (Fig. 4).

Fluorochrome banding patterns and 45S rDNA sites

CPD staining and DAPI counterstaining revealed distinct heterochromatin differentiation among the eight species (Figs 1, 2; Suppl. material 2; Table 2). In each species, all the chromosomal regions corresponding to the 45S rDNA sites which were confirmed by the subsequent FISH with the 45S rDNA probe showed CPD bands (Fig. 1A, C, E, G, I, K, M, O). All (peri) centromeric regions in *C. sativus, C. melo, C. lanatus, M. charantia* and *L. cylindrica* displayed CPD bands (Fig. 1A, C, E, I, K; Suppl. material 2: fig. S1A–C, E, F), while those in *B. hispida, L. siceraria* var. *hispida* and *C. moschata* did not display CPD bands (Fig. 1G, M, O; Suppl. material 2: fig. S1D, G, H). Particularly, in *C. sativus*, the terminals of the short arms of pairs 4, 5, 6 and 7 and the long arms of pairs 1, 3, 4, 5 and 7 displayed CPD bands (Figs 1A, 2A). In

Figure 4. PCoA for the eight Cucurbitaceae species based on *x*, 2*n*, TCL, M_{CA}, CV_{CL} and CV_{CI}. C.s., C.me., C.l., B.h., M.c., L.c., L.s., and C.mo. represent *C. sativus*, *C. melo*, *C. lanatus*, *B. hispida*, *M. charantia*, *L. cylindrica*, *L. siceraria* var. *hispida* and *C. moschata*, respectively. PCoA1 reflects the original data characteristics before the dimensionality reduction of 60.57%. PCoA2 reflected the character of the original data before the dimensionality reduction of 23.36%. The sum of the two percentages is 83.93%, indicating that the two-dimensional coordinate system can reflect the characteristics of 83.93% of the original data.

C. moschata, after the FISH procedure, DAPI counterstaining showed pericentromeric DAPI⁺ bands (called post-FISH DAPI⁺ bands) on chromosome pairs 7, 10, 11 and 18 (Fig. 1P). The total amount of the (peri)centromeric CPD bands in *C. sativus*, *C. melo*, *C. lanatus*, *M. charantia* and *L. cylindrica* are 13.89%, 21.62%, 31.25%, 25.72% and 25.89% of the karyotype length, respectively (Table 2; Suppl. material 3). The total amount of the terminal CPD bands in *C. sativus* is 21.98% of the karyotype length (Table 2; Suppl. material 3). The total amount of post-FISH DAPI⁺ bands in relation to the karyotype length is 4.88% in *C. moschata* (Table 2; Suppl. material 3). The size of the rDNA CPD bands, non-rDNA CPD bands and post-FISH DAPI⁺ bands varied among the chromosome pairs (Fig. 2; Table 2; Suppl. material 3).

FISH with the 45S rDNA probe onto the chromosomes previously stained by CPD is presented in Figure 1. The number and location of 45S rDNA sites are summarized in Table 2, and illustrated in Figure 2. There are obvious differences in number, size and location among the eight species (Table 2). Two 45S loci were detected in *C. melo*, *C. lanatus*, *B. hispida*, *M. charantia* and *L. siceraria* var. *hispida*, and five 45S loci were detected in *C. sativus*, *L. cylindrica* and *C. moschata* (Figs 1B, D, F, H, J, L, N, P, 2). There were twenty-five 45S rDNA loci in the eight taxa, of which 14 (accounting for 56%) were located at the terminals and 11 (accounting for 44%) were located in

the proximal regions of the respective chromosome arms (Fig. 2; Table 2). In *C. sativus*, the five 45S rDNA loci are located in the proximal regions of the short arms of chromosome pairs 1, 3 and 7 and the long arms of pairs 2 and 4. The 45S rDNA sites of pairs 1, 2 and 4 are major loci in which secondary constrictions appear in prophase and prometaphase cells, while those of pairs 3 and 7 are minor loci (Figs 1A, B, 2A). In *C. melo*, one 45S locus is distally located on the short arms of pair 1 and occupies the majority of the arms, another 45S locus is located in the proximal regions of the long arms of pair 2 (Figs 1C, D, 2B). There are secondary constrictions on the proximal side of the 45S locus of pair 1 and inside the 45S locus of pair 2. In *C. lanatus*, the two 45S rDNA loci are terminally located in the short arms of pair 6 and the long arms of pair 8, respectively (Figs 1E, F, 2C). In *B. hispida*, one 45S locus is located at the terminals of the short arms of pair 4 beside which secondary constrictions occur, another 45S locus is terminally located in the short arms of pair 7 (Figs 1G, H, 2D). The 45S loci of both *C. lanatus* and *B. hispida* account for more than half of the respective arms (Fig. 2C, D; Table 2). In *M. charantia*, one 45S locus is terminally located on the short arms of pair 2 and occupies the majority of the arms, another locus occupies the entire short arms of pair 4 (Figs 1I, J, 2E; Table 2). There are secondary constrictions beside the two 45S loci (Figs 1I, 2E). The five 45S loci of *L. cylindrica* are relatively small and located at the terminals of the short arms of pairs 1, 2, 5, 8 and 12 (Figs 1K, L, 2F). In *L. siceraria* var. *hispida*, one 45S locus is located at the terminals of the short arms of pair 1 which accounts for the majority of the arms and produces secondary constrictions in the interior of the sites, another locus is situated at the terminals of the short arms of pair 6 (Figs 1M, N, 2G). In *C. moschata*, three major 45S loci are located in the proximal regions of the long arms of pairs 2, 3 and 4, respectively, two minor loci are proximally placed on the long arms of pair 6 and the short arms of pair 12, respectively (Figs 1O, P, 2H). In particular, the size of the hybridization signals and CPD bands of the 45S rDNA sites of pair 3 vary significantly between two homologous chromosomes (Fig. 1O, P), indicating the heterozygosity of the *C. moschata* accession analyzed in this study.

We find that all mitotic chromosomes of *C. sativus* can be precisely identified by the combination of the 45S rDNA FISH signals and terminal CPD bands (Figs 1A, B, 2A). The features of each chromosome pair of *C. sativus* are as follows. Chromosome 1 has strong 45S rDNA signal in the proximal region of the short arm and terminal CPD band on the long arms. Chromosome 2 has strong 45S rDNA signal in the proximal region of the long arm and is devoid of CPD band at the terminals of both arms. Chromosome 3 has weak 45S rDNA signal in the proximal region of the short arm and terminal CPD band on the long arm. Chromosome 4 has strong 45S rDNA signal in the proximal region of the long arm and terminal CPD bands on both arms. Chromosome 5 has terminal CPD bands on both arms and is devoid of 45S rDNA signal. Chromosome 6 has terminal CPD band on the short arm and is devoid of 45S rDNA signal. Chromosome 7 has weak 45S rDNA signal in the proximal region of the short arm and terminal CPD bands on both arms.

Discussion

Karyotype features and 45S rDNA patterns

Precise chromosome measurement is essential for accurate karyotype analysis. Chromosomes should have morphologically distinct primary constrictions and clearly defined boundaries; otherwise, it is difficult to determine the length of chromosome arms and, consequently, to calculate chromosomal parameters (Kadluczka and Grzebelus 2021). In our previous cytogenetic investigations, it was found that the morphological differentiation of mitotic chromosomes depended on the degree of condensation (She et al. 2006, 2015, 2017, 2020; She and Jiang 2015). When condensation is insufficient, the terminals of chromosome arms may be still in the decondensation state, and then the boundary of chromosomes is not clear. However, when chromosomes condensed to the maximum, the morphological differences between chromosomes decreased. Therefore, it is important to select metaphase chromosomes with moderate degree of condensation for identification and measurement of chromosomes. This is especially true for species with small chromosomes. In addition, the landmarks produced by CPD staining and rDNA FISH facilitated chromosome identification (She et al. 2006, 2015, 2017, 2020; She and Jiang 2015). In this study, well morphologically differentiated metaphase chromosomes of the eight Cucurbitaceae crops were prepared using the EMF method and characterized by fluorochrome banding and 45S rDNA-FISH. Detailed karyotypes of the eight species were established with a combined dataset consisted of chromosome measurements, fluorochrome bands and 45S rDNA FISH signals. Furthermore, four different karyotype asymmetry indices of the eight species were simultaneously measured for the first time. Therefore, the newly constructed karyotypes of these species are more accurate, informative and comparable.

The current karyotype of *C. sativus* differs in chromosome size and the classification of chromosome morphotype from some previously reported karyotypes of this species. The range of chromosome size and TCL detected in our study are similar to those of Trivedi and Roy (1970), but much larger than those of Chen et al. (1998) and Koo et al. (2002). The karyotype formula obtained here is similar to those reported by Li (1989), Chen et al. (1998), Koo et al. (2002) and Waminal and Kim (2012) which also consisted of 12 m and 2 sm, but different from those reported by Trivedi and Roy (1970), Han et al. (2008) and Zhang et al. (2015b) in which all chromosomes were metacentric, and that reported by Hoshi et al. (1998) which was comprised of 10 m and 4 sm. In the current karyotype, chromosome 2 is the longest and chromosome 4 is submetacentric, while in the karyotype reported by Han et al. (2008), chromosome 3 is the longest and chromosome 4 is metacentric. We identified three pairs of satellite chromosomes (pairs 1, 2 and 4) in *C. sativus* by means of CPD staining. Secondary constructions had been observed in *C. sativus* by several investigators (Bhaduri and Bose 1947; Trivedi and Roy 1970; Ramachandran and Seshadri 1986; Li 1989), but the number of secondary constrictions have not been reliably confirmed using the conventional staining. The 45S rDNA pattern of *C. sativus* including the number, position and size of 45S sites revealed by us is consistent with those reported by Han et al (2008).

The chromosome size of *C. melo* detected by us is similar to that of Li (1989), but differs considerably from that of Trivedi and Roy (1970). Our karyotype of *C. melo* which is composed of both m and sm chromosomes is similar to that of Zhang et al. (2015a), but differs from those of Liu et al. (2010) and Li (1989) in which one or more pairs of subterminal (st) chromosomes were involved. In current study, all centromeres of *C. melo* were marked by the CPD bands, enabling the arm ratio of each chromosome to be accurately measured. The number and position of the 45S sites of *C. melo* detected in this study is coincident with the previous reports (Zhang et al. 2015a; Zhang et al. 2016). In our karyotype of *C. melo*, the chromosome pairs bearing the 45S sites are designated as pairs 1 and 2 according to the descending order of the length. They should correspond to chromosome pairs 4 and 10 in the karyotype of Zhang et al. (2015a), respectively. In addition, the arms of pair 2 that have 45S sites are identified as the long arms in our study, instead of the short arms as described by Zhang et al. (2015a).

The TCL of *C. lanatus* detected by us is larger than those of Waminal et al. (2011) and Trivedi and Roy (1970). Our karyotype formula of *C. lanatus* is accordance with those described by Li (1989) and Li et al. (2007), and slightly differs from that reported by Waminal et al. (2011) in which 14 m and 8 sm were involved. The 45S rDNA sites situated on chromosome pairs 6 and 8 of *C. lanatus* in our study should correspond to those mapped on pairs 8 and 4 by Ren et al. (2012), Guo et al. (2013) and Li et al. (2016), respectively. However, the rDNA-bearing arms of pair 8 was designated as the long arms in our study rather than the shorts arms (Ren et al. 2012; Guo et al. 2013; Li et al. 2016).

The karyotype formula of *B. hispida* obtained in our study resemble those reported by Li (1989) and Xu et al. (2007), but varies significantly from that of Waminal et al. (2011). The TCL of *B. hispida* detected by us is similar to that of Li (1989), but larger than that of Waminal et al. (2011). Our study demonstrates that both of the two 45S rDNA sites in *B. hispida* are located at the terminals of the short arms of the respective chromosomes instead of subtelomeric or interstitial regions of the respective short arms (Xu et al. 2007; Waminal et al. 2011), which is consistent with the result of Xie et al. (2019b).

Our karyotype formula of *M. charantia* is similar to that reported by Li (1989) and slightly different from those of Waminal and Kim (2012) and Li et al. (2007). The TCL detected by us is larger than that of Waminal and Kim (2012), and smaller than that of Li (1989). In this study, sequential CPD staining and 45S rDNA FISH reveal that the two 45S rDNA loci occupy the majority or the entire length of the respective short arms, providing a more accurate mapping of the 45S rDNA sites in *M. charantia* than previous studies (Li et al. 2007; Waminal and Kim 2012; Xie et al. 2019b).

The karyotype formula of *L. cylindrica* obtained by us is in accordance with that of Xu et al. (2007), but slightly differing from those of Li (1989) and Waminal and Kim (2012). The TCL of *L. cylindrica* detected in our study is similar to that of Li (1989), but is much larger than that of Waminal and Kim (2012). Our study demonstrates that all the five 45S rDNA loci are terminally located on the short arms of five chromosome pairs, being consistent with the result of Waminal and Kim (2012), but different from the result of Xu et al. (2007) in which the positions of the five 45S loci were identified as subtelomeric regions.

Our karyotype formula of *L. siceraria* var. *hispida* is coincident with that of Li (1989), but differs from the karyotype of *L. siceraria* reported by Waminal and Kim (2012). The TCL of *L. siceraria* var. *hispida* detected in this study is slightly smaller than that of Li (1989), but much larger than the TCL of *L. siceraria* detected by Waminal and Kim (2012). Our study reveals that the number of 45S rDNA sites in *L. siceraria* var. *hispida* is the same as in *L. siceraria*, and the rDNA sites are also located at the terminals of the short arms of two chromosome pairs (Waminal and Kim 2012; Li et al. 2016; Xie et al. 2019b).

The karyotype formula of *C. moschata* constructed by us consists of only m and sm chromosomes, being similar to that reported by Waminal et al. (2011). However, it is considerably different from those of Li (1989) and Xu et al. (2007) in which, except for m and sm chromosomes, four and eight st chromosomes were involved, respectively. The chromosome size of *C. moschata* detected by us is similar to that of Li (1989), but much larger than that of Waminal et al. (2011). In our study, the locations and sizes of the five 45S rDNA loci of *C. moschata* are determined accurately by the rDNA CPD bands, demonstrating that all the five 45S loci are at proximal localization instead of terminal localization (Xu et al. 2007; Waminal et al. 2011).

The discrepancies in karyotype feature and 45S rDNA pattern between our results and the previous reports are probably due to differences in the accessions analyzed, the condensation level of measured chromosomes, and the difficulty in identifying chromosomes using the mitotic chromosome spreads of lower quality in the previous studies.

Genome differentiation between species

The total chromosome length of the haploid complement (TCL) can be used as a proxy for genome size (Levin 2002). Previous studies found that the correlations between TCL and DNA content typically exceeded $r = 0.85$ within species, between congeneric species and among species in related genera (Levin 2002; Carta and Peruzzi 2016). In our study, the TCL of each of the eight taxa was measured using the chromosomes with the highest degree of condensation (She et al. 2015), and therefore, the TCLs of these species can be well comparable with each other. The correlation analysis using the SPSS 25.0 software (Suppl. material 4) reveals a high correlation between the difference in TCL and the change in nuclear DNA content within the eight taxa ($r = 0.899$, $p < 0.01$), providing new evidence of the feasibility of comparing genome size based on TCL values among species of related genera. For example, whether according to TCL or nuclear DNA content, *B. hispida* was the largest genome and *M. charantia* was the smallest genome, and *L. cylindrica* genome was about twofold larger than *M. charantia* genome (Urasaki et al. 2016; Xie et al. 2019a; Wu et al. 2020). However, it was also found that TCL and DNA content values were incompletely proportional to each in some other cases. For example, the DNA contents of *C. sativus* and *C. moschata* are almost equal (Huang et al. 2009; Sun et al. 2017), but the TCL of *C. moschata* is 1.6 times as much as that of *C. sativus*. Several studies have revealed that total chromosome volume, instead of TCL could be a descriptor of chromosome size, and more suitable

to reflect genome size (Kramer et al. 2021; Mehravi et al. 2022). This may account for the inconsistency. Increase in genome size may, in general, be attributed to transposable element amplification and to polyploidization. Genome sequencing studies revealed that repeat expansion led to large genome size in cucurbits (Garcia-Mas et al. 2012; Xie et al. 2019a; Wu et al. 2020). For example, *B. hispida* genome did not have any recent lineage-specific whole-genome duplication as other sequenced species in the tribe Benincaseae including *C. sativus*, *C. melo*, *C. lanatus* and *L. cylindrica*, the substantial accumulation of transposable elements and especially LTR retrotransposons contributes greatly to the large genome size of this species (Xie et al. 2019a; Wu et al. 2020).

The differences in CPD and DAPI[+] bands, with regard to presence, position and size, reveal distinct heterochromatin differentiation among the eight cucurbits studied. CPD staining reveals the occurrence of (peri)centromeric GC-rich heterochromatin in *C. sativus*, *C. melo*, *C. lanatus*, *M. charantia* and *L. cylindrica*, and terminal GC-rich heterochromatin in *C. sativus*. (Peri)centromeric and terminal GC-rich heterochromatin was previously detected in *C. sativus* using CMA/DAPI staining (Plader et al. 1998). According to the recent classification of Cucurbitaceae, *Cucumis*, *Benincasa*, *Citrullus* and *Lagenaria* Seringe, 1825 belong to tribe Benincaseae, and *Luffa* Miller, 1754, *Momordica* Linnaeus, 1753 and *Cucurbita* Linnaeus, 1753 belong to tribe Luffeae, Joliffieae and Cucurbiteae, respectively (Jeffrey 2005; Kocyan et al. 2007). *Coccinia grandis* (Linnaeus) Voigt, 1845 ($2n$ = 24), a species of tribe Benincaseae, showed centromeric GC-rich heterochromatin in the majority of chromosome pairs by CMA/DAPI staining (Bhowmick et al. 2012). These facts suggest that the presence of (peri)centromeric GC-rich heterochromatin is an ancestral genome feature that occurred before the divergence of Subfamily Cucurbitoideae. However, the inexistence of (peri)centromeric GC-rich heterochromatin in *B. hispida*, *L. siceraria* var. *hispida* and *C. moschata* seems to be contradiction with this speculation. A reasonable explanation is that the (peri)centromeric GC-rich heterochromatin of these three species has undergone a reduction of GC content after speciation, resulting in the disappearance of red CPD bands (She et al. 2006, 2015). The (peri)centromeric CPD bands in *C. melo* may result from staining of the centromere-specific repeats (CmCent) whose GC contents is rather high (56–57%) (Koo et al. 2010). The terminal GC-rich heterochromatin of *C. sativus* should be differentiated during speciation because *C. melo* has not such heterochromatin. In the *C. sativus* genome, two types of tandem repeats, Type I/II and Type IV, are located in the subtelomeric regions of the majority of chromosome arms in the complement, while tandem repeat type III and centromere-specific satellite 1 (CsCent1) are located in the centromeres of each chromosome (Han et al. 2008; Ren et al. 2009; Koo et al. 2010; Zhao et al. 2011). Given that Type I/II, Type III, Type IV, and CsCent1 have higher GC content (Han et al. 2008; Koo et al. 2010), the centromeric CPD bands may result from the staining of type III and CsCent1, and the terminal CPD bands may result from the staining of Type I/II and Type IV. The FISH signals of Type I/II and Type IV were detected all but one of chromosome arms (Han et al. 2008; Ren et al. 2009), while the terminal CPD bands were only detected on 9 out of 14 chromosome arms. This discrepancy is probably attributed to the lower

sensitivity of CPD staining compared with FISH technology or the difference in the accessions analyzed. The occurrence of post-FISH DAPI⁺ bands in *C. moschata* was a prominent indication of heterochromatic differentiation. DAPI⁺ bands revealed only after FISH procedure have been reported in many plant species (e.g. Morales et al. 2012; She et al. 2015), and should represent another kind of heterochromatin that is different from GC- and AT-rich heterochromatin (Barros e Silva and Guerra 2010).

The number, location and distribution of the 5S and 45S rDNA clusters in chromosomes are useful for deducing species history and phylogenetic relationship (Weiss-Schneeweiss et al. 2008). We statistically analyzed the number and position of the 45S rDNA sites from 58 species (subspecies or varieties) in Cucurbitaceae (a total sample of 64 karyotypes) that have been investigated by FISH up to now (Suppl. material 5). In the 64 karyotypes, 43 of which are of the species belonging to tribe Benincaseae (Kocyan et al. 2007), there were a total of 182 45S loci, of which 137 (accounting for 75.3%) were located in the terminal portions of chromosomes (including the sites occupying the whole arm), 41 loci in the proximal regions, and 4 in the pericentromeric or interstitial regions (Suppl. material 5); and the number of 45S rDNA sites per complement ranged from one pair up to seven pairs with the most frequent numbers of sites per karyotype being two pairs (accounting for 43.7%). Such distribution feature of 45S rDNA sites in Cucurbitaceae is basically consistent with the general distribution pattern in the entire angiosperms which shows that 45S rDNA sites occur preferentially on the short arms and in the terminal regions of chromosomes (Roa and Guerra 2012). Four of the eight species studied here, including *C. lanatus*, *B. hispida*, *L. siceraria* var. *hispida* and *M. charantia* had similar distribution of 45S rDNA sites (owning two 45S loci of terminal position), suggesting a close relationship between these four species. This was in accord with the results based on molecular phylogenetic analyses, which revealed that there were close relationships among the genera *Citrullus*, *Benincasa* and *Lagenaria* (Zhang et al. 2018; Xie et al. 2019a; Wu et al. 2020), and that *M. charantia* was more related to *C. lanatus* than to *C. sativus* and *C. melo* (Urasaki et al. 2016). In addition, among the 43 karyotypes of tribe Benincaseae, 22 karyotypes (accounting for 51.2%) had two 45S rDNA loci of terminal position (Suppl. material 5). Based on these facts, we speculate that the ancestral progenitor of tribe Benincaseae might bear two 45S loci that were located in the terminal portions of two chromosome pairs. Very recently, the ancestral karyotype of *Cucumis* was reconstructed using comparative oligo-painting, which owns two 45S rDNA loci that located in the terminal regions of two short arms (Zhao et al. 2021). The ancestral karyotype of lineage I of *Cucumis* ($2n = 2x = 24$), which has two 45S rDNA loci of terminal position, evolved to *C. melo* mainly by inversions, and evolved to *C. sativus* mainly by chromosome fusions and inversions (Zhao et al. 2021).

Concerning karyotype asymmetry is one of the most popular, cheap and widely used cytotaxonomic approach. Up to now, a variety of parameters and indices for evaluating karyotype asymmetry have been proposed, including the quali-quantitative one, Stebbins category (Stebbins 1971), as well as several quantitative indices (for details and references see Paszko 2006; Peruzzi and Eroglu 2013). Critical reviews have confirmed

that CV_{CL} is a powerful statistical parameter for estimating the interchromosomal asymmetry, M_{CA} is the most appropriate parameter for a measure of intrachromosomal asymmetry, and other quantitative indices are outdated, redundant, or statistically incorrect (Paszko 2006; Peruzzi and Eroglu 2013; Astuti et al. 2017). The best way in representing karyotype asymmetry relationships among taxa is by means of bidimensional scatter plot, where the two asymmetry estimators are put in the x and y axes and points represent each sample (Peruzzi et al. 2009; Peruzzi and Eroglu 2013; Dehery et al. 2020). Our results show that the karyotype asymmetry relationships among the eight Cucurbitaceae species studied can be best explained by means of the scatter plot of M_{CA} vs. CV_{CL}, confirming that this couple of indices are reliable to assess chromosome asymmetry.

In order to compare karyotypes and reconstructing karyological relationships among the eight species, we applied the methodology proposed by Peruzzi and Altınordu (2014), considering six quantitative parameters (x, $2n$, TCL, M_{CA}, CV_{CL}, CV_{CI}) and subjecting them to PCoA (Dehery et al. 2020; Kadluczka and Grzebelus 2021). Our results demonstrated PCoA with the six parameters was indeed a good way to establish the karyological relationships among the eight Cucurbitaceae species because the karyological relationships among the eight taxa established by PCoA was found to be basically accordant with the phylogenetic relationships revealed by DNA sequences. In the molecular phylogenetic trees, *C. lanatus* and *L. siceraria* were very closely related, and both of them were closely to *C. melo* and *C. sativus* (Kocyan et al. 2007; Zhang et al. 2018). Genome sequencing analysis revealed that *M. charantia* was more related to *C. lanatus* than to *C. sativus* and *C. melo* (Urasaki et al. 2016). However, we also observed something different from the molecular evolutionary trees. *B. hispida*, a species with close relationship with *C. lanatus* and *L. siceraria* in the molecular phylogenetic trees (Kocyan et al. 2007; Xie et al. 2019a), was distantly separated from the two species in the PCoA scatter plot. On the whole, to infer the direction of changes of karyotype evolution in Cucurbitaceae species, karyotype asymmetry study using the multivariate quantitative approach is recommended as one of the complementary characters besides the molecular taxonomic character.

The reported basic chromosome numbers of the Cucurbitaceae family ranged from $x = 7$ to 20, with $x = 11$ a prevalent number (Carta et al. 2020). Recent comparative analyses of six cucurbit genomes reveal that the *B. hispida* genome represents the most ancestral karyotype, with the predicted ancestral genome having 15 proto-chromosomes (Xie et al. 2019a). After *B. hispida*, the 15 ancestral chromosomes were either retained or form new chromosomes through chromosome arrangements such as fusions, fissions, inversions during the speciation and evolution of later species (Xie et al. 2019a). Recent studies revealed that *C. melo* and *C. sativus* were evolved from an ancestral karyotype ($x = 12$) by large-scale inversions, centromere repositioning and chromothripsis-like rearrangement (Zhao et al. 2021), and *C. moschata* resulted from an ancient allotetraploidization event (Sun et al. 2017). Alterations in chromosome symmetry may arise through chromosome arrangements including translocations, pericentric inversions, fusions or fissions (Schubert 2007), or through removes or addition of the same amount of DNA from/to both arms of chromosomes (Levin 2002; Peruzzi et

al. 2009). Although many alterations in number and structure of chromosomes as well as genome size occurred during speciation and evolution, the karyotype asymmetry of the eight species involved in this study had not changed significantly, and their karyotypes were all symmetrical. The reasons for this are worth further study.

Conclusions

Detailed karyotypes of eight Cucurbitaceae crops, *C. sativus*, *C. melo*, *C. lanatus*, *B. hispida*, *M. charantia*, *L. cylindrica*, *L. siceraria* var. *hispida* and *C. moschata*, were reconstructed using the dataset of chromosome measurements, fluorochrome bands and 45S rDNA FISH signals. Comparative karyotyping revealed distinct variations in the karyotypic parameters, and the patterns of fluorochrome bands and 45S rDNA sites among species. The karyological relationships among the eight taxa based on six karyological parameters was basically accordant with the phylogenetic relationships revealed by DNA sequences, indicating that karyotype asymmetry study using the multivariate quantitative approach is one of the complementary characters for inferring the direction of changes of karyotype evolution in Cucurbitaceae species.

Competing interests

The authors have declared that no competing interests exist.

Acknowledgements

References

Achigan-Dako EG, Fuchs J, Ahanchede A, Blattnerv FR (2008) Flow cytometric analysis in *Lagenaria siceraria* (Cucurbitaceae) indicates correlation of genome size with usage types and growing elevation. Plant Systematics and Evolution 276(9): 9–19. https://doi.org/10.1007/s00606-008-0075-2

Astuti G, Roma-Marzio F, Peruzzi L (2017) Traditional cytotaxonomic studies: can they still provide a solid basis in plant systematics? Flora Mediterranea 27: 91–98. https://doi.org/10.7320/FlMedit27.091

Barros e Silva AE, Guerra M (2010) The meaning of DAPI bands observed after C-banding and FISH procedures. Biotechnic & Histochemistry 85(2): 115–125. https://doi.org/10.1080/10520290903149596

Beevy SS, Kuriachan P (1996) Chromosome numbers of South Indian Cucurbitaceae and a note on the cytological evolution in the family. Journal of Cytology and Genetics 31(1): 65–71.

Bhaduri PN, Bose PC (1947) Cyto-genetical investigations in some common cucurbits, with special reference to fragmentation of chromosomes as a physical basis of speciation. Journal of Genetics 48(2): 237–256. https://doi.org/10.1007/BF02989384

Bhowmick BK, Jha TB, Jha S (2012) Chromosome analysis in the dioecious cucurbit *Coccinia grandis* (L.) Voigt. Chromosome Science 15(1–2): 9–15. https://doi.org/10.11352/scr.15.9

Bisognin DA (2002) Origin and evolution of cultivated cucurbits. Ciência Rural 32(5): 715–723. https://doi.org/10.1590/S0103-84782002000400028

Carta A, Bedini G, Peruzzi L (2020) A deep dive into the ancestral chromosome number and genome size of flowering plants. New Phytologist 228(3): 1097–1106. https://doi.org/10.1111/nph.16668

Carta A, Peruzzi L (2016) Testing the large genome constraint hypothesis: plant traits, habitat and climate seasonality in Liliaceae. New Phytologist 210(2): 709–716. https://doi.org/10.1111/nph.13769

Chen JF, Staub JE, Adelberg JW, Jiang J (1999) Physical mapping of 45S rRNA genes in *Cucumis* species by fluorescence *in situ* hybridization. Canadian Journal of Botany 77(3): 389–393. https://doi.org/10.1139/b98-226

Chen JF, Staub JE, Jiang J (1998) A reevaluation of karyotype in cucumber (*Cucumis sativus* L.). Genetic Resources and Crop Evolution 45(4): 301–305. https://doi.org/10.1023/A:1008603608879

Dane F, Tsuchiya T (1976) Chromosome studies in the genus *Cucumis*. Euphytica 25(1): 367–374. https://doi.org/10.1007/BF00041569

de Moraes AP, dos Santos Soares-Filho W, Guerra M (2007) Karyotype diversity and the origin of grapefruit. Chromosome Research 15(1): 115–121. https://doi.org/10.1007/s10577-006-1101-2

Dehery SK, Panda E, Saha PR, Sinha RK, Das AB (2021) Chromosome diversity and karyotype asymmetry analysis in four cultivated triploid and three diploid wild genotypes of *Musa* from north-East India. The Nucleus 64(2): 167–179. https://doi.org/10.1007/s13237-020-00334-z

Garcia-Mas J, Benjak A, Sanseverino W, Bourgeois M, Mir G, González VM, Hénaff E, Câmara F, Cozzuto L, Lowy E, Alioto T, Capella-Gutiérrez S, Blanca J, Cañizares J, Ziarsolo P, Gonzalez-Ibeas D, Rodríguez-Moreno L, Droege M, Du L, Alvarez-Tejado M, Lorente-Galdos B, Melé M, Yang L, Weng Y, Navarro A, Marques-Bonet T, Aranda MA, Nuez F, Picó B, Gabaldón T, Roma G, Guigó R, Casacuberta JM, Arús P, Puigdomènech P (2012) The genome of melon (*Cucumis melo* L.). Proceedings of the National Academy of Sciences of the United States of America 109(29): 11872–11877. https://doi.org/10.1073/pnas.1205415109

Guerra M (2012) Cytotaxonomy: the end of childhood. Plant Biosystems 146(3): 703–710. https://doi.org/10.1080/11263504.2012.717973

Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, Zheng Y, Mao L, Ren Y, Wang Z, Min J, Guo X, Murat F, Ham BK, Zhang Z, Gao S, Huang M, Xu Y, Zhong S, Bombarely A, Mueller LA, Zhao H, He H, Zhang Y, Zhang Z, Huang S, Tan T, Pang E, Lin K, Hu Q,

Kuang H, Ni P, Wang B, Liu J, Kou Q, Hou W, Zou X, Jiang J, Gong G, Klee K, Schoof H, Huang Y, Hu X, Dong S, Liang D, Wang J, Wu K, Xia Y, Zhao X, Zheng Z, Xing M, Liang X, Huang B, Lv T, Wang J, Yin Y, Yi H, Li R, Wu M, Levi A, Zhang X, Giovannoni JJ, Wang J, Li Y, Fei Z, Xu Y (2013) The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. Nature Genetics 45(1): 51–58. https://doi.org/10.1038/ng.2470

Han Y, Zhang T, Thammapichai P, Weng Y, Jiang J (2015) Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides. Genetics 200(3): 771–779. https://doi.org/10.1534/genetics.115.177642

Han Y, Zhang Z, Huang S, Jin W (2011) An integrated molecular cytogenetic map of *Cucumis sativus* L. chromosome 2. BMC Genetics 12: 18. https://doi.org/10.1186/1471-2156-12-18

Han Y, Zhang Z, Liu C, Liu J, Huang S, Jiang J, Jin W (2009) Centromere repositioning in cucurbit species: implication of the genomic impact from centromere activation and inactivation. Proceedings of the National Academy of Sciences of the United States of America 106(35): 14937–14941. https://doi.org/10.1073/pnas.0904833106

Han YH, Zhang Z, Liu JH, Lu JY, Huang SW, Jin WW (2008) Distribution of the tandem repeat sequences and karyotyping in cucumber (*Cucumis sativus* L.) by fluorescence *in situ* hybridization. Cytogenetic and Genome Research, 122(1): 90–98. https://doi.org/10.1159/000151320

Hasterok R, Jenkins G, Langdon T, Jones RN, Maluszynska J (2001) Ribosomal DNA is an effective marker of *Brassica* chromosomes. Theoretical and Applied Genetics 103(4): 486–490. https://doi.org/10.1007/s001220100653

Hoshi Y, Plader W, Malepszy S (1998) New C-banding pattern for chromosome identification in cucumber (*Cucumis sativus* L.). Plant Breeding 177(1): 77–82. https://doi.org/10.1111/j.1439-0523.1998.tb01452.x

Hoshi Y, Plader W, Malepszy S (1999) Physical mapping of 45S rRNA gene loci in the cucumber (*Cucumis sativus* L.) using fluorescence *in situ* hybridization. Caryologia 52(1–2): 49–57. https://doi.org/10.1080/00087114.1998.10589153

Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas WJ, Wang X, Xie B, Ni P, Ren Y, Zhu H, Li J, Lin K, Jin W, Fei Z, Li G, Staub J, Kilian A, van der Vossen EA, Wu Y, Guo J, He J, Jia Z, Ren Y, Tian G, Lu Y, Ruan J, Qian W, Wang M, Huang Q, Li B, Xuan Z, Cao J, Asan, Wu Z, Zhang J, Cai Q, Bai Y, Zhao B, Han Y, Li Y, Li X, Wang S, Shi Q, Liu S, Cho WK, Kim JY, Xu Y, Heller-Uszynska K, Miao H, Cheng Z, Zhang S, Wu J, Yang Y, Kang H, Li M, Liang H, Ren X, Shi Z, Wen M, Jian M, Yang H, Zhang G, Yang Z, Chen R, Liu S, Li J, Ma L, Liu H, Zhou Y, Zhao J, Fang X, Li G, Fang L, Li Y, Liu D, Zheng H, Zhang Y, Qin N, Li Z, Yang G, Yang S, Bolund L, Kristiansen K, Zheng H, Li S, Zhang X, Yang H, Wang J, Sun R, Zhang B, Jiang S, Wang J, Du Y, Li S (2009) The genome of the cucumber, *Cucumis sativus* L. Nature Genetics 41(12): 1275–1281. https://doi.org/10.1038/ng.475

Jeffrey C (2005) A new system of Cucurbitaceae. Botanicheskii Zhurnal 90: 332–335.

Kadluczka D, Grzebelus E (2021) Using carrot centromeric repeats to study karyotype relationships in the genus *Daucus* (Apiaceae). BMC Genomics 22(1): 508. https://doi.org/10.1186/s12864-021-07853-2

Kramer EM, Tayjasanant PA, Cordone B (2021) Scaling laws for mitotic chromosomes. Frontiers in Cell and Developmental Biology 9: 684278. https://doi.org/10.3389/fcell.2021.684278

Kocyan A, Zhang LB, Schaefer H, Renner SS (2007) A multi-locus chloroplast phylogeny for the Cucurbitaceae and its implications for character evolution and classification. Molecular Phylogenetics and Evolution 44(2): 553–577. https://doi.org/10.1016/j.ympev.2006.12.022

Koo DH, Choi HW, Cho J, Hur Y, Bang JW (2005) A high resolution karyotype of cucumber (*Cucumis sativus* L. 'Winter Long') revealed by C-banding, pachytene analysis, and RAPD-aided fluorescence *in situ* hybridization. Genome 48(3): 534–540. https://doi.org/10.1139/g04-128

Koo DH, Hur Y, Jin DC, Bang JW (2002) Karyotype analysis of a Korean cucumber cultivar (*Cucumis sativus* L. cv. Winter Long) using C-banding and bicolor fluorescence *in situ* hybridization. Molecules and Cells 13(3): 413–418. https://doi.org/10.1046/j.1524-475X.2002.10410.x

Koo DH, Nam YW, Choi D, Bang JW, de Jong H, Hur Y (2010) Molecular cytogenetic mapping of *Cucumis sativus* and *C. melo* using highly repetitive DNA sequences. Chromosome Research 18(3): 325–336. https://doi.org/10.1007/s10577-010-9116-0

Levin DA (2002) The role of chromosomal change in plant evolution. Oxford University Press, New York, 49–52.

Li KP, Wu YX, Zhao H, Wang Y, Lü XM, Wang JM, Xu Y, Li ZY, Han YH (2016) Cytogenetic relationships among *Citrullus* species in comparison with some genera of the tribe Benincaseae (Cucurbitaceae) as inferred from rDNA distribution patterns. BMC Evolutionary Biology 16: 85. https://doi.org/10.1186/s12862-016-0656-6

Li MX, Chen RY (1985) A suggestion on the standardization of karyotype analysis in plants. Journal of Wuhan Botanical Research 3(4): 297–302.

Li Q, Ma L, Huang J, Li L (2007) Chromosomal Localization of ribosomal DNA sites and karyotype analysis in three species of Cucurbitaceae. Journal of Wuhan University (Nat. Sci. Ed.) 53(4): 449–456.

Li RQ (1989) Studies on karyotypes of vegetables in China. Wuhan University Press, Wuhan, 106–131.

Lima-de-Faria A (1980) Classification of genes, rearrangements and chromosomes according to the chromosome field. Hereditas 93(1): 1–46. https://doi.org/10.1111/j.1601-5223.1980.tb01043.x

Liu C, Liu J, Li H, Zhang Z, Han Y, Huang S, Jin W (2010) Karyotyping in melon (*Cucumis melo* L.) by cross-species fosmid fluorescence *in situ* hybridization. Cytogenetic and Genome Research 129(1–3): 241–249. https://doi.org/10.1159/000314343

Mehravi S, Karimzadeh G, Kordenaeej A, Hanifei M (2022) Mixed-ploidy and dysploidy in *Hypericum perforatum*: A karyomorphological and genome size study. Plants 11: 3068. https://doi.org/10.3390/plants11223068

Morales AG, Aguiar-Perecin MLR, Mondin M (2012) Karyotype characterization reveals an up and down of 45S and 5S rDNA sites in *Crotalaria* (Leguminosae-Papilionoideae) species of the section Hedriocarpae subsection Macrostachyae. Genetic Resources and Crop Evolution 59(2): 277–288. https://doi.org/10.1007/s10722-011-9683-8

Moscone EA, Lein F, Lambrou M, Fuchs J, Schweizer D (1999) Quantitative karyotyping and dual color FISH mapping of 5S and 18S-25S rDNA probes in the cultivated *Phaseolus* species (Leguminosae). Genome 42(6): 1224–1233. https://doi.org/10.1139/g99-070

Paszko B (2006) A critical review and a new proposal of karyotype asymmetry indices. Plant Systematics and Evolution 258(1–2): 39–48. https://doi.org/10.1007/s00606-005-0389-2

Pellerin RJ, Waminal NE, Belandres HR, Kim HH (2018a) Karyotypes of three exotic cucurbit species based on triple-color FISH analysis. Korean Journal of Horticultural Science & Technology 36(3): 417–425. https://doi.org/10.12972/kjhst.20180041

Pellerin RJ, Waminal NE, Kim HH (2018b) Triple-color FISH karyotype analysis of four Korean wild Cucurbitaceae species. Korean Journal of Horticultural Science & Technology 36(1): 98–107. https://doi.org/10.12972/kjhst.20180011

Peruzzi L, Altınordu F (2014) A proposal for a multivariate quantitative approach to infer karyological relationships among taxa. Comparative Cytogenetics 8(4): 337–349. https://doi.org/10.3897/CompCytogen.v8i4.8564

Peruzzi L, Carta A, Altınordu F (2017) Chromosome diversity and evolution in *Allium* (Amaryllidaceae, Allioideae). Plant Biosystems 151(2): 212–220. https://doi.org/10.1080/11263504.2016.1149123

Peruzzi L, Eroglu H (2013) Karyotype asymmetry: again, how to measure and what to measure? Comparative Cytogenetics 7(1): 1–9. https://doi.org/10.3897/compcytogen.v7i1.4431

Peruzzi L, Leitch IJ, Caparelli KF (2009) Chromosome diversity and evolution in Liliaceae. Annals of Botany 103(3): 459–475. https://doi.org/10.1093/aob/mcn230

Plader W, Hoshi Y, Malepszy S (1998) Sequential fluorescent staining with CMA and DAPI for somatic chromosome identification in cucumber [*Cucumis sativus* L.]. Journal of Applied Genetics 39(3): 249–258.

Ramachandran C, Seshadri VS (1986) Cytological analysis of the genome of cucumber (*Cucumis sativus* L.) and muskmelon (*Cucumis melo* L.). Zeitschrift für Pflanzenzuechtung 96(1): 25–38.

Reddy UK, Aryal N, Islam-Faridi N, Tomason YR, Levi A, Nimmakayala P (2013) Cytomolecular characterization of rDNA distribution in various *Citrullus* species using fluorescent *in situ* hybridization. Genetic Resources and Crop Evolution 60(7): 2091–2100. https://doi.org/10.1007/s10722-013-9976-1

Ren Y, Zhang Z, Liu J, Staub JE, Han Y, Cheng Z, Li X, Lu J, Miao H, Kang H, Xie B, Gu X, Wang X, Du Y, Jin W, Huang S (2009) An integrated genetic and cytogenetic map of the cucumber genome. PLoS ONE 4(6): e5795. https://doi.org/10.1371/journal.pone.0005795

Ren Y, Zhao H, Kou Q, Jiang J, Guo S, Zhang H, Hou W, Zou X, Sun H, Gong G, Levi A, Xu Y (2012) A high resolution genetic map anchoring scaffolds of the sequenced watermelon genome. PLoS ONE 7(1): e29453. https://doi.org/10.1371/journal.pone.0029453

Roa F, Guerra M (2012) Distribution of 45S rDNA sites in chromosomes of plants: structural and evolutionary implications. BMC Evolutionary Biology 12: 225. https://doi.org/10.1186/1471-2148-12-225

Schubert I (2007) Chromosome evolution. Current Opinion in Plant Biology 10(2): 109–115. https://doi.org/10.1016/j.pbi.2007.01.001

Schweizer D (1976) Reverse fluorescent chromosome banding with chromomycin and DAPI. Chromosoma 58(4): 307–324. https://doi.org/10.1007/BF00292840

She CW, Jiang XH (2015) Karyotype analysis of *Lablab purpureus* (L.) Sweet using fluorochrome banding and fluorescence *in situ* hybridization with rDNA probes. Czech Journal of Genetics and Plant Breeding 51(3): 110–116. https://doi.org/10.17221/32/2015-CJGPB

She CW, Jiang XH, Ou LJ, Liu J, Long KL, Zhang LH, Duan WT, Zhao W, Hu JC (2015) Molecular cytogenetic characterisation and phylogenetic analysis of the seven cultivated *Vigna* species (Fabaceae). Plant Biology 17(1): 268–280. https://doi.org/10.1111/plb.12174

She CW, Liu JY, Song YC (2006) CPD staining: an effective technique for detection of NORs and other GC-rich chromosomal regions in plants. Biotechnic & Histochemistry 81(1): 13–21. https://doi.org/10.1080/10520290600661414

She CW, Mao Y, Jiang XH, He CP (2020) Comparative molecular cytogenetic characterization of five wild *Vigna* species (Fabaceae). Comparative Cytogenetics 14(2): 243–264. https://doi.org/10.3897/CompCytogen.v14i2.51154

She CW, Wei L, Jiang XH (2017) Molecular cytogenetic characterization and comparison of the two cultivated *Canavalia* species (Fabaceae). Comparative Cytogenetics 11(4): 579–600. https://doi.org/10.3897/compcytogen.v11i4.13604

Siljak-Yakovlev S, Peruzzi L (2012) Cytogenetic characterization of endemics: past and future. Plant Biosystems 146(3): 694–702.

Singh AK, Roy RP (1974) Karyological studies in *Cucumis* (L.). Caryologia 27(2): 153–160. https://doi.org/10.1080/00087114.1974.10796570

Stebbins GL (1971) Chromosomal Evolution in Higher Plants. Addison-Wesley, London, 220 pp.

Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, Zhang J, Zhang H, Gong G, Jia Z, Zhang F, Tian J, Lucas WJ, Doyle JJ, Li H, Fei Z, Xu Y (2017) Karyotype stability and unbiased fractionation in the paleo-allotetraploid *Cucurbita* genomes. Molecular Plant 10(10): 1293–1306. https://doi.org/10.1016/j.molp.2017.09.003

Sun J, Zhang Z, Zong X, Huang S, Li Z, Han Y (2013) A high-resolution cucumber cytogenetic map integrated with the genome assembly. BMC Genomics 14: 461. https://doi.org/10.1186/1471-2164-14-461

Trivedi RN, Roy RP (1970) Cytological studies in *Cucumis* and *Citrullus*. Cytologia 35(4): 561–569. https://doi.org/10.1508/cytologia.35.561

Turkov VD, Shelepina GA, Pushnov MB (1975) The identification of individual chromosome in the caryotype of the cucumber (*Cucumis sativus* L.) on the base of their linear differentiation. Cytologia 40(1): 31–34. https://doi.org/10.1508/cytologia.40.31

Urasaki N, Takagi H, Natsume S, Uemura A, Taniai N, Miyagi N, Fukushima M, Suzuki S, Tarora K, Tamaki M, Sakamoto M, Terauchi R, Matsumura H (2017) Draft genome sequence of bitter gourd (*Momordica charantia*), a vegetable and medicinal plant in tropical and subtropical regions. DNA Research 24(1): 51–58. https://doi.org/10.1093/dnares/dsw047

Waminal NE, Kim HH (2012) Dual-color FISH karyotype and rDNA distribution analyses on four Cucurbitaceae species. Horticulture Environment & Biotechnology 53(1): 49–56. https://doi.org/10.1007/s13580-012-0105-4

Waminal NE, Kim HH (2015) FISH Karyotype analysis of four wild Cucurbitaceae species using 5S and 45S rDNA probes and the emergence of new polyploids in *Trichosanthes kirilowii* Maxim. Korean Journal of Horticultural Science & Technology 33(6): 869–876. https://doi.org/10.7235/hort.2015.15101

Waminal NE, Kim NS, Kim HH (2011) Dual-color FISH karyotype analyses using rDNAs in three Cucurbitaceae species. Genes & Genomics 33(5): 521–528. https://doi.org/10.1007/s13258-011-0046-9

Weiss-Schneeweiss H, Tremetsberger K, Schneeweiss GM, Parker JS, Stuessy TF (2008) Karyotype diversification and evolution in diploid and polyploid South American *Hypochaeris* (Asteraceae) inferred from rDNA localization and genetic fingerprint data. Annals of Botany 101(7): 909–918. https://doi.org/10.1093/aob/mcn023

Wu H, Zhao G, Gong H, Li J, Luo C, He X, Luo S, Zheng X, Liu X, Guo J, Chen J, Luo J (2020) A high-quality sponge gourd (*Luffa cylindrica*) genome. Horticulture Research 7(1): 128. https://doi.org/10.1038/s41438-020-00350-9

Xie D, Xu Y, Wang J, Liu W, Zhou Q, Luo S, Huang W, He X, Li Q, Peng Q, Yang X, Yuan J, Yu J, Wang X, Lucas WJ, Huang S, Jiang B, Zhang Z (2019a) The wax gourd genomes offer insights into the genetic diversity and ancestral cucurbit karyotype. Nature Communication 10(1): 5158. https://doi.org/10.1038/s41467-019-13185-3

Xie WJ, Huang J, Ma J (2019b) Localization of 45S and 5S rDNA sequences on chromosomes of 20 species of Cucurbitaceous plants. Journal of South China Agricultural University 40(6): 74–81. http://dx.doi.org/10.7671/j.issn.1001-411X.201811048

Xu YH, Yang F, Cheng YL, Ma L, Wang JB, Li LJ (2007) Comparative analysis of rDNA distribution in metaphase chromosomes of Cucurbitaceae species. Hereditas (Beijing) 29(5): 614–620. https://doi.org/10.1360/yc-007-0614

Yagi K, Pawełkowicz M, Osipowski P, Siedlecka E, Przybecki Z, Tagashira N, Hoshi Y, Malepszy S, Pląder W (2015) Molecular cytogenetic analysis of *Cucumis* wild species distributed in southern Africa: physical mapping of 5S and 45S rDNA with DAPI. Cytogenetic and Genome Research 146(1): 80–87. https://doi.org/10.1159/000433572

Zhang X, Zhou T, Yang J, Sun J, Ju M, Zhao Y, Zhao G (2018) Comparative analyses of chloroplast genomes of Cucurbitaceae species: lights into selective pressures and phylogenetic relationships. Molecules 23(9): 2165. https://doi.org/10.3390/molecules23092165

Zhang Y, Cheng C, Li J, Yang S, Wang Y, Li Z, Chen J, Lou Q (2015a) Chromosomal structures and repetitive sequences divergence in *Cucumis* species revealed by comparative cytogenetic mapping. BMC Genomics 16(1): 730. https://doi.org/10.1186/s12864-015-1877-6

Zhang YX, QF Lou, Li ZA, Wang YZ, Zhang ZT, Li J, Chen JF (2015b) Rapid karyotype analysis of cucumber varieties based on genomic *in situ* hybridization. Scientia Agricultura Sinica 48(2): 398–406. http://dx.doi.org/10.3864/j.issn.0578-1752.2015.02.20

Zhang ZT, Yang SQ, Li ZA, Zhang YX, Wang YZ, Cheng CY, Li J, Chen JF, Lou QF (2016) Comparative chromosomal localization of 45S and 5S rDNAs and implications for genome evolution in *Cucumis*. Genome 59(7): 449–457. https://doi.org/10.1139/gen-2015-0207

Zhao Q, Meng Y, Wang P, Qin X, Cheng C, Zhou J, Yu X, Li J, Lou Q, Jahn M, Chen J (2021) Reconstruction of ancestral karyotype illuminates chromosome evolution in the genus *Cucumis*. The Plant Journal 107(4): 1243–1259. https://doi.org/10.1111/tpj.15381

Zhao X, Lu J, Zhang Z, Hu J, Huang S, Jin W (2011) Comparison of the distribution of the repetitive DNA sequences in three variants of *Cucumis sativus* reveals their phylogenetic relationships. Journal of Genetics and Genomics 38(1): 39–45. https://doi.org/10.1016/j.jcg.2010.12.005

ORCID

Chao-Wen She https://orcid.org/0000-0003-1935-5509
Xiang-Hui Jiang https://orcid.org/0000-0003-0923-210X

Supplementary material 1

The plant materials

Authors: Chao-Wen She, Xiang-Hui Jiang, Chun-Ping He

Data type: table (docx. file)

Link: https://doi.org/10.3897/compcytogen.v17.i1.99236.suppl1

Supplementary material 2

CPD-stained mitotic metaphase chromosomes with the maximum condensation degree

Authors: Chao-Wen She, Xiang-Hui Jiang, Chun-Ping He

Data type: figure (docx. file)

Explanation note: CPD-stained mitotic metaphase chromosomes with the maximum condensation degree from *C. sativus* (**A**), *C. melo* (**B**), *C. lanatus* (**C**), *B. hispida* (**D**), *M. charantia* (**E**), *L. cylindrica* (**F**), *L. siceraria* var. *hispida* (**G**) and *C. moschata* (**H**). Scale bars: 10 μm.

Link: https://doi.org/10.3897/compcytogen.v17.i1.99236.suppl2

Supplementary material 3

Chromosome measurements of the eight Cucurbitaceae crops obtained from five metaphases per species

Authors: Chao-Wen She, Xiang-Hui Jiang, Chun-Ping He

Data type: table (docx. file)

Link: https://doi.org/10.3897/compcytogen.v17.i1.99236.suppl3

Supplementary material 4

The correlational analysis between the difference in TCL and the change in nuclear DNA content within the eight Cucurbitaceae crops using the SPSS 25.0 software

Authors: Chao-Wen She, Xiang-Hui Jiang, Chun-Ping He

Data type: table (docx. file)

Link: https://doi.org/10.3897/compcytogen.v17.i1.99236.suppl4

Supplementary material 5

The number and position of 45S rDNA locus in Cucurbitaceae species

Authors: Chao-Wen She, Xiang-Hui Jiang, Chun-Ping He

Data type: table (docx. file)

Link: https://doi.org/10.3897/compcytogen.v17.i1.99236.suppl5

SHORT COMMUNICATION

Assessing ploidy levels and karyotype structure of the fire ant *Solenopsis saevissima* Smith, 1855 (Hymenoptera, Formicidae, Myrmicinae)

Ananda Ribeiro Macedo de Andrade[1], Danon Clemes Cardoso[1],
Maykon Passos Cristiano[1]

1 *Genetics and Evolution of Ants Research Group, Universidade Federal de Ouro Preto, Ouro Preto, MG – 35400-000, Brazil*

Corresponding author: Danon Clemes Cardoso (danon@ufop.edu.br); Maykon Passos Cristiano (maykon@ufop.edu.br)

Citation: de Andrade ARM, Cardoso DC, Cristiano MP (2023) Assessing ploidy levels and karyotype structure of the fire ant *Solenopsis saevissima* Smith, 1855 (Hymenoptera, Formicidae, Myrmicinae). Comparative Cytogenetics 17: 59–73. https://doi.org/10.3897/compcytogen.17.100945

Abstract

The family Formicidae is composed of ants that organize themselves into castes in which every individual has a joint organizational function. *Solenopsis* Westwood, 1840 is an ant genus with opportunistic and aggressive characteristics, known for being invasive species and stings that cause burning in humans. This genus is particularly difficult to classify and identify since its morphology provides few indications for species differentiation. For this, a tool that has been useful for evolutionary and taxonomic studies is cytogenetics. Here, we cytogenetically studied *Solenopsis saevissima* Smith, 1855 from Ouro Preto, Minas Gerais, Brazil. We evaluated the occurrence of polyploid cells in individuals and colonies by conventional cytogenetics. A total of 450 metaphases were analyzed and counted. Chromosome counts of individuals and colonies showed varied numbers of ploidies, from n = 16 to 8n = 128. The karyomorphometrical approach allowed determination of the following karyotypes: n = 10 m + 4 sm + 2 st, 2n = 20 m + 8 sm + 4 st, and 4n = 40 m + 16 sm + 8 st. Polyploidy can be found naturally in individuals and colonies and may represent an adaptative trait related to widespread distribution and invasion ability of new habitats.

Keywords

Evolution, Fire ant, Invasive species, Karyotype, Polyploidy

Introduction

Ants are recognized as some of the most successful organisms among invertebrates, being widely distributed throughout the world (Hölldobler and Wilson 1990). Formicidae is a clade that includes all ants and is the only family in which all species have the characteristic known as eusociality. This means that colonies are organized in castes, exhibit division of labor with overlapping generations (Wilson 1990). The queen, or sometimes queens in polygynous species, is responsible for reproduction, while the workers build the nest, defend the colony, and are responsible for obtaining and handling resources (Wilson 1998). The division of labor can also be related to the morphology of each worker, in which the sizes and ages of the workers will define which function they perform within the nest (Haight 2010).

Myrmicinae is the most diverse subfamily and includes the genus *Solenopsis* Westwood, 1840, which are known as "fire ants". This popular name based on their aggressiveness and painful sting, which is due to the accumulation of allergenic proteins and alkaloids in their venom (Fox et al. 2010). They are native to South America (Buren 1972) but have great potential for habitat invasion. Some species of *Solenopsis* are currently found in Central America, North America, and Oceania (Callcott and Collins 1996; Holway et al. 2002).

Although ants are essential organisms within their ecosystems as they participate in maintaining the soil, nutrient cycling, and other ecosystem services (Lobry De Bruyn 1999), members of the genus *Solenopsis* are also responsible for great damage, both in agriculture and their effects on humans and animals, and can be considered pests (Santos 2016). The species *Solenopsis saevissima* Smith, 1855 is responsible for 35% of the reports of insect bites (Fox et al. 2012). They are also invasive, and when fire ants arrive in a new environment, they become harmful to other native species, which can be removed by competition due to the aggressiveness imposed by them (Wojcik et al. 2001).

The genus *Solenopsis* comprises more than 190 described species worldwide. They are cosmopolitan and taxonomically difficult. According to Fox et al. (2010), Pitts et al. (2005), and Shoemaker et al. (2006), workers lack morphological features for precise classification, and the morphological differences in some groups are not easily perceptible. In this context, cytogenetic and molecular data can provide useful markers for the systematics and taxonomy of this ant group.

Cytogenetics is a field of study interested in understanding the structure and function of the chromosomes (Speicher and Carter 2005). How the genome of an organism is organized into a defined number of DNA molecules is one of the most basic pieces of information that is reflected by the karyotype of the species. Thus, cytogenetic studies are relevant for evolutionary and taxonomic knowledge since the analysis of karyotypes can help distinguish species, and therefore complement phylogenetic and evolutionary analyses (Lukhtanov et al. 2006; Lorite and Palomeque 2010). For instance, a particular example of how cytogenetics can be used in the taxonomy of ants is the genus *Amoimyrmex* Cristiano, Cardoso et Sandoval, 2020 (Cristiano et al. 2020),

i.e., a new genus of leaf-cutting ants discovered by integrating cytogenetics, molecular genetics, and morphology. Numerous other taxonomic issues for which cytogenetics could be useful are still to be addressed. Today, only 7% of ants have been cytogenetically analyzed (Cardoso et al. 2018a), which represents less than 1,000 species from more than 16,000 species known so far (Bolton 2022).

Even considering the small number of species cytogenetically studied, ants show an extreme karyotype diversity varying from the haploid number n = 1 (Crosland and Crozier 1986) to n = 60 chromosomes (Mariano et al. 2008). Considering only *Solenopsis*, two main karyotypes were recovered, n = 11 and n = 16 chromosomes (Cardoso et al. 2018a). Wurm et al. (2011) presented information about the genome of *Solenopsis invicta* Buren, 1972, making it possible to understand its genomic structure, such as the identification of gene duplications and the multifunctionality of vitellogenin genes. Polyploid cells were already reported in insects and were suggested to be regulated by the endoreplication system (Fox and Duronio 2013). Endoreplication is a process that results in polytene chromosomes that have thousands of DNA strands. Polyploid organisms are common in plants (Otto and Whitton 2000), however they are rare in animals (White 1973; Clark and Wall 1996). Nevertheless, polyploidy is a heritable condition where an organism possess more than two complete sets of chromosomes. Polyploid cells can be identified through cytogenetic evidence and further confirmed by flow cytometry (FCM).

Polyploidy in ants has already been reported, but the studies do not describe whether and how the karyotype varies within the colony. In the present study, we describe the karyotype of the species *Solenopsis saevissima* from Brazil and evaluate whether and how the karyotype varies within individuals and the colony. We also perform a karyomorphometric analysis to precisely determine the karyotype structure and provide quantitative data for *S. saevissima* chromosomes. Additionally, we used flow cytometry analysis to determine the ploidy level of brain cells of *S. saevissima*. These data will certainly help our understanding of the ant's genome evolution, taxonomy, and systematics.

Material and methods

Species sampling

Solenopsis saevissima colonies were sampled in Ouro Preto, Minas Gerais, Brazil (20°17'15"S, 43°30'29"W) located in the southeast region at over 1,150 m of altitude. Sampling occurred from October to December 2020, the period when broods were available. The nests were identified according to the description by Porter and Tschinkel (1987), who mentioned nests as mounds of soil located in grassy, sunny, open areas (Fig. 1). We marked the colonies as 1–4. The ants were collected with the aid of gloves and a shovel, stored in a plastic container, and taken to the laboratory for further processing. We never collected the entire colony, allowing the brood to recover.

Figure 1. *Solenopisis saevissima* mound located in a grassy field in the campus of Morro do Cruzeiro, Universidade Federal de Ouro Preto, Ouro Preto - MG, Brazil. Scale bar: 3 cm.

Sample preparation and obtention of mitotic cells

The colony fractions of *S. saevissima* were taken directly to the laboratory, and while alive, the post-defecating larvae (without meconium; or pre-pupae) were isolated. As described by Imai et al. (1988) and detailed by Cardoso et al. (2017), the cerebral ganglion of the larvae was removed and transferred to a container containing hypotonic colchicine solution (0.005% w/v colchicine in 1% sodium citrate solution) and incubated for 60 min in the dark. The time of incubation was adjusted considering the frequency of metaphases and standard condensation pattern (see Cristiano et al. 2017). Therefore, the ganglia were placed on a slide and smashed until smooth with the aid of two needles to release the cells. Metaphase spreads were obtained by dropping solutions on smashed tissue serially: first, solution 1 (acetic acid:ethanol:distilled water; 3:3:4), followed by solution 2 (acetic acid:ethanol; 1:1), and finally solution 3 (acetic acid 100%). After air drying, the slides were labeled with the respective colony code.

The slides were stained with Giemsa (4%) to observe the chromosomes under an optical microscope. Metaphases were photographed using a Zeiss Axio Imager Z2 microscope coupled to an AxioCam MRc image capture system. A total of 450 photos were captured of the metaphases found on the slides from the four different colonies (N1, N3, N4, and N5). The number of chromosomes was counted in all captured photos. A minimum of ten well-spread haploid (n) (males) and diploid metaphases (2n)

(females) were assembled and submitted to karyomorphometrical analysis according to the description by Cristiano et al. (2017). For each chromosome, the total length (TL), short arm (S), and long arm (L) were measured, calculated as the distance between the arm telomere and the centromere. The total length (KL) of the karyotype was calculated from the sum of the total length (TL) of all chromosomes. The relative size (RL) was calculated in relation to the total size of all chromosomes with the formula (TL × 100 / ∑TL). The ratio (*r*) between the length of the long arm and short arm was given by the formula (*r* = L / S) and used to classify the chromosomes as metacentric (m), submetacentric (sm), and subtelocentric (st) as described by Levan et al. (1964).

Genome size (in picograms, pg) was estimated by flow cytometry in individuals from the four colonies following the protocol established by Moura et al. (2020). Cerebral ganglia of the post-defecating larvae from workers and the internal standard (*Drosophila melanogaster*) were detached and immersed in 100–300 μL of Galbraith buffer and ground to release the cell nuclei. Subsequently, 600 μL of the buffer was added, filtered through a 40 μm nylon mesh, and stained by adding 6.5 μL of propidium iodide solution and 3.5 μl RNAse and analyzed after 15 min. The analyses were performed on a FACSCalibur (BD Biosciences, San José, USA) cytometer at Universidade Federal de Ouro Preto, equipped with a laser source (488 nm) and the histograms were obtained by the BD Cell Quest software. For each sample, at least 10,000 nuclei were analyzed regarding their relative fluorescence intensity. Three independent replicates (three individuals per colony) were conducted and histograms with a coefficient of variation above 5% were rejected. Histograms were analyzed using the Flowing 2.5.1 software (http://www.flowingsoftware.com). The genome size of each *S. saevissima* was calculated using the 1C-value (0.18 pg) of *Drosophila melanogaster* and the values were obtained according to the equation by Doležel and Bartoš (2005) and converted to megabase pairs (1 pg = 978 Mbp).

Results

The chromosome counts for the *S. saevissima* individuals analyzed here were n = 16 (22 metaphases), 2n = 32 (122 metaphases), 4n = 64 (26 metaphases), and 8n = 128 (a single metaphase) (Fig. 2) considering all colonies. Our observations confirm that we can commonly find polyploid cells in the brain ganglion of immatures of *S. saevissima*. All counts (n = 452) are summarized in Fig. 3, showing the distribution of metaphases around the modal chromosome numbers 16, 32, and 64. The karyomorphometric data from haploid and diploid karyotypes are given in Suppl. material 1: tables S1, S2. Accurate karyomorphometric analysis from the polyploid metaphases was unlikely, but a particular 4n metaphase was evaluated and the resulting measurements are given in Suppl. material 1: table S3. The karyotypic formulas found were n = 10m + 4sm + 2st, 2n = 20m + 8sm + 4st, and 4n = 40m + 16sm + 8st. The two largest metacentric and submetacentric chromosome pairs showed secondary constrictions. Polyploid cells were observed in all colonies at the similar frequency (Suppl. material 1: table S4).
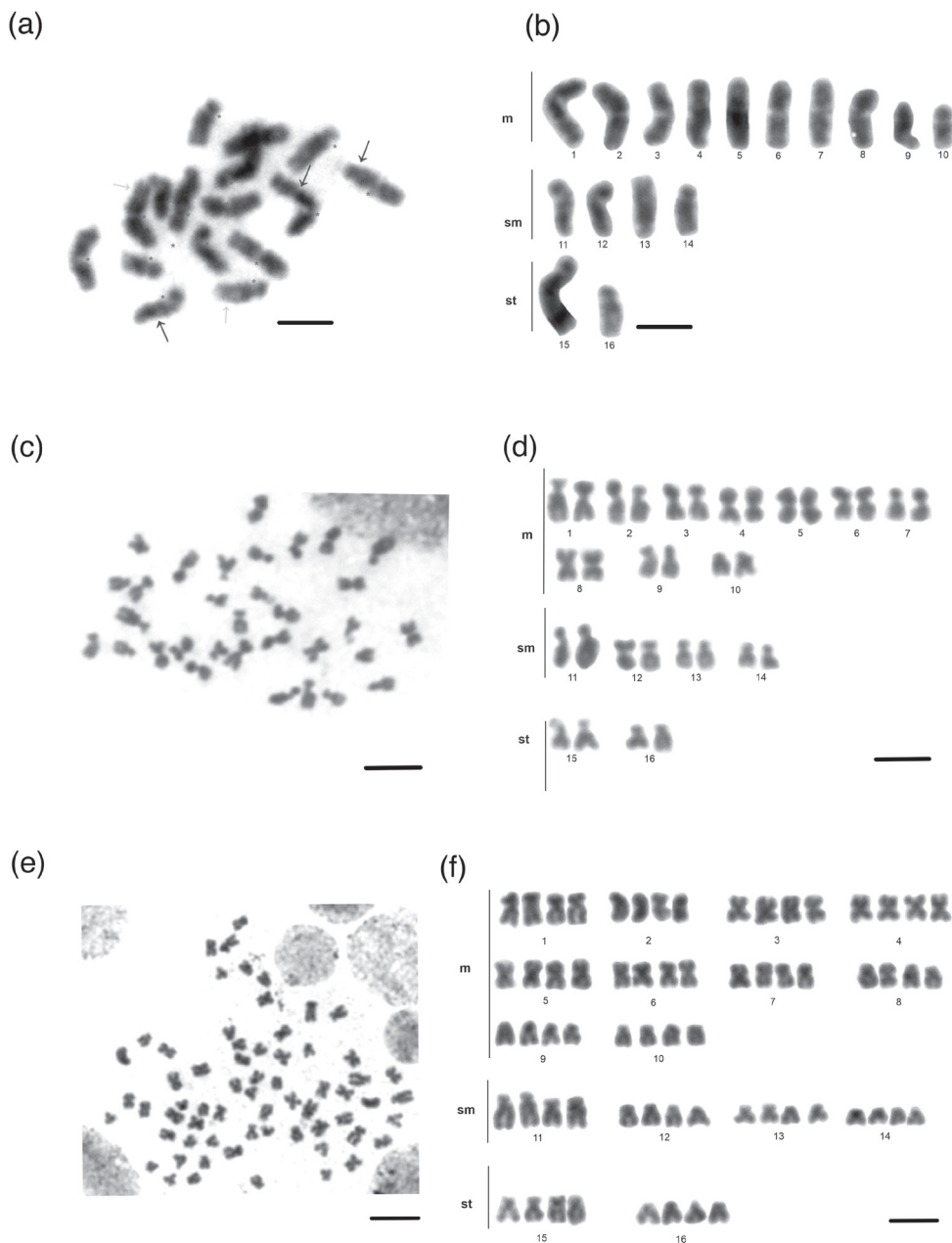
Figure 2. Chromosomes of *Solenopisis saevissima* **a** metaphase **b** haploid karyotype; n = 16 **c** metaphase **d** diploid karyotype, 2n = 32 **e** metaphase; and **f** tetraploid karyotype, 4n = 64. Asterisks, grey and black arrows indicate centromeres as well as smaller and larger heterochromatic segments respectively. Scale bars: 10 μm.

The nuclei isolated from the brain tissue were properly recovered given the histograms showing peaks from cells at different stages of the cell cycle: the higher peak G0/G1 (unreplicated DNA in the nuclei – 2C) and lower peak G2 (replicated DNA – 4C).
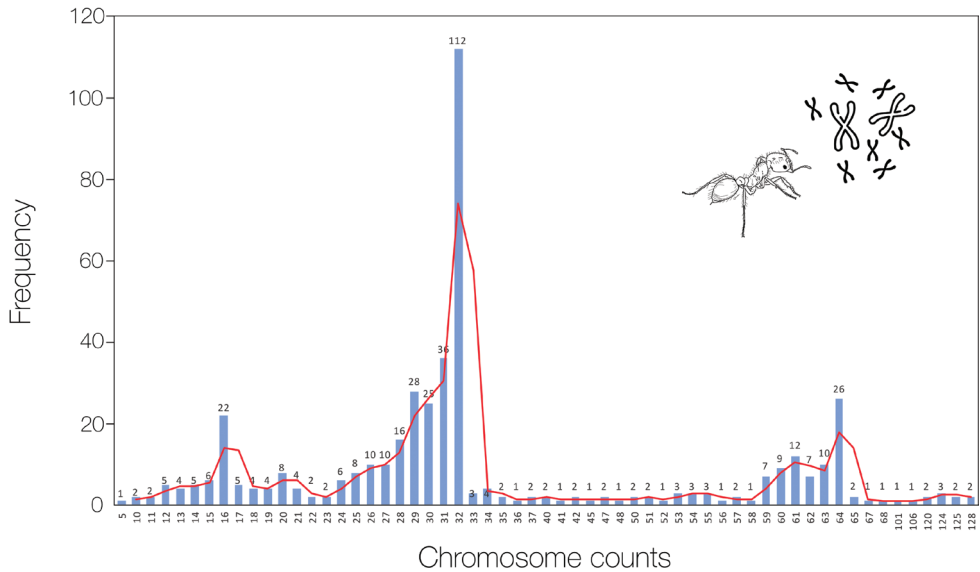
Figure 3. Chromosome count frequency of *Solenopisis saevissima* throughout all 452 metaphases. The highest frequency was observed in the modal haploid (n = 16) and diploid (2n = 32) karyotypes together with the less frequent 4n = 64. The red line represents the tendency curve. Variations are due to the technique employed to obtain mitotic chromosomes.

Additional peaks were observed after the common G0/G1 and G2 peaks for which the nuclei occupy a well-defined series of regions, equally spaced in terms of fluorescence and corresponding to 8C and 16C nuclei (see Fig. 4). The population of nuclei declines from 2C to 16C, representing the other ploidies observed both for the internal standard *D. melanogaster* as well for *S. saevissima*, indicating endoreduplication or polyploid cells as expected (Fig. 4). The genome size of *S. saevissima* was 0.51 ± 0.015 pg or 498.78 Mbp.

Discussion

Here we observed a chromosome number variation in *S. saevissima* from n = 16 to 8n = 128 chromosomes. These counts agree with previous descriptions (Murakami et al. 2021). The typical chromosome number recovered from other *Solenopsis* species, such as *Solenopsis geminata* (Fabricius, 1804), *Solenopsis richteri* Forel, 1909, and *Solenopsis invicta* is n = 16 (Cardoso et al. 2018a), which suggests that the chromosome number of n = 16 was the regular count of the haploid karyotype of *S. saevissima*. The other description from Uruguay also reported the same chromosome number (Goñi et al. 1983). The genome size estimates agree with previous data (Moura et al. 2021), and the 2C, 4C, and 8C values were clearly recovered by our flow cytometry analysis. Here we demonstrated that ploidy of cells varies among individuals within the colonies. Polyploid cells have been reported in other ants (see the reviews by Crozier 1975; Imai et al. 1977; and Lorite and Palomeque 2010), but not often. For example, although regularly studied
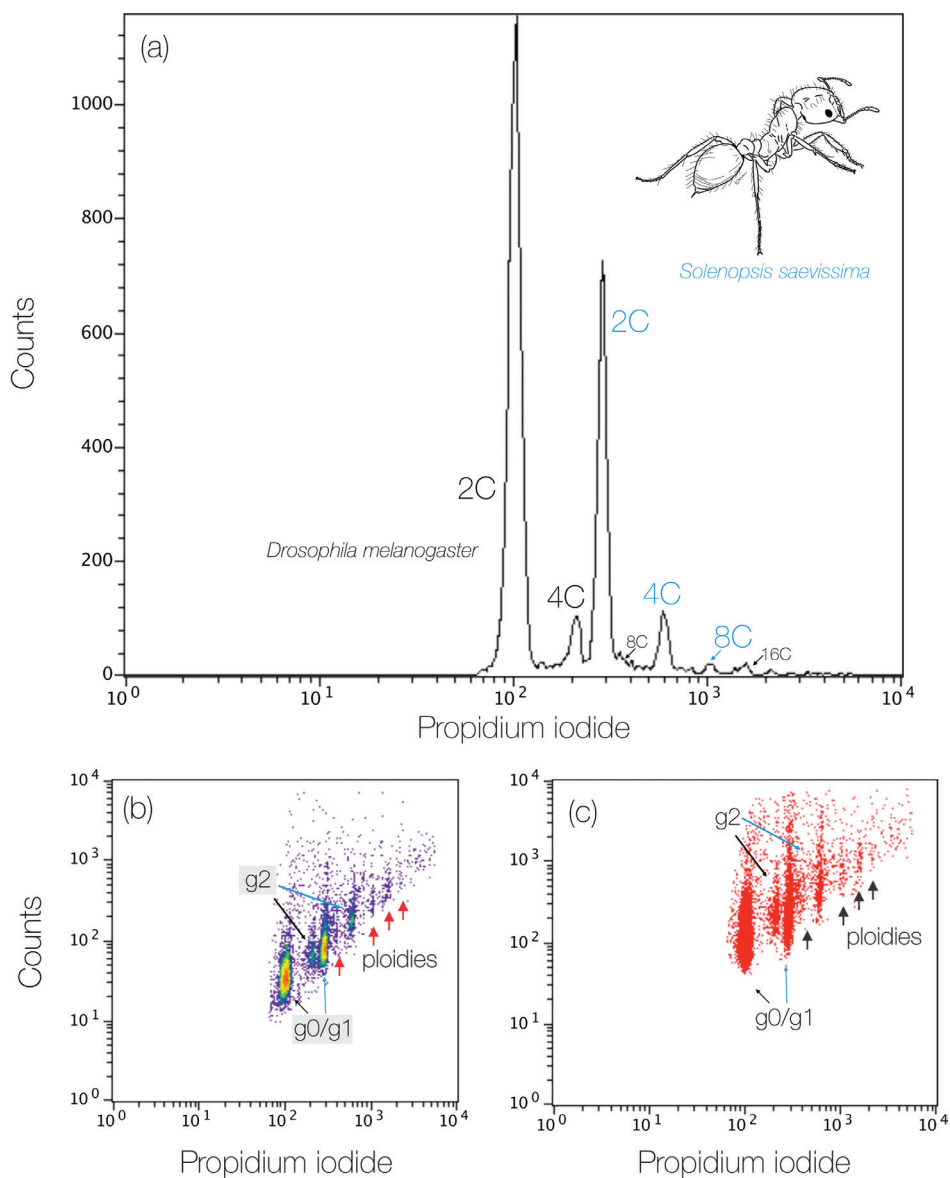
Figure 4. Genome size of *Solenopisis saevissima* showing ploidy variations estimated by flow cytometry **a** histogram highlighting the peaks from 2C to 8C (blue lettering refers to *S. saevissima* and black lettering refers to the internal standard) **b** density plot **c** dot plot containing many events, within which the nuclei occupy a well-defined series of regions, equally spaced in terms of fluorescence and corresponding to 2C, 4C, 8C, and 16C nuclei. Colors in the density plot indicate nuclei population density, with red as the highest and blue as the lowest.

from a cytogenetic point of view, polyploidy was not evidenced among fungus-farming ants (Cardoso and Cristiano 2021) or recovered by flow-cytometry studies (Moura et al. 2021), suggesting that polyploid cells may be restricted to some ant linages and not

widespread within Formicidae. However, the lack of polyploid records may be due to the small coverage of cytogenetic studies and the high diversity of ant species. In her doctoral thesis, Silva (2016) used flow cytometry to demonstrate that there is a reversion of the presence of polyploid cells throughout the developmental stages of *S. saevissima* from larvae to pupae and adult workers, suggesting that polyploid cells occur only in the immature phase. This is expected since ants are holometabolous insects that do not change after metamorphosis. Thus, we hypothesized that the presence of polyploid cells in the immature and mutable stages may contribute to the colony's fitness advantage.

The polyploid cells observed in the brood phase of *S. saevissima* may promote some benefit resulting in the developmental rate of the immature workers, which in turn will result in the number of workers. This could be analogous to the way polyploid cells occur in the salivary tissue of dipterans (Rodman 1967; Wells and Andrew 2015), who depend on salivary secretions to feed. The colony can grow faster and exploit available resources by reaching maturity for reproduction. *Solenopsis saevissima* is a recruiting species, and workers signal and convene other workers at resources (Yong-Yue et al. 2012), which can be exploited faster and invested into growing the colony and ultimately sexuals for reproduction (see Peeters and Molet 2010 for ant colony life cycle details). The faster a colony grows and exploits the environment to produce sexuals that will establish new colonies, the higher the fitness. Here, ploidy is indicated as a potential cytogenetic feature that allows *S. saevissima* to spread and colonize new areas, but such an idea requires experimental testing in future field studies.

Considering the stage where polyploid cells were found, it apparently results from endomitosis, which consists of normal G1 and G2 phases, but with incomplete mitosis. This means that the cytokinesis step does not occur at the end of the cell cycle, the chromosomes accumulate, thus generating polyploid cells (Lee et al. 2008). Indeed, studies in animal and even plant developmental systems have revealed conserved mechanisms that control the generation of polyploidy, and a reasonable expectation is that polyploid cells, through endoreplication, may provide key biological functions during developmental stages (Fox and Duronio 2013).

A recent study on *Solenopsis* by Murakami et al. (2021) compared species in native and invaded areas. Their results showed differences in chromosomal morphology between the analyzed populations, mainly in ploidy, suggesting a possible generalized hybridization between ants native to South and North America. Evidence of hybridization in this genus has already been reported by Taber and Cokendolpher (1988) and Ross and Shoemaker (2005). The former suggests that species in the US can hybridize with *S. invicta*, *S. geminata*, and *S. molesta* (Say, 1836). Hybridization in genetically close species can generate disarrangements in the cytoplasm, duplicating the genome and consequently resulting in polyploidy (Fujiwara et al. 1997).

Based on cytogenetic evidence, Murakami et al. (2021) suggested that invasive *Solenopsis* species, when settling in new environments, hybridize with closely related, or even genetically distant species. This process resulted in various chromosome numbers. Such a mechanism may promote an increase in the genetic diversity of the population and the acquisition of adaptive genes that will better acclimate species to the invaded environment (Chen 2010).

Our study complements the importance of understanding the chromosomal biology of ants. This approach can also help understand species' life histories and contributes to the analysis of invasive species. Here, we found cytogenetic evidence that may reflect the species' biology. *Solenopsis* ants are aggressive competitors, opportunistic scavengers nesting in open areas in urban and natural preserved environments (Lofgren et al. 1975) and are well-adapted to anthropized areas.

The external morphologies of *S. saevissima* and its congeners do not provide suitable traits to recognize potential cryptic species (Fox et al. 2012). Thus, karyotyping determines the number and morphology of chromosomes, proving to be a good tool for understanding genetic barriers within inconspicuous groups (Cristiano et al. 2017; Cardoso et al. 2018b). In the present study, cytogenetic analysis of *Solenopsis saevissima* yielded the same chromosome number, which was observed previously. Further, it appears that a chromosome number of n = 16 is a common karyotype feature of *Solenopsis* spp.

Data availability statement

All relevant data are within the manuscript and its Supporting Information files (Suppl. material 1). All other information can be requested from the corresponding authors.

Acknowledgements

References

Bolton B (2022) An online catalog of the ants of the world. https://antcat.org [Accessed 12.07.2022]

Buren WF (1972) Revisionary studies on the taxonomy of the imported fire ants. Journal of the Georgia Entomological Society 7: 1–26.

Callcott AA, Collins HL (1996) Invasion and range expansion of imported fire ants (Hymenoptera: Formicidae) in North America from 1918–1995. Florida Entomologist 79: 240–251. https://doi.org/10.2307/3495821

Cardoso DC, Cristiano MP (2021) Karyotype diversity, mode, and tempo of the chromosomal evolution of Attina (Formicidae: Myrmicinae: Attini): Is there an upper limit to chromosome number? Insects 12(12): 1084. https://doi.org/10.3390/insects12121084

Cardoso DC, Heinze J, Moura MN, Cristiano MP (2018b) Chromosomal variation among populations of a fungus-farming ant: implications for karyotype evolution and potential restriction to gene flow. BMC Evolutionary Biology 18(1): 146. https://doi.org/10.1186/s12862-018-1247-5

Cardoso DC, Pereira TTP, Cordeiro AL, Cristiano MP (2017) Cytogenetic data on the agro-predatory ant *Megalomyrmex incisus* Smith, 1947 and its host, *Mycetophylax conformis* (Mayr, 1884) (Hymenoptera, Formicidae). Comparative Cytogenetics 11(1): 45–53. https://doi.org/10.3897/CompCytogen.v11i1.10842

Cardoso DC, Santos HG, Cristiano MP (2018a) The Ant Chromosome database - ACdb: an online resource for ant (Hymenoptera: Formicidae) chromosome researchers. Myrmecological News 27: 87–91. https://doi.org/10.25849/myrmecol.news_027:087

Chen ZJ (2010) Molecular mechanisms of polyploidy and hybrid vigor. Trends in Plant Science 15(2): 57–71. https://doi.org/10.1016/j.tplants.2009.12.003

Clark MS, Wall WJ (1996) Chromosome form and function. In: Clark MS, Wall WJ (Eds) Chromosomes: the complex code. Chapman and Hall, London, 61–65. https://doi.org/10.1007/978-94-009-0073-8_2

Cristiano MP, Cardoso DC, Fernandes-Salomão TM (2013) Cytogenetic and molecular analyses reveal a divergence between *Acromyrmex striatus* (Roger, 1863) and other congeneric species: taxonomic implications. PLoS ONE 8(3): e59784. https://doi.org/10.1371/journal.pone.0059784

Cristiano MP, Cardoso DC, Sandoval-Gómez VE, Simões-Gomes FC (2020) *Amoimyrmex* Cristiano, Cardoso and Sandoval, gen. nov. (Hymenoptera: Formicidae): a new genus of leaf-cutting ants revealed by multilocus molecular phylogenetic and morphological analyses. Austral Entomology 59: 643–676. https://doi.org/10.1111/aen.12493

Cristiano MP, Pereira TTP, Simões LP, Sandoval-Gómez VE, Cardoso DC (2017) Reassessing the chromosome number and morphology of the turtle ant *Cephalotes pusillus* (Klug, 1824) using karyomorphometrical analysis and observations of new nesting behavior. Insects 8(4): 114. https://doi.org/10.3390/insects8040114

Crosland MW, Crozier RH (1986) *Myrmecia pilosula*, an ant with only one pair of chromosomes. Science 231: 1278. https://doi.org/10.1126/science.231.4743.1278

Crozier R (1975) Hymenoptera. In: John B (Ed.) Animal Cytogenetics, vol. 3. Insecta 7. Gebrüder Borntraeger, Berlin, Stuttgart, 95 pp.

Doležel J, Bartoš J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. Annals of Botany 95: 99–110. https://doi.org/10.1093/aob/mci005

Fox EGP, Bueno OC, Yabuki AT, de Jesus CM, Solis DR, Rossi ML, Nogueira NL (2010) General morphology and ultrastructure of the venom apparatus and convoluted gland of the fire ant, *Solenopsis saevissima*. Journal of Insect Science 10(1): 24. https://doi.org/10.1673/031.010.2401

Fox DT, Duronio RJ (2013) Endoreplication and polyploidy: Insights into development and disease. Development 140: 3–12. https://doi.org/10.1242/dev.080531

Fox EGP, Pianaro A, Solis DR, Delabie JHC, Vairo BC, Machado EA, Bueno OC (2012) Intraspecific and intracolonial variation in the profile of venom alkaloids and cuticular hydrocarbons of the fire ant *Solenopsis saevissima* Smith (Hymenoptera: Formicidae). Psyche: A Journal of Entomology 2012: 398061. https://doi.org/10.1155/2012/398061

Fujiwara A, Abe S, Yamaha E, Yamazaki F, Yoshida MC (1997) Uniparental chromosome elimination in the early embryogenesis of the inviable salmonid hybrids between masu salmon female and rainbow trout male. Chromosoma 106(1): 44–52. https://doi.org/10.1007/s004120050223

Goñi B, De Zolessi LC, Imai HT (1983) Karyotypes of thirteen ant species from Uruguay (Hymenoptera, Formicidae). Caryologia 36: 363–371. https://doi.org/10.1080/00087114.1983.10797677

Haight KL (2010) Worker size and nest defense in *Solenopsis invicta* (Hymenoptera: Formicidae). Annals of the Entomological Society of America 103(4): 678–682. https://doi.org/10.1603/AN09161

Hölldobler B, Wilson, EO (1990) The ants. Harvard University Press, 732 pp. https://doi.org/10.1007/978-3-662-10306-7

Holway DA, Lach L, Suarez AV, Tsutsui ND, Case TJ (2002) The causes and consequences of ant invasions. Annual Review of Ecology and Systematics 33: 181–233. https://doi.org/10.1146/annurev.ecolsys.33.010802.150444

Imai HT, Crozier RH, Taylor RW (1977) Karyotype evolution in Australian ants. Chromosoma 59: 341–393. https://doi.org/10.1007/BF00327974

Imai HT, Taylor RW, Crozier RH (1988) Modes of spontaneous chromosomal mutation and karyotype evolution in ants with reference to the minimum interaction hypothesis. The Japanese Journal of Genetics 63: 159–185. https://doi.org/10.1266/jjg.63.159

Lee G, Chung SJ, Park IS, Lee JS, Kim J, Kim DS, Kang S (2008) Variation in the phenotypic features and transcripts of color mutants of chrysanthemum (*Dendranthema grandiflorum*) derived from gamma ray mutagenesis. Journal of Plant Biology 51: 418–423. https://doi.org/10.1007/BF03036063

Levan A, Fredga K and Sandberg AA (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201–220. https://doi.org/10.1111/j.1601-5223.1964.tb01953.x

Lobry De Bruyn, LA (1999) Ants as bioindicators of soil function in rural environments. Agriculture, Ecosystems and Environment 74: 425–441. https://doi.org/10.1016/S0167-8809(99)00047-X

Lofgren CS, Banks WA, Glancey BM (1975) Biology and control of imported fire ants. Annual Review of Entomology 20: 1–30. https://doi.org/10.1146/annurev.en.20.010175.000245

Lorite P, Palomeque T (2010) Karyotype evolution in ants (Hymenoptera: Formicidae), with a review of the known ant chromosome numbers. Myrmecological News 13: 89–102.

Yong-Yue L, Bi-Qiu W, Ling Z, Yi-Juan X (2014) Comparison of Foraging Ability Between *Solenopsis invicta* and *Tapinoma melanocephalum* (Hymenoptera: Formicidae). Sociobiology 59(3): 1015–1024.

Lukhtanov VA, Dincă V, Talavera G, Vila R (2011) Unprecedented within-species chromosome number cline in the Wood White butterfly *Leptidea sinapis* and its significance for karyotype evolution and speciation. BMC Evolutionary Biology 11: 109. https://doi.org/10.1186/1471-2148-11-109

Lukhtanov VA, Vila R, Kandul NP (2006) Rearrangement of the *Agrodiaetus dolus* species group (Lepidoptera, Lycaenidae) using a new cytological approach and molecular data. Insect Systematics & Evolution 37: 1–10. https://doi.org/10.1163/187631206788838563

Mariano CSF, Pompolo SG, Barros LAC, Mariano-Neto E, Campiolo S, Delabie JHC (2008) A biogeographical study of the threatened ant *Dinoponera lucida* Emery (Hymenoptera: Formicidae: Ponerinae) using a cytogenetic approach. Insect Conservation and Diversity 1: 161–168. https://doi.org/10.1111/j.1752-4598.2008.00022.x

Moura MN, Cardoso DC, Cristiano MP (2021) The tight genome size of ants: Diversity and evolution under ancestral state reconstruction and base composition. Zoological Journal of the Linnean Society 193: 124–144. https://doi.org/10.1093/zoolinnean/zlaa135

Moura MN, Cardoso DC, Lima Baldez BC, Cristiano MP (2020) Intraspecific variation in the karyotype length and genome size of fungus-farming ants (genus *Mycetophylax*), with remarks on procedures for the estimation of genome size in the Formicidae by flow cytometry. PLoS ONE 15(8): e0237157. https://doi.org/10.1371/journal.pone.0237157

Murakami T, Paris C, Chirino M, Sasa C, Sakamoto H, Higashi S, Sato K (2021) Unusual chromosome numbers and polyploidy in invasive fire ant populations. Genetica 149(4): 203–215. https://doi.org/10.1007/s10709-021-00128-4

Otto SP, Whitton J (2000) Polyploid incidence and evolution. Annual Review of Genetics 34: 401–437. https://doi.org/10.1146/annurev.genet.34.1.401

Peeters C, Molet M (2010) Colonial Reproduction and Life Histories. In: Lach L, Parr C, Abbott K (Eds) Ant Ecology, Oxford Academic, 159–176. https://doi.org/10.1093/acprof:oso/9780199544639.003.0009

Pitts JP, Mchugh JV, Ross KG (2005) Cladistic analysis of the fire ants of the *Solenopsis saevissima* species-group (Hymenoptera: Formicidae). Zoologica Scripta 34: 493–505. https://doi.org/10.1111/j.1463-6409.2005.00203.x

Porter SD, Tschinkel WR (1987) Foraging in *Solenopsis invicta* (Hymenoptera: Formicidae): effects of weather and season. Environmental Entomology 16: 802–808. https://doi.org/10.1093/ee/16.3.802

Rodman TC (1967) DNA replication in salivary gland nuclei of *Drosophila melanogaster* at successive larval and prepupal stages. Genetics 55: 375–386. https://doi.org/10.1093/genetics/55.3.375

Ross KG, Shoemaker DD (2005) Species delimitation in native South American fire ants. Molecular Ecology 14: 3419–3438. https://doi.org/10.1111/j.1365-294X.2005.02661.x

Santos MN (2016) Research on urban ants: approaches and gaps. Insectes Sociaux 63(3): 359–371. https://doi.org/10.1007/s00040-016-0483-1

Shoemaker DD, Ahrens ME, Ross KG (2006) Molecular phylogeny of fire ants of the *Solenopsis saevissima* species-group based on mtDNA sequences. Molecular Phylogenetics and Evolution 38: 200–215. https://doi.org/10.1016/j.ympev.2005.07.014

Silva APA (2016) Cytogenetic Description of 13 morphospecies of *Solenopsis* Westwood, 1840 (Hymenoptera: Formicidae). Ph.D. Dissertation, Universidade Federal de Viçosa, 56 pp. [In Portuguese]

Speicher MR, Carter NP (2005) The new cytogenetics: blurring the boundaries with molecular biology. Nature Reviews Genetics 6(10): 782–792. https://doi.org/10.1038/nrg1692

Taber SW, Cokendolpher JC (1988) Karyotypes of a dozen ant species from the southwestern U.S.A. (Hymenoptera: Formicidae). Caryologia 41: 93–102. https://doi.org/10.1080/00 087114.1988.10797851

White MJD (1973) Animal Cytology and Evolution. 3rd ed. Cambridge University Press, London–New York, 961 pp.

Wilson EO (1978) Division of Labor in Fire Ants Based on Physical Castes (Hymenoptera: Formicidae: *Solenopsis*). Journal of the Kansas Entomological Society 51(4): 615–636. https://www.jstor.org/stable/25083857

Wilson EO (1990) Success and dominance in ecosystems: the case of the social insects. Oldendorf/Luhe, Federal Republic of Germany: Ecology Institute, 104 pp.

Wojcik DP, Allen CR, Brenner RJ, Forys EA, Jouvenaz DP, Lutz RS (2001) Red Imported Fire Ants: Impact on Biodiversity. American Entomologist 47: 16–23. https://doi.org/10.1093/ae/47.1.16

Wurm Y, Wang J, Riba-Grognuz O, Corona M, Nygaard S, Hunt BG, Ingram KK, Falquet L, Nipitwattanaphon M, Gotzek D, Dijkstra MB, Oettler J, Comtesse F, Shih CJ, Wu WJ, Yang CC, Thomas J, Beaudoing E, Pradervand S, Flegel V, Cook ED, Fabbretti R, Stockinger H, Long L, Farmerie WG, Oakey J, Boomsma JJ, Pamilo P, Yi SV, Heinze J, Goodisman MAD, Farinelli L, Harshman K, Hulo N, Cerutti L, Xenarios I, Shoemaker DW, Keller L (2011) The genome of the fire ant *Solenopsis invicta*. Proceedings of the National Academy of Sciences of the United States of America 108(14): 5679–5684. https://doi.org/10.1073/pnas.1009690108

ORCID

Ananda Ribeiro Macedo de Andrade https://orcid.org/0009-0009-8321-4365
Danon Clemes Cardoso https://orcid.org/0000-0002-2811-2536
Maykon Passos Cristiano https://orcid.org/0000-0001-7850-9155

Supplementary material 1

Results from the karyomorphometrical analyses of *Solenopsis saevissima* and Chromosome counts frequency by individual and colony of *Solenopsis saevissima*

Authors: Ananda Ribeiro Macedo de Andrade, Danon Clemes Cardoso, Maykon Passos Cristiano

Data type: tables (PDF file)

Explanation note: table S1: Results from the karyomorphometrical analyses of *Solenopsis saevissima* n = 16 chromosomes. TL: total length; L: long arm length; S: short arm length; RL: relative length; *r*: arm ratio (= L/S); m: metacentric; sm: submetacentric; st: subtelocentric; table S2: Results from the karyomorphometrical analyses of *Solenopsis saevissima* 2n = 32 chromosomes. TL: total length; L: long arm length; S: short arm length; RL: relative length; *r*: arm ratio (= L/S); m: metacentric; sm: submetacentric; st: subtelocentric; table S3: Results from the karyomorphometrical analyses of *Solenopsis saevissima* 4n = 64 chromosomes. TL: total length; L: long arm length; S: short arm length; RL: relative length; *r*: arm ratio (= L/S); m: metacentric; sm: submetacentric; st: subtelocentric; table S4; Chromosome counts frequency by individual and colony of *Solenopsis saevissima*.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/compcytogen.17.100945.suppl1

COMPARATIVE
Cytogenetics
International Journal of Plant & Animal Cytogenetics,
Karyosystematics, and Molecular Systematics

A peer-reviewed open-access journal

Intraspecific divergence of diploid grass *Aegilops comosa* is associated with structural chromosome changes

Ekaterina D. Badaeva[1,2], Violetta V. Kotseruba[3], Andnrey V. Fisenko[1],
Nadezhda N. Chikida[4], Maria Kh. Belousova[4], Peter M. Zhurbenko[3],
Sergei A. Surzhikov[2], Alexandra Yu. Dragovich[1]

1 *N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkina str. 3, GSP-1, Moscow 119991, Russia* **2** *Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilova str. 32, GSP-1, Moscow 119334, Russia* **3** *Komarov Botanical Institute, Russian Academy of Sciences, Prof. Popova str. 2, Saint Petersburg 197376, Russia* **4** *N.I. Vavilov Institute of Plant Genetic Resources (VIR), Ministry of Science and Higher Education, Bolshaya Morskaya str. 42-44, Saint Petersburg 190000, Russia*

Corresponding author: Ekaterina D. Badaeva (katerinabadaeva@gmail.com)

Abstract

Aegilops comosa Smith in Sibthorp et Smith, 1806 is diploid grass with MM genome constitution occurring mainly in Greece. Two morphologically distinct subspecies – *Ae. c. comosa* Chennaveeraiah, 1960 and *Ae. c. heldreichii* (Holzmann ex Boissier) Eig, 1929 are discriminated within *Ae. comosa*, however, genetic and karyotypic bases of their divergence are not fully understood. We used Fluorescence in situ hybridization (FISH) with repetitive DNA probes and electrophoretic analysis of gliadins to characterize the genome and karyotype of *Ae. comosa* to assess the level of their genetic diversity and uncover mechanisms leading to radiation of subspecies. We show that two subspecies differ in size and morphology of chromosomes 3M and 6M, which can be due to reciprocal translocation. Subspecies also differ in the amount and distribution of microsatellite and satellite DNA sequences, the number and position of minor NORs, especially on 3M and 6M, and gliadin spectra mainly in the α-zone. Frequent occurrence of hybrids can be caused by open pollination, which, along with genetic heterogeneity of accessions and, probably, the lack of geographic or genetic barrier between the subspecies, may contribute to extremely broad intraspecific variation of GAA_n and gliadin patterns in *Ae. comosa*, which are usually not observed in endemic plant species.

Introduction

Aegilops comosa Smith ex Sibthorp et Smith, 1806 is annual diploid grass (2n=2x=14) with the MM genome constitution, which grows mainly in coastal and inland Greece, rarely – in coastal regions of Albania and Former Yugoslavia (Zhukovsky 1928; Eig 1929; Van Slageren 1994; Kilian et al. 2011). Several scattered populations have been found in Turkey (Zhukovsky 1928; Van Slageren 1994). Recently *Ae. comosa* was discovered also in Cyprus and Bulgaria (Van Slageren 1994).

Two morphologically distinct forms are discriminated within *Ae. comosa*; usually they are treated as subspecies of *Ae. comosa*: subsp. *comosa* Chennaveeraiah, 1960, thereafter *comosa*, and subsp. *heldreichii* (Holzmann et Boissier) Eig, 1929 thereafter *heldreichii* (Eig 1929; Hammer 1980; Kilian et al. 2011). Some taxonomists however recognize them as two distinct species: *Ae. comosa* and *Ae. heldreichii* (Boissier) Holzmann 1884 (Zhukovsky 1928; Chennaveeraiah 1960; Boguslavsky and Golik 2003), or as varieties of *Ae. comosa* [var. *comosa* Boissier, 1884 and var. *subventricosa* Jaubert et Spach ex Bornmüller, 1898 (Van Slageren 1994)]. Subspecies of *Ae. comosa* grow together, often in a mix with *Ae. caudata* Linneaus, 1753 on roadsides, grasslands and hillsides, sometimes in cultivated fields (Zhukovsky 1928). Except for Greece, *Ae. comosa* is uncommon to rare throughout its range.

Based on morphological similarity of *Ae. comosa* (both *comosa* and *heldreichii*) with *Ae. uniaristata* Visiani, 1852, P. Zhukovsky (1928) placed them into a common section *Comopyrum* Zhukovsky, 1928. These species however are genetically distinct and carry different types of nuclear and cytoplasmic genomes – M and N, respectively (Kimber et al. 1983; Kimber and Tsunewaki 1988). Radiation of *Ae. comosa* and *Ae. uniaristata* was accompanied by different structural chromosomal rearrangements (Molnár et al. 2016; Li et al. 2021; Said et al. 2021), which led to significant karyotype divergence of these species (Chennaveeraiah 1960; Teoh and Hutchinson 1983; Badaeva et al. 1996a, 2011; Friebe et al. 1996; Iqbal et al. 2000; Song et al. 2020; Li et al. 2021). Subspecies *comosa* and *heldreichii* are characterized by similar karyotype structures (Chennaveeraiah 1960), however meiotic analysis of their F_1 hybrid showed that they differ by one reciprocal translocation (Kihara 1940).

C-banding proved to be effective tool in phylogenetic analyses of the Triticeae. This method was employed to characterize karyotypes of all diploid and polyploid *Aegilops* Linneaus, 1753 species including *Ae. comosa* (Teoh and Hutchinson 1983; Friebe et al. 1992, 1993; Friebe et al. 1996; Badaeva et al. 1998, 2002, 2004), and differences between subspecies *comosa* and *heldreichii* in the amount and distribution of C-bands have been reported (Teoh et al. 1983; Friebe et al. 1996). On the contrary, other researchers failed to discriminate *comosa* from *heldreichii* using FISH with GAA_n microsatellite probe, which produces C-banding-like pattern (Song et al. 2020).

In earlier classifications C-banded *Ae. comosa* chromosomes were arranged in a decreasing length and designated with capital letters A – G (Teoh et al. 1983; Georgiou et al. 1992). First genetic nomenclature of the M-genome chromosomes was developed by B. Friebe et al. (1996) based on similarities of their C-banding patterns with homoeologous chromosomes of other *Aegilops* species. This system was later proved by chromosome sorting and single-gene FISH (Molnár et al. 2016; Said et al. 2021). Some controversy remained in classifying chromosomes 2M and 5M, which are indistinguishable by flow sorting due to similar morphology and same DNA content. Even more discrepancies in chromosome designations exist with the nomenclature developed by C. Liu et al. (2019) on the basis of analysis of addition and substitution wheat-*Ae. comosa* lines using FISH and PLUG markers. Thus, chromosome classification of this species still needs verification.

Aegilops comosa plays an important role in the evolution of polyploid *Aegilops*. Based on "analyzer" method H. Kihara (Kihara 1947, 1954) hypothesized that it gave rise to five tetraploid *Aegilops* species (bold **M** indicates genome modification): *Ae. crassa* Boissier, 1846 (DD**M**M), *Ae. columnaris* Zhukovsky, 1928 (UU**M**M), *Ae. neglecta* Requien ex Bertoloni, 1834 (UU**M**M), *Ae. biuncialis* Visiani, 1842 (UU**M**M), and *Ae. geniculata* Roth, 1787 (UU**M**M). Recent studies, however, did not confirm the presence of the M-genome in *Ae. crassa*, *Ae. columnaris* and *Ae. neglecta* (Resta et al. 1996; Badaeva et al. 2004; Edet et al. 2018; Abdolmalaki et al. 2019), but it was proved for *Ae. biuncialis* and *Ae. geniculata* (Kihara 1954; Kimber et al. 1988; Resta et al. 1996; Tsunewaki 1996; Friebe et al. 1999; Badaeva et al. 2004; Molnár and Molnár-Láng 2010; Molnár et al. 2011; Abdolmalaki et al. 2019; Said et al. 2022). All these papers reported significant genome modifications in karyotypes of these two tetraploid species, which seemed to proceed differently in *Ae. biuncialis* and *Ae. geniculata* (Badaeva et al. 2004; Said et al. 2022). *Aegilops comosa* as well as its tetraploid derivatives exhibited an extremely broad intraspecific variation of C-banding and/ or GAA_n labeling patterns (Teoh et al. 1983; Georgiou et al. 1992; Friebe et al. 1996; Badaeva et al. 2004; Song et al. 2020), which may impede delimitation of taxa boundaries and tracking evolutionary changes in karyotype of polyploids using these markers.

Among a broad range of botanical, cytogenetic, biochemical and molecular markers employed for evaluating intraspecific and interspecific diversity of wild and cultivated plant species, seed storage proteins (gliadins) appear to be relatively cheap, but informative markers for polymorphism analysis. Gliadins (Gli) belong to protein fraction prolamines, which is characterized by high glutamine and proline amino acid content and by specific molecular structure (size, domen composition, biochemical properties) (Shewry and Halford 2002). Electrophoretic spectra of gliadins allow discrimination, with high effectiveness, of lines, cultivars and varieties of tetraploid and hexaploid wheat; gliadin profiles are used to assess samples' heterogeneity and to evaluate phylogenetic relationships between species and accessions (Metakovsky et al. 2018). So far, polymorphism analyses of *Aegilops* based on gliadin loci were mainly focused on *Ae. tauschii* Cosson, 1850, the D-genome donor of common wheat (Yan et al. 2003; Dudnikov 2018; Badaeva et al. 2019a), and only few publications dealt with other *Aegilops* species (Cole et al. 1981; Medouri et al. 2015; Garg et al. 2016). On

the other hand, owing to similar structure of gliadin loci in wheat and *Aegilops* (Dong et al. 2016; Huo et al. 2018) and extremely high polymorphism, gliadins can serve as supplementary markers for the assessing genetic variability of *Ae. comosa* and clarifying phylogenetic relationships between the subspecies.

Aegilops comosa possesses a number of agronomically valuable traits such as pest and disease resistance (Riley et al. 1968; Gill et al. 1985; Boguslavsky and Golik 2003) and salt tolerance (Xu et al. 1996), which can potentially be used for wheat improvement. In late 60ᵗʰ of XX R. Riley, V. Chapman, and R.O.Y.Johnson introduced resistance to yellow rust from *Ae. comosa* into wheat cultivar "Compare" by genetically induced homoeologous recombination (Riley et al. 1968). Several wheat-*Ae. comosa* amphiploid and introgression lines have been developed in China and UK (Weng 1995; Liu et al. 2019; Zuo et al. 2020); some of these lines showed good resistance to yellow rust and powdery mildew (Liu et al. 2019). However, modifications of the M-genome chromosomes over the course of evolution, in particular, species-specific translocations, prevent the direct utilization of *Ae. comosa* gene pool in wheat breeding (Nasuda et al. 1998). Manipulations with genetic material of *Ae. comosa* require a deeper understanding of the genome of this species, its chromosomal structure and the range of polymorphism.

The aim of our study was a comparative analysis of *Ae. comosa* subsp. *comosa* and subsp. *heldreichii* on a broad sample of accessions of different geographical origins using FISH with fifteen DNA probes and electrophoretic analysis of seed storage proteins (gliadins) in order to characterize polymorphism and reveal mechanisms leading to divergence of subspecies.

Material and methods

Material

Thirty-six accessions of *Ae. comosa* including 20 accessions of *comosa* and 16 accessions of *heldreichii* collected from different regions of Greece and Turkey (Fig. 1) and maintained in genetic collections of N. I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St.-Petersburg, Russia, and Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany were used in our study (Suppl. material 9). Accessions of *comosa* and *heldreichii* show clear differences in spike morphology (Fig. 2).

Thus, spikes of subsp. *comosa* plants are slender, narrowly cylindrical, tapering toward apex, with 3–4 fertile and 0–2 rudimentary spikelets. Glumes of lateral spikelets have one tooth and one short awn, the apical spikelet has three well-developed awns, the central one of 4–11 cm long and lateral – 2.5–3.5 cm long. Spikes of subsp. *heldreichii* plants are shorter and stouter, not or hardly tapering toward the apex, with one rudimentary and 1–3 fertile spikelets. Lateral spikelets are urceolate, the apical one is obconical. Glumes are ovate, the lateral glumes with broadly triangular tooth on abaxial site and short awn on adaxial side. Apex of apical glume extends into three
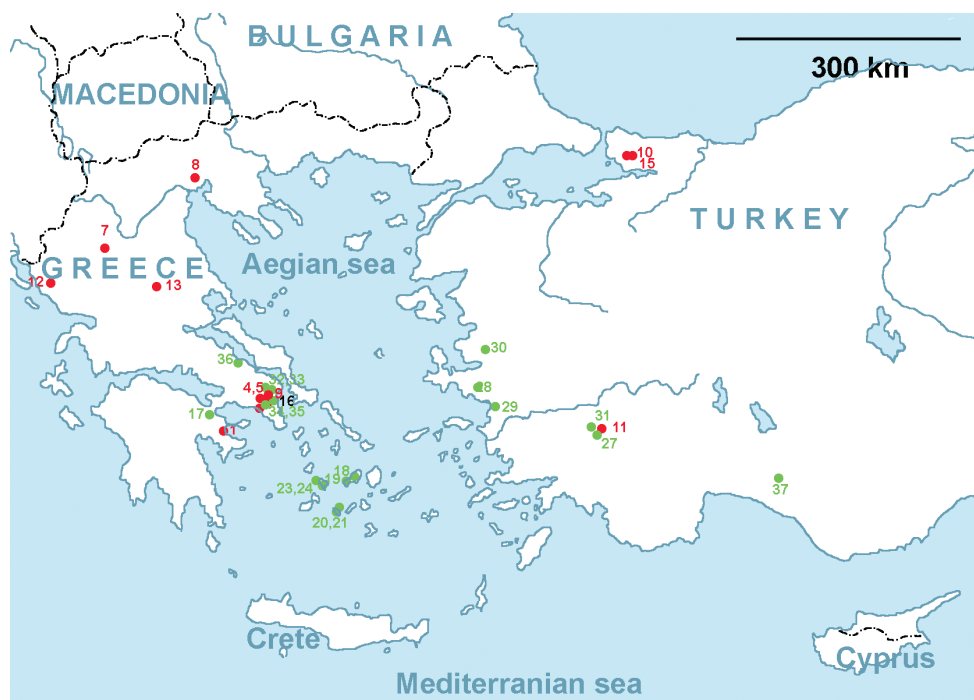
Figure 1. Geographical location of *Ae. comosa* subsp. *comosa* (green dots) and *Ae. comosa* subsp. *heldreichii* (red dots) accessions with known collection sites. The green (subsp. *comosa*) and red (subsp. *heldreichii*) numerals specify the accession numbers according to Suppl. material 9.

3–3.5 cm-long cetulose awns. Lateral awns are shorter and more slender, often reduced to teeth or even absent (Van Slageren 1994). Accessions of both subspecies significantly vary in spike length and color (Fig. 2), and accession K-3809 (subsp. *comosa*) is characterized by longest spike with black color.

DNA probes

Fifteen oligo-probes were used in FISH analysis. Microsatellite probes were labeled with either 6-FAM (GTT_{10}, GAA_{10}) or Cy3/TAMRA (GAA_{10}, ACT_{10}, AC_{20}) from the 5'-end. Oligo-18S was designed based on conservative region of the 18S rRNA gene. Melting temperature and potential secondary structures were calculated using OligoCalc (Kibbe 2007). Names and nucleotide sequences of other probes are listed in Table 1. The probes GTT_{10}, oligo-42, oligo-44, oligo-45, oligo-18SrDNA (o-18S), oligo-pSc119.2 were synthesized in Evrogen (Moscow, Russia); GAA_{10}-FAM, pTa-71-2 (pTa71), oligo-pTa-794 (5SrDNA), oligo-k566 were synthesized in Syntol (Moscow, Russia); ACT_{10}, GAA_{10}-Cy3, oligo-pAs1-1 (pAs1), oligo-pTa-713, oligo-pTa-535 were synthesized in the Laboratory of Biological Microchips at the Engelhardt Institute of Molecular Biology, Moscow, Russia.
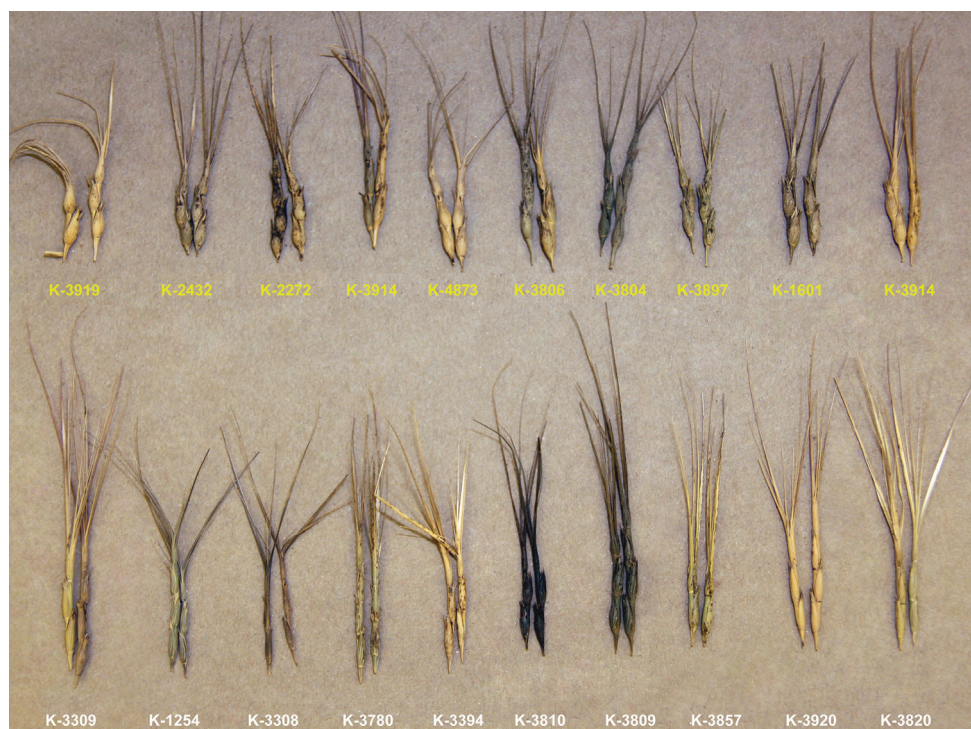
Figure 2. Comparison of spike morphology of *Ae. comosa heldreichii* (top raw) and *comosa* (bottom raw).

FISH analysis

The seeds are germinated on moist filter paper in Petri dishes at 24 °C. The seedlings with ~0.5 cm roots are transferred into 1.25 mM solution of hydroxyurea for 18 h, washed thoroughly with distilled water and grown in Petri dishes with distilled water for 5 h, as described in (Badaeva et al. 2017). The roots are cut and pretreated in ice water for 24 h and then fixed in the solution of ethanol : glacial acetic acid (3 : 1). Fixed roots are kept in fixative solution at -20 °C until use.

Metaphase cells are prepared by squashing, coverslips are removed after freezing in liquid nitrogen, and slides are kept in 96% ethanol at -20 °C. Fluorescence in situ hybridization is carried out according to previously published protocol (Badaeva et al. 2017). The slides are examined on a Zeiss Imager D1 epifluorescent microscope. Metaphase plates are captured with a 100^x objective using black and white digital camera Axiocam HRm using a software AxioVision, release 4.8. The images are processed using Adobe Photoshop, version 7.0.

Electrophoretic analysis of seed storage proteins (gliadins)

Electrophoresis (EP) in polyacrylamide gel (PAAG) according to the previously published protocol (Metakovsky and Novoselskaya 1991) was employed to obtain gliadin

Table 1. Oligo-probes used in FISH analysis.

Probe name	Sequence	Amount of probe (ng/ slide)	Reference
Oligo-pTa-71	FAM/5'- GGG CAA AAC CAC GTA CGT GGC ACA CGC CGC CTA-3'	21.1	Tang et al. 2014
Oligo-18S	FAM/5'- CTC GGA TAA CCG TAG TAA TTC TAG AGC TAA TAC GTG CAA CAA ACC CCG-3'	40.5	Current paper
Oligo-5S rDNA	Cy3/5'-TCA GAA CTC CGA AGT TAA GCG TGC TTG GGC GAG AGT AGT AC-3'	27.1	Yu et al. 2019
Oligo-GAA$_n$	TAMRA (or FAM)/5'-GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA-3'	21,4	Cuadrado et al. 2008a
Oligo-GTT$_n$	FAM/5'-GTT GTT GTT GTT GTT GTT GTT GTT GTT GTT-3'	19.5	Cuadrado et al. 2008a
Oligo-ACT$_n$	Cy3/5'-ACT ACT ACT ACT ACT ACT ACT ACT ACT ACT-3'	20.1	Cuadrado et al. 2008a
Oligo-AC	TAMRA/5'-AC AC AC AC AC AC AC AC AC AC AC AC AC AC AC AC AC AC AC AC AC AC-3'	18.4	Cuadrado et al. 2008a
Oligo-pSc119.2	FAM/5'- CCG TTT TGT GGA CTA TTA CTC ACC GCT TTG GGG TCC CAT AGC TAT -3'	28.3	Tang et al. 2014
Oligo-pAs1	Cy3/5'-CCT TTC TGA CTT CAT TTG TTA TTT TTC ATG CAT TTA CTA ATT ATT TTG AGC TAT AAG AC-3'	36.7	Tang et al. 2014
Oligo-pTa-713	Cy3/5'- GTC GCG GTA GCG ACG ACG GAC GCC GAG ACG AGC ACG TGA CAC CAT TCC CAC CCT GTC TA-3'	37.9	Tang et al. 2016
Oligo-pTa-535	Cy3/5'- AAA AAC TTG ACG CAC GTC ACG TAC AAA TTG GAC AAA CTC TTT CGG AGT ATC AGG GTT TC-3'	37.4	Tang et al. 2014
Oligo-k566	FAM/5'- ATC CTA CCG AGT GGA GAG CGA CCC TCC CAC TCG GGG GCT TAG CTG CAG TCC AGT ACT CG-3'	37.1	Tang et al. 2016
Oligo-45	TAMRA/5'-CGG CCG CTC CGC GCG TCG CCA TCG GTT GGT CAC CTC ATC ACC ACT-3'	28.2	Tang et al. 2018a
Oligo-42	FAM/5'-CTC GCT CGC CCA GCT GCT GCT ACT CCG GCT CTC GCT CGA TCG-3'	26.1	Tang et al. 2018a
Oligo-44	TAMRA/5'-TAG CTC TAC AAG CTA GTT CAA ATA ATT TTA CAC TAG AGT TGA AC-3'	27.88	Tang et al. 2018a

spectra of 26 accessions of *Ae. comosa*. Since no information on the composition and inheritance of the blocks of gliadin components in this species was available from literature, electrophoretic spectra of all accessions must be compared with an etalon sample with the known genetic control of components. In wheat and related species the gliadin spectrum of bread wheat cultivar Bezostaya-1 serves as an etalon (Novoselskaya-Dragovich et al. 2018).

The gliadin spectra of the Triticeae are traditionally divided into four zones, α, β, γ and ω-zones, depending on electrophoretic mobility of individual polypeptides (Woychik et al. 1961). Peptides from the ω-zone are coded by genes located on group 1 chromosomes, while those from the α-zone – by genes of group 6 chromosomes. Components from β and γ zones are controlled by chromosomes of both genetic groups (Dong et al. 2016; Huo et al. 2018). Based on this information we presume that electrophoretic components from the α-zone of the spectra of *Ae. comosa* accessions are encoded by 6M, whereas from the ω-zone, by 1M chromosome.

Results

### Intraspecific diversity of *Ae. comosa* in karyotype structure

Two subspecies (*comosa* and *heldreichii*) of *Ae. comosa* have similar karyotype structures, which include metacentric, submetacentric and subacrocentric chromosomes (Fig. 3 a–c). Despite overall karyotypic similarity, we observe variation in the number and morphology of satellite (SAT) chromosomes (Fig. 3d). Most accessions have two pairs of satellite chromosomes differing in morphology (Fig. 3a, b), but few genotypes carry only one SAT pair (Fig. 3c).

The satellite on one pair is always small, and this chromosome is classified as 1M. The satellite on the second pair – 6M, is much larger and appears on physically longer arm. Comparison of the SAT chromosomes allows to divide *Ae. comosa* accessions into four groups (Fig. 3d). Group I is characterized by long submetacentric chromosome 1M with a small SAT attached to the short arm. The SAT on 6M is very large and occupies nearly half of the arm length. Most accessions from this group belong to subsp. *heldreichii*, and this combination of satellite chromosomes is designated "*heldreichii*-like". Group II differs from Group I in shorter satellite length (approximately 1/3 to 1/4 of the arm) on the chromosome 6M. Accessions from this group belong predominantly to subsp. *comosa* and we designate this combination of satellite chromosomes as "*comosa*-like". Groups III and IV include representatives of both subspecies. Group III shows altered morphology of both SAT chromosome pairs, whereas Group IV contains just one pair of SAT chromosomes – 1M (K-3308, K-3309, K-4873; Fig. 3d). Group IV includes one of the three analyzed genotypes of AE 1254, one of the two genotypes of each K-3824, K-3920 (all *comosa*), and all K-3806 genotypes (*heldreichii*). All these genotypes are characterized by heteromorphic pair of 6M chromosomes: one homolog carries large, while the second – much smaller satellite on the long arms (Fig. 3, "heterozygote"), this group presumably represents hybrids between the subspecies.

Diversity of *Ae. comosa* in localization of rDNA clusters

Clusters of rDNA were mapped on chromosomes of 36 *Ae. comosa* accessions by FISH with probes oligo-pTa71-2 (thereafter pTa71), oligo-pTa-794 (5S rDNA), and oligo-18S (thereafter o-18S). Comparion of labeling patterns obtained using pTa71 and o-18S probes reveals intrinstic feature. The pTa71 visualizes all minor and major rDNA loci (Figs 4a, 5a), while o-18S fails to reveal major NORs and *Ae. comosa*-specific minor NORs located terminally on 2MS, 3MS and 5MS arms (Figs 4b, 5i). By contrast, o-18S is more efficient in detecting minor NORs located in interstitial and pericentromeric regions of many M-genome chromosomes (Figs 4b, 5i; Suppl. material 1: fig. S1c16, h04–h06, h10).

In karyotype of *Ae. comosa* major NORs are located on chromosomes 1M and 6M (Fig. 6). Signals on 1M and 6M detected with pTa71 probe are usually equal in size, however in AE 1254 (*comosa*) and K-3806 (*heldreichii*) the signals on 6M are significantly smaller compared to 1M (Fig. 4a, Suppl. material 1: fig. S1, c06, h03).
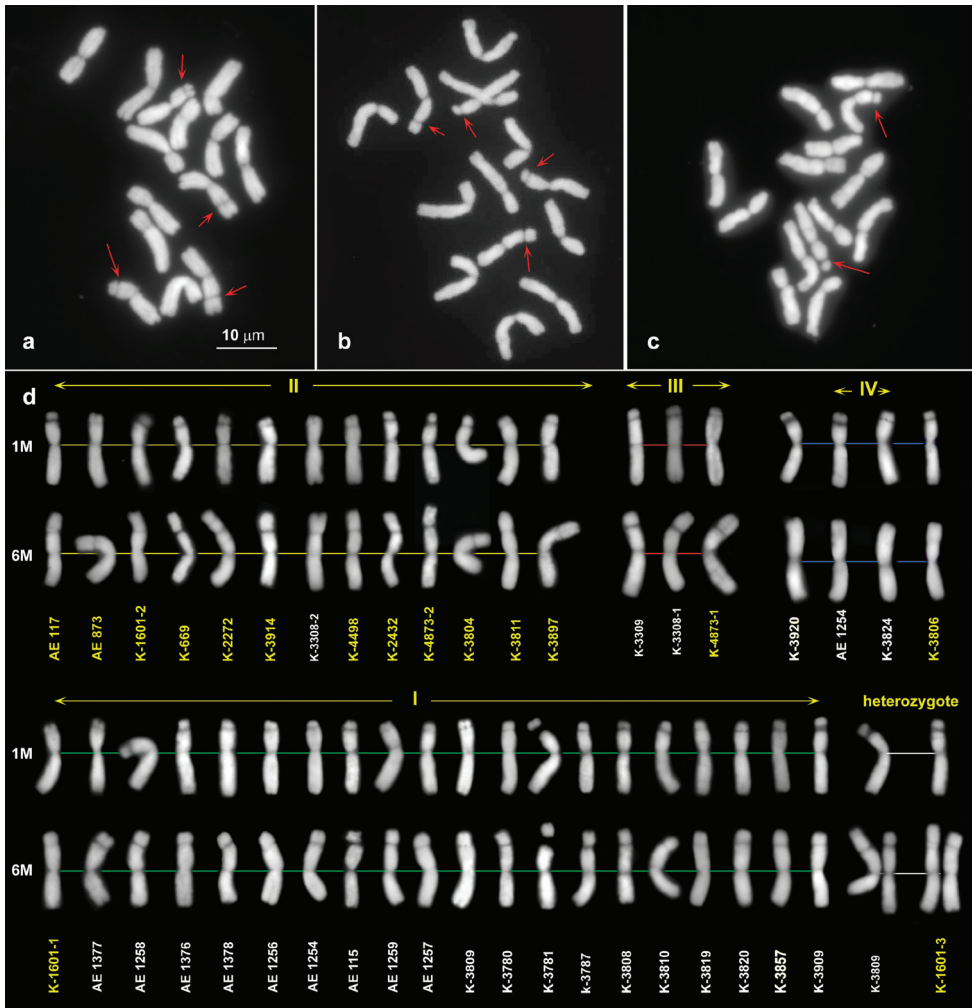
Figure 3. Metaphase cells of *Ae. comosa* subsp. *heldreichii*, AE 783 (**a**) and subsp. *comosa*, AE 1258 (**b**) and AE 1254 (**c**). Satellite chromosomes are shown with red arrows. (**d**) structural diversity of the SAT chromosomes in *Ae. comosa*: accessions of *comosa* are designated with white numbers, *heldreichii* – with yellow numbers. Scale bar: 10 μm.

As mentioned above, these accessions carry only one pair of SAT chromosomes; thus, inactivation of NORs on 6M is associated with/ or caused by elimination of the 45S rRNA gene sequences from the respective loci.

The pTa71 probe produces distinct signals in subtelomeric regions of short arms of 2M, 3M, and 5M chromosomes (minor NORs) in all accessions of both *comosa* and *heldreichii*, however, these loci are not visualized by o-18S probe. The latter ribosomal probe however reveals weak signals in pericentromeric region of 1ML, distal region of 2ML and 4MS, and two minor loci in the proximal half of 7M short arm; all these
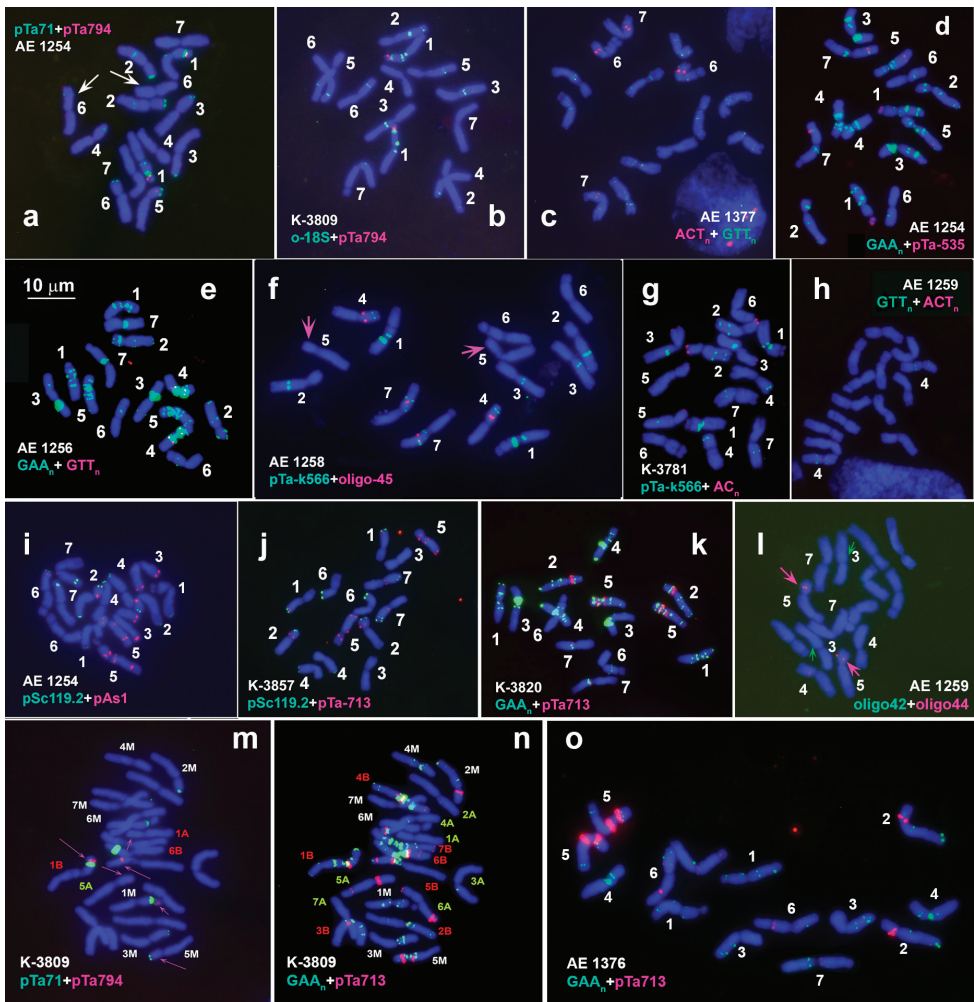
Figure 4. Localization of repeated DNA sequences on metaphase chromosomes of *Ae. comosa* subsp. *comosa*. Accession numbers and probe names are shwon on each plate; probe color corresponds to signal color. Chromosomes are designated with numerals according to genetic groups. Scale bar: 10 μm.

sites are common for both subspecies. By contrast, several sites discriminate *comosa* from *heldreichii* (Fig. 6, Suppl. material 1: fig. S1). These are: (1) the locus of variable intensity in a proximal part of chromosome 3M is observed in nearly all (with one exception) *comosa* genotypes, but not in *heldreichii*; (2) most *comosa* carry two minor pericentromeric o-18S sites in the opposite arms of chromosome 1M, while *heldreichi* possesses signal only in the long arm; (3) minor pericentromeric rDNA site on chromosome 4M is usually located in the long arm of *comosa*, but in the short arm of 4M[h] of *heldreichii*; (4) most *comosa* accessions possess two weak intercalary pTa71 sites in the long arm of 2M, and only one distal site is present in *heldreichii*.
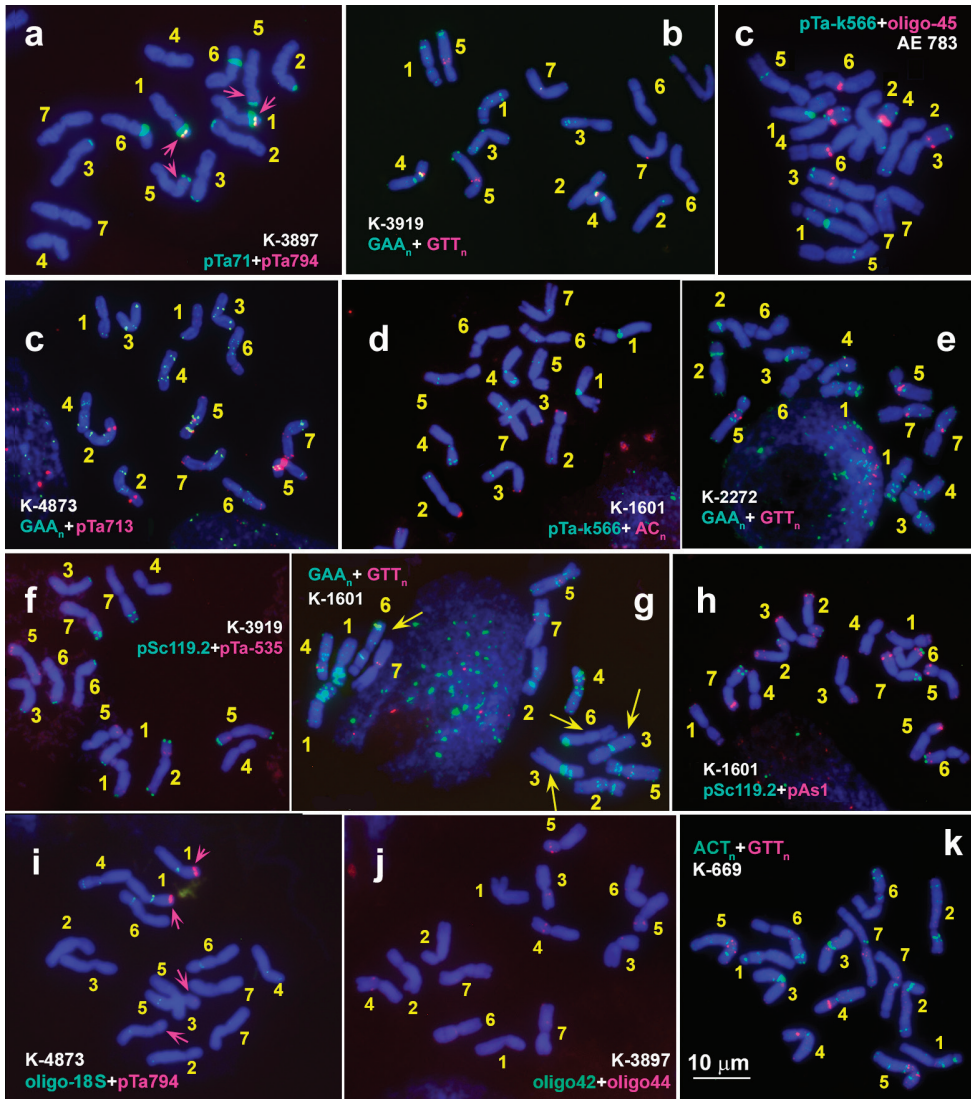
Figure 5. Localization of repeated DNA sequences on metaphase chromosomes of *Ae. comosa* subsp. *heldreichii*. Accession numbers and probe names are shown on each plate; probe color corresponds to signal color. Chromosomes are designated with numerals according to genetic groups. Scale bar: 10 μm.

Two pairs of 5S rDNA loci with unequal size are revealed in all *Ae. comosa* accessions. The signal on 1M is much larger than that on 5M and arrears distally to NOR. The signal on 5M is very faint, especially in *heldreichii*, and occurs in the middle of short arm (Fig. 6).

Based on hybridization pattern of 5S and 45S rDNA probes we identify a reciprocal translocation between 1M and 6M chromosomes in K-3308 (Suppl. material 1:
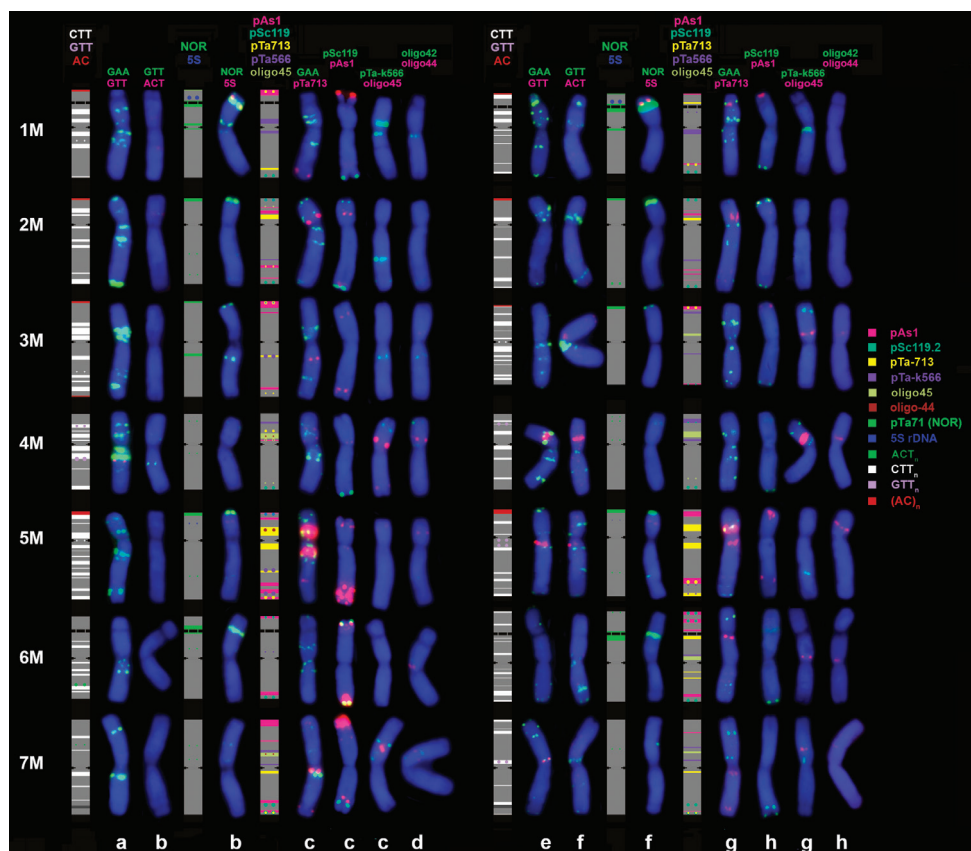
Figure 6. Idiograms and chromosomal images showing the distribution of repetitive DNA families on chromosomes of *comosa* (left side) and *heldreichii* (right side). Probe names are shown on the top of the figure; probe color corresponds to signal color. Accession codes are given in the bottom: **a** AE 1257 **b** AE 1259 **c** AE 1258 **d** AE 1377 **e** AE 117 **f** K-669 **g** AE 783 **h** K-3897.

fig. S1c13), which caused the alterations in morphology of the SAT chromosomes observed in all Group 3 accessions.

Diversity of *Ae. comosa* in the distribution of microsatellite probes

Four microsatellite probes: GAA_{10}, GTT_{10}, ACT_{10}, and AC_{20} were mapped on chromosomes of *Ae. comosa*. Hybridization with GAA_{10} (Figs 4d, e, k, 5b, c, e, g) results in patterns similar to C-banding (C-banding-like patterns). Thus, *comosa* chromosomes carry numerous GAA_n signals of variable size in subterminal, interstitial and pericentromeric chromosome regions. Compared to *comosa*, *heldreichii* possess lesser number of $(GAA)_n$ sites, most of which with low or moderate intensity. Hybridization signals in *heldreichii* appear predominantly in subtelomeric and pericentromeric chromosome regions, rarer interstitially (Fig. 6).

Giant GAA$_n$ signals exceeding the respective sites on 1M of *comosa* are detected on the short arm of 1M^h of few *heldreichii* accessions. Both *comosa* and *heldreichii* exhibit an extremely broad polymorphism of GAA$_n$-labeling patterns (Suppl. material 2: fig. S2, Suppl. material 3: fig. S3); differences between accessions, between individual plants within the accession (Suppl. material 3: fig. S3h–i, t, u, Suppl. material 4: fig. S4d–f, k, l) and even between homologous chromosomes (heteromorphism of homologues) have frequently been observed (Fig. 5g; Suppl. material 3: fig. S3e). Using the GAA$_n$ probe we confirm the 1M-6M translocation in K-3308 (*comosa*) and identify a similar translocation in K-4873 (*heldreichii*) accession (Suppl. material 2: fig. S2d, e, Suppl. material 3: fig. S3j).

Most obvious differences between the subspecies show chromosomes 3M and 3M^h (Fig. 6), which differ even in size and morphology. 3M of *comosa* is larger and more asymmetric (submetacentric) than 3M^h of *heldreichii* (metacentric), due to the loss of a distal part of the long arm. The short arm of 3M (*comosa*) carries two prominent, often fused GAA$_n$ clusters in a proximal part and several smaller sites in the distal and, rarer, in the proximal third of the long arm. A slightly deviant hybridization pattern of 3M in K-3857 (*comosa*) can be caused by large pericentric inversion (Fig. 4q). Chromosome 3M^h of *heldreichii* shows highly polymorphic labeling patterns (Suppl. material 3: fig. S3), however three intercalary GAA$_n$-sites in the long arm are constantly present. Most proximal GAA$_n$ site varies in signal intensity from huge (K-669, K-1601, K-2272, K-2432) to medium or even small (K-3914, K-3919).

Among all *comosa* accessions, K-3857, AE 1376 and AE 1377 from Greece exhibit most deviant GAA$_n$ patterns (Suppl. material 2: fig. S2n, p, q). According to karyotype analysis, K-1601 is highly heterogeneous and consists of genotypes, belonging to *comosa* and *heldreichii* types, and some seedling prove to be hybrids, including hybrids between the subspecies (Suppl. material 3: fig. S3e). Hybrids between the subspecies and even with an unknown tetraploid wheat species have been identified within the accession K-3809 (Fig. 3m, n).

Distribution of GTT$_{10}$ probe discriminates subspecies *comosa* from *heldreichii*. Two small, but sharp intercalary signals appear in short and long arms of 4M in all *comosa* accessions (Fig. 4e). These GTT$_n$ sites are located at the borders of large pericentromeric GAA$_n$ complex (Suppl. material 2: fig. S2g). All studied *heldreichii* accessions possess distinct GTT$_n$ signals on three chromosome pairs, 4M (middle of short arm), 5M and 7M (pericentromeric regions) (Fig. 5b, e, g; Suppl. material 3: fig. S3a). In addition, very weak, fuzzy hybridization signals of GTT$_n$ probe may appear in regions overlapping with GAA$_n$ in both *comosa* and *heldreichii* accessions.

Poor hybridization with ACT$_{10}$ probe is observed in eight *Ae. comosa* accessions examined in our study; chromosomal regions corresponding to C-bands/ GAA$_n$ sites show a little brighter intensity (Figs 4c, h, 5k). Sharp signals appear on chromosomes 6ML and 7ML of only one *comosa* accession – AE 1377 (Fig. 4c).

Hybridization of (AC)$_{20}$ repeat on chromosomes of *comosa* and *heldreichii* results in similar patterns: small subtelomeric signals appear on chromosome arms 1MS, 2MS, 3MS/ 3ML, and 5MS (Suppl. material 4: fig. S4g, h, p). These are the sites in which minor subtelomeric NORs are also detected (Figs 4g, 5d; Suppl. material 1: fig. S1 c01, c03, h02).

Polymorphism in the distribution of satellite DNAs: pTa-713 family

Most prominent pTa-713 clusters appear in the short arm of 2M and in very proximal parts of both arms of 5M. Two smaller sites are observed in subtelomeric and pericentromeric regions of the long arm of 7M of all studied accessions of *Ae. comosa* (Figs 4j, k, 5c, 6). Signals in a terminal part of 1MS, a distal part of 1ML, and in subtelomeric region of 5ML are detected in several accessions belonging to both subspecies, however, some polymorphic pTa-713 sites are specific for either *comosa* or *heldreichii* (Suppl. material 5: fig. S5, Suppl. material 6: fig. S6).

Thus, sixteen of 20 accessions of *comosa* carry small signal in a proximal third of 3ML, while the signal in the middle of 5ML is detected in more than a half *comosa* accessions (Suppl. material 5: fig. S5). AE 1376 and AE 1377 (*comosa*) show the unique pTa-713 patterns on 3M chromosome consisting of clear signal in the terminus of short arm and fuzzy signal in a distal part of the long arm (Suppl. material 5: fig. S5g, h, shown with pink arrows), which are not observed in *comosa* and *heldreichii* accessions, except K-3811 (*heldreichii*) (Suppl. material 6: fig. S6e). Most *comosa* also possess two faint signals in the short arm of 4M. Two *comosa*'s lack these marker sites, but carry a small signal in the distal part of 4ML arm (Suppl. material 5: fig. S5o, t). Many small sites in unusual positions appear in Turkish accession K-3787 (*comosa*).

All *heldreichii* accessions show characteristic labeling pattern of chromosome 6M^h, which carries clear pTa713 signal in the short arm, either adjacent to secondary constriction (8 of 13 accessions) or in the middle of satellite (2 accessions), and two signals in the proximal half of the long arm (Suppl. material 6: fig. S6). We consider such pattern as *heldreichii*-specific, although it is not observed in two "*comosa*-like" genotypes of K-1601. These genotypes carry "short" satellite on 6M^h, similar to chromosome 6M of *comosa*. A unique prominent pTa-713 site is detected in a distal part of 5M^{h}L of K-3914 (*heldreichii*) (Suppl. material 6: fig. S6a).

Polymorphism in the distribution of satellite DNAs: pSc119.2 family

Signals of pSc119.2 probe appear in subtelomeric regions of 1ML, 2MS+2ML, and 6ML arms and in distal (*comosa*) or terminal + distal parts of 7ML (mainly *heldreichii*) of all studied accessions of *Ae. comosa* (Figs 4i, j, 5f, h). The pSc119.2 signals in the distal part of 4MS and a terminus of 4ML occur more frequently in *comosa* (Fig. 6; Suppl. material 7: fig. S7). No pSc119.2 signals appear on 5M chromosome of 26 studied accessions of *Ae. comosa* belonging to both subspecies, and only three accessions (all *heldreichii*) carry small pSc119.2 site in the terminus of 3M^{h}L (Suppl. material 7: fig. S7o, r).

Polymorphism in the distribution of satellite DNAs: pAs1 and pTa-535 families

Comparison of labeling patterns obtained using pAs1 and pTa-535 probes (e.g., Fig. 4d, I, red color) shows that all pTa-535 sites overlap with pAs1 sites, however many

pAs1 sites are not detected by pTa-535 (Suppl. material 3: fig. S3t). Based on this observation, we choose pAs1 for polymorphism assessment in *Ae. comosa*. In general, *comosa* exhibits more intense hybridization with pAs1, however broad variation of labeling patterns observed in both subspecies preclude their discrimination using this probe (Suppl. material 7: fig. S7). Prominent pAs1 clusters occur in terminus of the SAT of 1M chromosome of AE 115, AE 1256, AE 1258, AE 1378, K-3309, K-3787, K-3819, K-3820 (*comosa*, see Suppl. material 7: fig. S7a, c, h, i) and K-2432 (*heldreichii*, see Suppl. material 7: fig. S7l) accessions. Accessions AE 1256, AE 1259, AE 1378, K-3787, K-3819, K-3820, K-3920, and K-3857 contain large signals in subtelomeric region of 3MS (Suppl. material 7: fig. S7c, f, i), K-3820 and AE 1258 – in subterminal part of 6ML, and accessions AE 1258, AE 1259, K-3309, K-3787 (*comosa*) and K-1601 (*heldreichii*) – in subtelomeric region of the 7M short arm (Suppl. material 7: fig. S7e, f, h). Prominent interstitial pAs1 clusters emerge in a distal part of 5M long arm of AE1254 and AE1258 (*comosa*), and K-3914 (*heldreichii*). All chromosomes except 4M possess medium to small hybridization sites located in interstitial and more frequently in distal and subtelomeric chromosome regions. Hybridization with pSc119.2 and pAs1 probes confirms the translocation 1M-6M in *comosa* accession K-3309 (Suppl. material 7: fig. S7h).

Polymorphism in the distribution of satellite DNAs: pTa-k566 and oligo-45 families

The pTa-k566 sequence hybridizes to all chromosomes of *Ae. comosa* and labeling patterns varied between the accessions. Polymorphism observed on 1M, 2M, 3M, 4M, 6M, and 7M chromosomes is found to be subspecies-specific (Fig. 6, Suppl. material 4: fig. S4). Chromosomes 7M of *comosa* and *heldreichii* with similar probe distribution differ only in signal intensities, whereas other chromosomes carry pTa-k566 sites specific for only one of the subspecies. Chromosome 1M of *comosa* contains pTa-k566 sites at the both sides of the centromere, but only one site occurs on 1M^hL of *heldreichii*. Median and distal pTa-k566 signals are present in the long arm of 2M of *comosa*, while only distal of them appears in *heldreichii*. Chromosome 3M of *comosa* carries prominent pTa-k566 site in the proximal region of the long arm, while the respective site on 3M^h of *heldreichii* is much weaker and the chromosome possess the second site distally in the short arm. Chromosome 4M of *comosa* lacks small proximal pTa-k566 site in the short arm, which occurs on 4M^h of all *heldreichii* accessions. *comosa* differs from *heldreichii* in the presence of clear pTa-k566 site in the proximal regions of 6M short arm, although it is also detected in two of 11 studied *comosa* accessions, namely K-3787 and AE 1377 (Suppl. material 4: fig. S4). Noteworthy that all pTa-k566 sites overlap with o-18S loci.

Chromosomes 1M and 2M do not possess signals of oligo-45 probe. Labeling patters of oligo-45 on chromosomes 4M and 7M of both *Ae. comosa* subspecies are similar, but signal intensity on 7MS is stronger in *comosa*. Some oligo-45 sites appear to be subspecies-specific. Thus, small interstitial signal in a distal part of 5MS is present only

in *comosa*, whereas chromosomes 3M^h and 6M^h of *heldreichii* carry distinct oligo-45 sites proximally in the short arm (Suppl. material 4: fig. S4). Accessions K-4873 (*heldreichii*) is heterozygous in labeling pattern of oligo-45 on 5M^h chromosome, whereas AE 1377 (*comosa*) – of 6M chromosome (Suppl. material 4: fig. S4b, j). Besides this, accession AE 1377, being classified as *comosa*, carries variants of 1M, 2M, 3M, and 4M chromosomes typical for *heldreichii*, while 5M and 7M – typical for *comosa* (Suppl. material 4: fig. S4b).

## Polymorphism in the distribution of satellite DNAs: oligo-42 and oligo-44 families

No signals of oligo-42 probe are detected on chromosomes of five *heldreichii* accessions, and very weak, inconsistent signals are found on chromosomes of very few *comosa* accessions. Most frequently, signals occur in the long arm of 3M (Fig. 4l), but in AE 1377 oligo-42 signals are also observed on other chromosomes (Fig. 6). Positions of these sites coincide with the signals of o-18S probe. Owing to weakness and inconsistency of oligo-42 signals, this probe is considered non-informative for FISH-analysis of *Ae. comosa*.

All *Ae. comosa* accessions contain clear oligo-44 site approximately in the middle of 5M short arm (Figs 4l, 5j; Suppl. material 4: fig. S4). Additional, weaker signals emerge in the proximal part of 4M short arm and in a proximal third of 7M short arm. Besides them, many, but not all *heldreichii* accessions carry minor oligo-44 signals proximally in 3M short arm and in pericentromeric region of 6M short arm. Both *heldreichii*-specific sites are recorded in AE 1377 (*comosa*), which however exhibits also several *comosa*-specific karyotype features.

Polymorphisms of electrophoretic patterns of gliadins

Genetic variability of *Ae. comosa* was also assessed using electrophoretic analysis of gliadins in 13 *comosa* and 13 *heldreichii* accessions. Comparison of gliadin profiles of all 26 accessions reveals an extremely broad intraspecific variation of *Ae. comosa*: each accession shows the unique profile. Spectra of *comosa* accessions are usually more "enriched in components" compared to *heldreichii* accessions (Fig. 7). This trend is manifested in a larger number of protein bands and their higher intensities, which is especially clear in the α-zone of electrophoretic spectra controlled by chromosomes of homoeologous group 6. From the other side, this trend is not mandatory because gliadin profiles of some *comosa* accessions (e.g., AE 1257, AE 1376, AE 1377) are very poor, whereas several *heldreichii* accessions exhibit rich spectra (K-3919; K-4498, some genotypes of K-1601).

Several accessions analyzed in this study are genetically homogeneous and consist of genotypes with identical spectra (e.g., K-3914, K-3309, K-3920, AE 1376). Other accessions are found to be heterogeneous and show two or even more gliadin profiles (K-3857, K-3809 (*comosa*), K-1601 and K-2432 (*heldreichii*)). The broadest variation of gliadin patterns is detected in K-1601. We identified six variants of electrophoretic
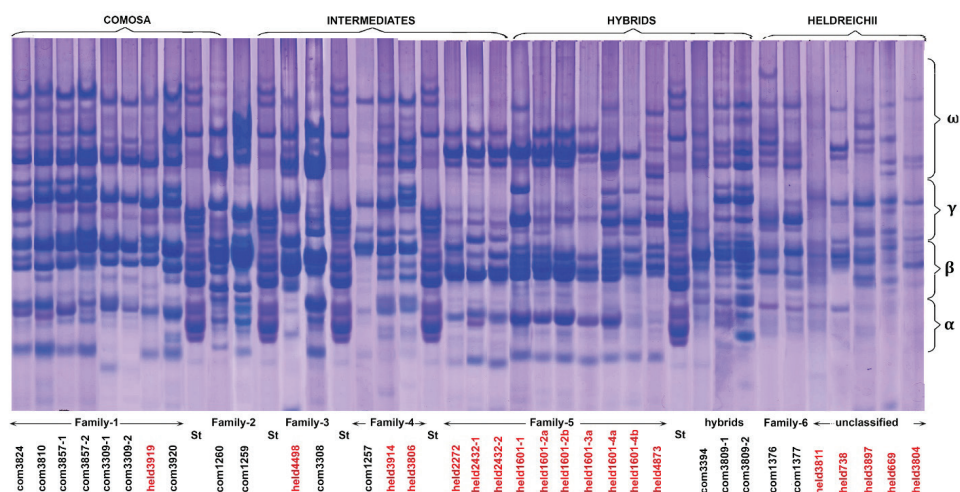
Figure 7. Electrophoretic spectra of *Ae. comosa* subsp. *comosa* (black) and subsp. *heldreichii* (red) accessions and their subdivision on families. St – the etalon spectra of wheat Bezostaya-1.

spectra among eight individual grains taken from four spikes; they differ in position of polypeptide bands in γ- and ω-zones (Fig. 8). The spectra 1a and 1b are more typical for *comosa*, whereas 2a and 3b (different spikes) – for *heldreichii*. The spectrum 3a contains bands present on 2a and 3b and probably represents a hybrid between these two genotypes. Similarly, the spectrum 1a may correspond to hybrid between 1b and 3b (Fig. 8). Another heterogeneous accession – K-3809 (*comosa*) is characterized by spectra highly enriched with gliadin components, which differ significantly from other *Ae. comosa* in the number, position and size of protein bands. This genotype can represent a hybrid of *Ae. comosa* with unknown 4× wheat species.

Comparison of gliadin patterns of different accessions of *Ae. comosa* shows that several combinations of polypeptide bands on electrophoretic spectra always appear together and are inherited as blocks of components (Fig. 7). Such coincidence means that (1) these bands are controlled by a common locus, which, in turn, indicates (2) that accessions under comparison are genetically related. Based on the presence of such common blocks we discriminate six gliadin Families among 26 *Ae. comosa* accessions (Fig. 7): (1): K-3824; K-3309, K-3810, K-3824; K-3857, K-3920 – *comosa*, K-3919 – *heldreichii*; (2): AE 1259, AE 1260 – *comosa*; (3): K-3308 – *comosa*, K-4498 – *heldreichii*; (4): AE 1257 – *comosa*, K-3806, K-3914 – *heldreichii*; (5): K-1601, K-2272, K-2432, K-4873 – *heldreichii*; (6): AE-1377, AE-1376 – *comosa*.

Families include accessions mainly from one subspecies. Thus, Families 1 and 2, which exhibited the richest spectra, consist of predominantly *comosa* accessions. Families 4 and 6 and partially Family 5 having relatively poor spectra are mainly composed by *heldrechii* accessions. Although both accessions from Family 6 have been assigned to subspecies *comosa*, their gliadin spectra share more common features with the spectra of *heldrechii* accessions AE 783, K-669, K-3804, K-3811, and K-3897 than with those
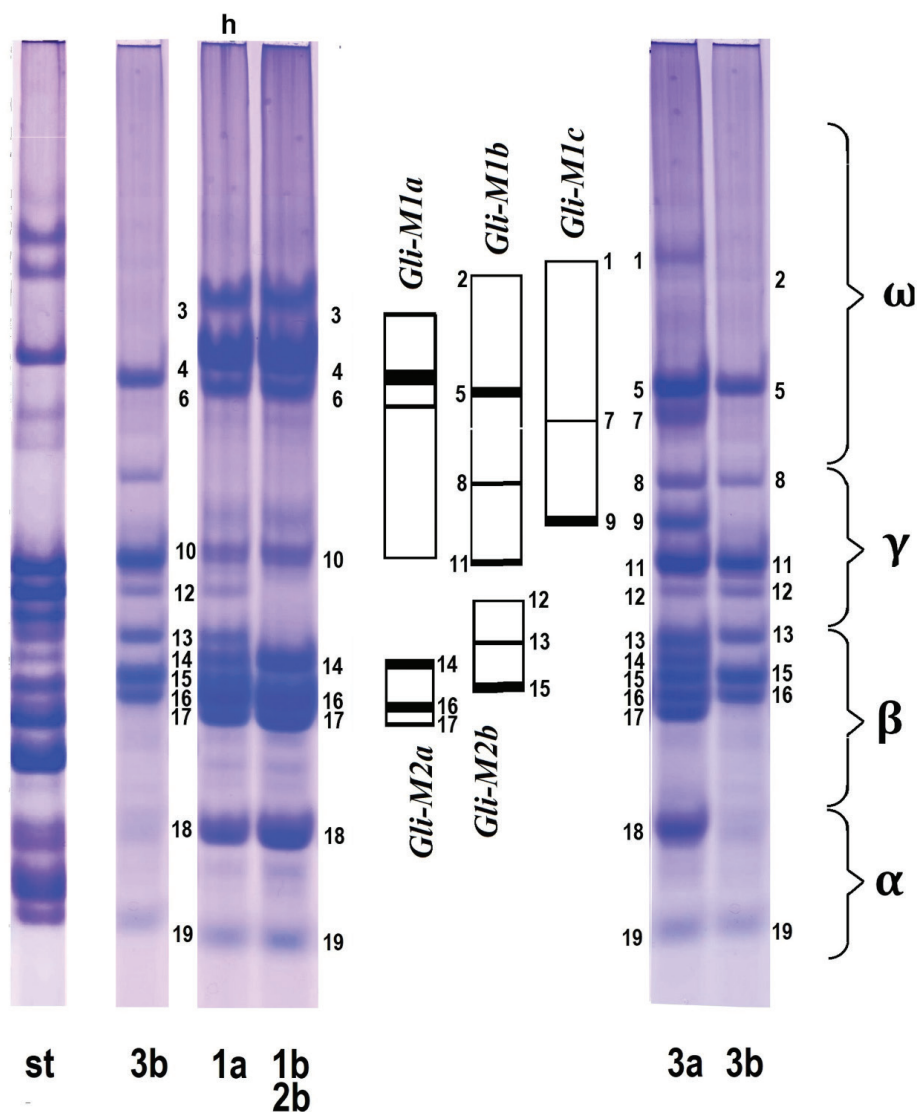
Figure 8. Diversity of the gliadin spectra detected in different grains taken from three individual spikes (1a and 1b; 2a and 2b, 3a and 3b) of *heldreichii* accession K-1601. St – an etalon wheat cultivar Bezostaya-1. Square blocks in the middle of the figure specify the blocks of gliadin components identified in eight genotypes of K-1601 (*Gli-M1a*, *Gli-M1b*, *Gli-1c*; *Gli-M2a*, *Gli-M2b*).

of *comosa*. Each of the five abovementioned *heldrechii* accessions show unique gliadin profile, which cannot be assigned to either one of the families due to dissimilarities in position and intensity of polypeptide bands on electrophoretic spectra. High ratio of *heldrechii* accessions with the unique gliadin spectra is an indicative of higher variability of this subspecies compared to subspecies *comosa*.

Discussion

Our current results and the data available from literature (Song et al. 2020) demonstrate that the levels of intraspecific diversity in *Ae. comosa* significantly vary depending on markers used for their assessment. Thus, FISH with GAA_n probe and gliadin electrophoresis uncover an extremely broad polymorphism of this species. High effectiveness of these markers for the analysis of intra- and interspecific diversity, evaluation of population structure, for characterization of individual genotypes has been proved in many publications (Metakovsky et al. 1989; Badaeva et al. 1990, 2015b, 2022; Masci et al. 1992; Pomortsev et al. 2011; Keskin et al. 2015; Jiang et al. 2017; Majka et al. 2017; Song et al. 2020). We find that, from one side, distribution of GAA_n sites is species- and chromosomes specific allowing identification of all individual chromosomes. From the other side, each accession carries a unique combination of polymorphic GAA_n sites and unique gliadin spectrum comprising their individuality.

Comparison of the GAA_n patterns of *Ae. comosa* chromosomes obtained in a current study and those reported previously (Molnár et al. 2016; Song et al. 2020) with the C-banding patterns (Teoh et al. 1983; Friebe et al. 1996; Badaeva et al. 1999) show that positions of GAA_n overlap with location of C-bands. Two other microsatellite sequences – GTT_n and ACT_n, which are known to be the components of constitutive heterochromatin in chromosomes of wheat and other grass species (Pedersen and Langridge 1997; Cuadrado et al. 2000; Cuadrado et al. 2008a, b; Luo et al. 2017, 2018; Zhang et al. 2022) are detected in minor quantities. The $(AC)_n$ sites co-localize with positions of minor NORs visualized with pTa71 probe on 1MS, 2MS, 3MS/3ML, and 5MS arms. These facts indicate that heterochromatin regions detected using Giemsa C-banding on *Ae. comosa* chromosomes are composed by predominantly GAA_n repeat, which is typical for the Triticeae, except diploid *Thinopyrum* Á.Löve, 1980 einkorn wheat and *Ae. tauschii*. Chromosomes of these species contain only little amounts of microsatellite sequences or do not possess them at all (Badaeva et al. 2015a, 2019a, b; Linc et al. 2017; Li et al. 2018; Zhao et al. 2018; Ebrahimzadegan et al. 2021).

Although labeling patterns of GAA_n probe prove to be highly informative for *Ae. comosa* chromosome identification and authentication of gene bank accessions, they are too polymorphic and complicated for broad-scale phylogenetic analyses. A similar complexity and ambiguity is found for gliadin profiles. The appropriate markers should be relatively simple and easy to score and should generate specific and reproducible patterns. Eight out of 15 FISH probes used in our study fit these criteria: the 5S and 45S rDNAs, pAs1, pSc119.2, pTa-713, pTa-k566, oligo-44, and oligo-45 probes (oligo-42 and ACT_n are found to be low informative for the analysis of *Ae. comosa* chromosomes due to weak and inconsistent labeling patterns). We used these eight probes for verification of the M-genome chromosome classification and for the assessment of intraspecific diversity of *Ae. comosa*.

Our study reveals an interesting feature of two oligo-probes designed for the detection of 45S rDNA loci. The oligo-pTa71-2 probe developed by Tang et al. (2014) is homologous to wheat rDNA 25S-18S intergenic region *Eco*RI-*Bam*HI fragment

(X-7841.1). This oligo-probe proves to be effective for the analysis of wheat and *Ae-gilops* species, however it fails to detect NORs in most other plant taxa, including, for example, barley, oat, or *Erantus* (Ranunculaceae) (Mitrenina et al. 2021). Owing to this, we tried to design a new oligo-probe for rDNA loci (based on genome sequence of *Aegilops tauschii*), which will be applicable for many plant species. Nucleotide sequence of a newly designed oligo-18S is homologous to highly conservative region of 18S rDNA gene of *Aegilops*, *Triticum*, *Hordeum*, *Musa*, and *Iris* (Suppl. material 8: fig. S8d) and it was able to detect major NORs on barley chromosomes, although signals were very weak. The o-18S hybridizes to chromosomes of many *Aegilops* species, however it was more effective for the detection of minor NORs (Suppl. material 8: fig. S8c). In contrast to pTa71 probe obtained from plasmid DNA or oligo-pTa71, o-18S does not detect major NORs and fails to reveal marker terminal 45S rDNA loci on 2M, 3M and 5M chromosomes (Suppl. material 8: fig. S8a, b). Currently we have no explanation of this phenomenon.

Correct chromosome classification is essential for phylogenetic analyses. Nomen-clatures suggested for *Ae. comosa* chromosome classification have been built on different principles. In early studies, the authors followed the rules of cytological nomencla-ture: chromosomes are arranged according to decreasing length and arm ratio (Chen-naveeraiah 1960; Teoh et al. 1983; Georgiou et al. 1992). Currently most chromosome classifications of the Triticeae species are based on homoeology with common wheat chromosomes – genetic nomenclature. Although genetic nomenclature of *Ae. comosa* chromosomes have been suggested in several publications (Badaeva et al. 1996a; Friebe et al. 1996; Molnár et al. 2011, 2016; Liu et al. 2019; Song et al. 2020; Said et al. 2021), classification of some chromosomes is still controversial. Two chromosomes – 2M and 5M, prove to be most difficult for discrimination owing to same DNA content, similar morphology, pAs1 labeling patterns, heterochromatin content and distribution.

Chromosome 2M was first identified and assigned to genetic group 2 by S. Nas-uda et al. (1998) using RFLP, GISH and C-banding techniques in *Ae. comosa*, wheat cultivar Compair, and wheat-*Ae. comosa* 2A/ 2M and 2D/ 2M translocation lines. Classification of 2M and 5M was further validated by M. Said et al. (2021) based on FISH mapping of tandem repeats and wheat single-gene probes. In a current study we confirmed classification of 2M and 5M by using DNA probes pTa794 (5S rDNA) and oligo-44. This is because the 5S rDNA loci in *Triticum* Linneaus, 1753 and *Aegilops* oc-cur only on group 1 and 5 chromosomes (Dvořák et al. 1989), and in diploid *Aegilops* species the respective signals appear in a distal part of group 1 chromosomes and in the middle of short arm of group 5 chromosomes (Badaeva et al. 1996b). The 5S rDNA signal on 5M localizes in the middle of short arm, thus this chromosome belongs to genetic group 5. No signals of 5S rDNA probe are detected on another chromosome, which is designated 2M (Fig. 9).

Another validation of genetic group for chromosome 5M come from hybridiza-tion pattern of oligo-44 probe. This probe was developed from chromosome-specific tandem repeats of wheat (Tang et al. 2018b) and mapped on all group 5 chromosomes,
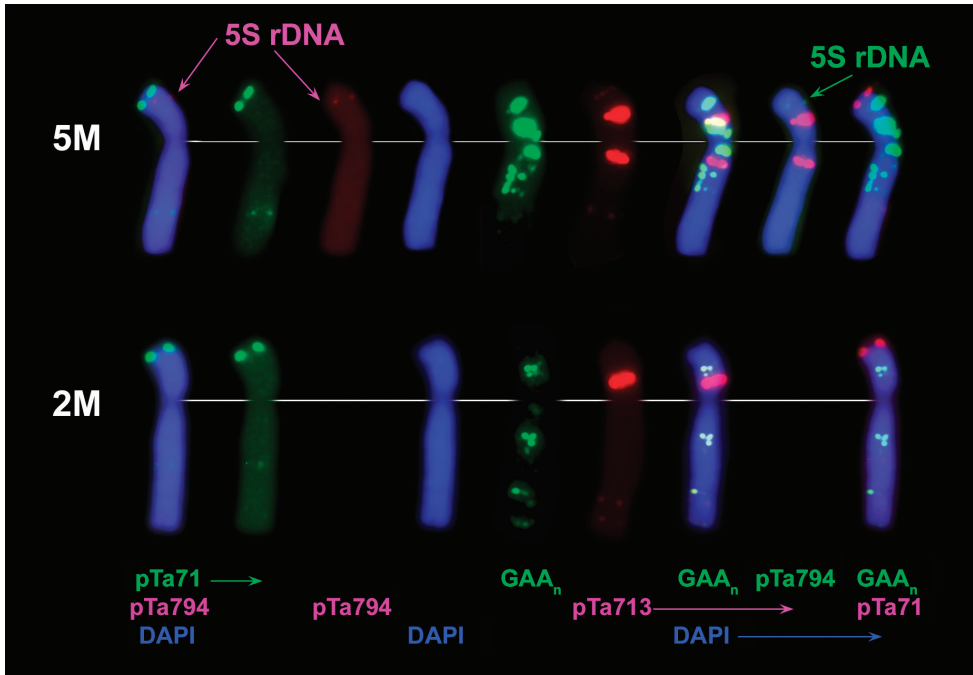
Figure 9. Characterization of 2M and 5M chromosomes of *Ae. comosa* using sequential FISH with (round 1) pTa794 (5S rDNA), pTa71 (45S rDNA), followed by (round 2) GAA$_{10}$ and pTa-713 probes. Probe names are shown on the bottom, probe color corresponds to signal color. Arrows point to 5S rDNA loci on chromosome 5M.

3A and 7A of Chinese Spring. A little later, T. Lang at al. (2019) using bioinformatics tools identified the homologous minisatellite repeat, Ta-3A1 from the wheat genome assembly and localized it on group 5 chromosomes of wheat (2x, 4x and 6x), rye, *Aegilops*, *Dasypyrum* (Cosson et Durieu) T.Durand, 1888 and *Thinopyrum* species. The CL241 probe, another homolog of oligo-44 isolated from *Ae. crassa* genome, was also mapped to all group 5 chromosomes of *Ae. tauschii* (2n=2x=14, DD), tetraploid and hexaploid *Ae. crassa* (2n=4x=28, $D^1D^1X^{cr}X^{cr}$ and 2n=6x=42, $D^1D^1X^{cr}X^{cr}D^2D^2$) (Kroupin et al. 2022).

Interestingly, that in *Ae. comosa* we reveal minor oligo-44 signals on chromosomes 3M and 7M, the homoeologs of 3A and 7A, which also possess the TA-3A1 sites in karyotypes of diploid, tetraploid and hexaploid wheat species, but not in *Aegilops* (Lang et al. 2019).

Another interesting outcome from our study is a possibility of discrimination of the two *Ae. comosa* subspecies using chromosomal markers. From one side, all thirty-six accessions of *Ae. comosa* included in the analyses show similar karyotype structures, distribution of rDNA loci, labeling patterns of repetitive DNA sequences indicating that they all belong to one biological species. Despite these similarities, we found clear

and reproducible differences between the subspecies in morphology, C-banding and FISH patterns of two chromosome pairs – 3M and 6M. We hypothesize that these two chromosomes were involved in the subspecies-specific translocation identified earlier in *Ae. comosa* by mean of meiotic analysis (Kihara 1940). This translocation probably led to the shortage of the long arm of 3M and the increase in the satellite length on 6M of *heldreichii* compared to *comosa*.

Interestingly, all *heldreichii* accessions carrying large SAT on chromosome 6M showed relatively poor gliadin profiles (Fig.8). Considering these facts, we suggest that the 3M-6M translocation can change the functioning of gliadin loci on 6M^h chromosome resulting in depletion of gliadin spectra in the α-zone. Intraspecific differentiation of *Ae. comosa* (Groups I and II) in the morphology of SAT chromosomes observed in our study and reported previously (Karataglis 1975) is caused by this subspecies-specific translocation, whereas the Group III probably evolved due to reciprocal 1M-6M translocation identified here in *comosa* K-3308, K-3309, K-3787 (all from Turkey) and K-4873 (*heldreichii*) accessions using FISH with rDNA, pTa-713, and GAA$_n$ probes.

Peculiarities of C-banding (Teoh et al. 1983; Friebe et al. 1996; Badaeva et al. 1999)/ GAA$_n$ patterns also allow discrimination of *Ae. comosa* subspecies indicating that their divergence was accompanied by amplification/ elimination and re-distribution of microsatellite repeats. These results however contradict the observations of Z. Song et al. (2020), who have not revealed any differences between the subspecies in GAA$_n$ labeling patterns. These authors, however, did not provide complete karyotypes of accessions used in their analyses, which does not permit the direct comparison of our data.

Differences between the subspecies are most clearly detected using a combination of pTa-k566 and oligo-45 probes; the pTa-713 also shows subspecies-specific patterns. These three probes prove to be best choice for the precise discrimination of *comosa* from *heldreichii* accessions using FISH markers. Most *Ae. comosa* chromosomes show rather conservative patterns, while diagnostic sites appear mainly on 3M or 6M. Different repeats are often accumulated in a single cluster. Such complex cluster, composed of 45S rDNA, pTa-713, pTa-k566, oligo-42, and, rarely, GAA$_n$ appears in a proximal third of the long arm of 3M (*comosa*) (Fig. 6). These repeats are mainly absent or found in minor quantities in the long arm of 3M^h (*heldreichii*), which contains small cluster consisting of oligo-42 and oligo-45 in the short arm. Clear signals of pTa-713, oligo-44 and oligo-45 probes are detected on chromosome 6M^h of *heldreichii*, but they are lacking in 6M of *Ae. comosa*. Although pAs1 probe cannot reliably discriminate *comosa* from *heldreichii*, pSc119.2-labeling pattern of 7M shows differences between the subspecies, although with few exceptions. Thus, two *comosa* accessions – AE 1377 and K-3809, carry two pSc119.2 sites on 7ML, which is typical for *heldreichii*. These accessions however are deviant and share karyotypic features of both subspecies. Cytogenetic analysis reveals that some grains of K-3809 represent hybrids of *Ae. comosa* with unknown tetraploid wheat, which can explain a deviant gliadin profile of this

accession, and one genotype possessing heteromorphic chromosome 5M is probably a derivative of *comosa* × *heldreichii* cross.

The significant role of hybridization in evolution and diversification of *Ae. comosa* is supported by other studies (Georgiou et al. 1992; Song et al. 2020). Thus, these papers described many heterozygotes in *Ae. comosa*, and in a current study we found genotypes segregating in labeling patterns of one to all seven chromosome pairs (see Suppl. material 1: fig. S1, h10, Suppl. material 3: fig. S3e, Suppl. material 4: fig. S4j), which point to their hybrid origin. Accession K-1601 (*heldreichii*) shows the highest heterogeneity: each of the seven genotypes examined by FISH and 5/8 genotypes analyzed by gliadin electrophoresis show unique patterns. Karyotypic features of some K-1601 genotypes correspond to *heldreichii* subspecies, while other share similarities with *comosa*. A similar trend is uncovered by gliadin analysis. Both FISH and gliadin electrophoresis identify many heterozygotes in K-1601, which may represent recent hybrids, including hybrids between the subspecies. It should be mentioned, however, that no variation in spike morphology (all *heldreichii*-like, see Fig. 2) has been identified between individual plants of this accession.

Three accessions, AE 1376, AE 1377 and K-3857 assigned to subspecies *comosa* based on botanical characters, combine chromosomal features of both *Ae. comosa* subspecies assuming that they might have hybrid origin. Gliadin analysis supports closer relations of AE 1376 and AE 1377 with *heldreichii* than with *comosa* indicating that taxonomical position of these accessions should be verified. Probably these forms emerged via hybridization of *comosa* and *heldreichii* followed by karyotype stabilization toward *heldreichii* (AE 1376, AE 1377) or *comosa* (K-3857) parent. In contrast to K-1601 or K-3809, these three accessions are cytogenetically stable and genetically uniform. Most likely, they emerged via *comosa* × *heldreichii* hybridization long time ago, and hybrid forms become stabilized over generations. Based on these facts we suggest that hybridization, including hybridization between subspecies, plays an important role in broadening genome diversity of this grass. It can be facilitated by following factors:

- *comosa* and *heldreichii* often grow together in mix stands (Zhukovsky 1928; Van Slageren 1994);
- heading and flowering time of *comosa* and *heldreichii* overlap (Boguslavsky and Golik 2003);
- *Ae. comosa* is considered as autogamous species, however, open pollination could be more common event than it is usually believed (Georgiou et al. 1992; Boguslavsky and Golik 2003);
- hybrids between the subspecies are partially fertile (Kihara 1940).

Summarizing results of a current study, we recommend the following set of markers for the precise identification of individual chromosomes and for discrimination of *Ae. comosa* subspecies using FISH markers (Table 2).

Table 2. Probe combinations for the M-genome chromosome identification and discrimination of *Ae. comosa* subspecies (according to Fig. 6).

Chr #	Markers common for subspecies	Markers discriminating subspecies
1M	Major NOR (pTa-71) in short arm; 5S rDNA locus in the satellite; terminal $(AC)_n$ site in satellite; pSc119.2 site in long and pAs1/ pTa-535 site in short arm; proximal o-18S/ pTa-k566 site in long arm.	Proximal o-18S/ pTa-k566 site in the short arm (*comosa*)
2M	Minor NOR in short arm overlapping with $(AC)_n$ site; pSc119.2 signals in both arms; pAs1 signals in short and long arms; distal o-18S/ pTa-k566 site in long arm; large pTa-713 cluster in short arm.	Intercalary o-18S/ pTa-k566 site in the middle of long arm (*comosa*).
3M	Minor NOR in short arm overlapping with $(AC)_n$ site; subterminal pAs1/pTa-535 cluster(s) of various intensity in short (and long) arm(s)	Metacentric (*heldreichii*) vs submetacentric (*comosa*); GAA_n patterns; cluster pTa71+pTa-713+pTa-k566+o-18S in long arm (*comosa*)/ cluster pTa-k566 in short arm (*heldreichii*); oligo-45 site in short arm (*heldreichii*).
4M	Minor distal NOR in short arm; prominent oligo-45 cluster in short and small – in long arm; Oligo-44 site overlapping with oligo-45; One-two faint pTa-713 sites in short arm.	pTa71/ o-18S in short (*heldreichii*) vs long (*comosa*) arms, proximal pTa-k566 site in short arm (*heldreichii*); two faint (*comosa*) vs one clear (*heldreichii*) GTT_n sites in short arm
5M	Minor NOR in short arm overlapping with $(AC)_n$; 5S rDNA site in the middle of short arm; Two prominent pTa-713 clusters; pTa-k566 site in long arm; oligo-44 site in short arm; pSc119.2 signals are mainly absent; pAs1 sites distally in the long arm and terminally in the short arm.	pericentromeric GTT_n cluster (*heldreichii*); oligo-45 site in short arm (*comosa*).
6M	Satellite in physically longer arm carrying major NOR; terminal pSc119.2 and distal pAs1 sites in the arm, opposite to NOR.	Medium (*comosa*) vs large (*heldreichii*) satellite; small pTa-713 sites in both arms (*heldreichii*); oligo-44, oligo-45, and pTa-k566 sites in the SAT arm (*heldreichii*).
7M	proximal pTa71/ o-18S/ pTa-k566 sites in short arm; oligo-45 site in short arm; pTa-713 sites in subtelomeric and proximal regions of long arm.	two (*heldreichii*) vs one (*comosa*) pSc119.2 sites in long arm; GTT_n cluster in proximal part of short arm (*heldreichii*); o-18S, pTa-k566, oligo-45 signal intensities

Conclusions

FISH and gliadin electrophoresis reveal broad intraspecific polymorphism of GAA_n patterns and gliadin profiles of *Ae. comosa* allowing not only genetic authentication of gene bank accessions, but also discrimination between the subspecies. Application of these markers however will be too complicated for the broad-scale phylogenetic analyses.

By using group-specific FISH markers, we justify classification of 2M and 5M chromosomes of *Ae. comosa* and suggest a set of DNA probes for the precise identification of each of the seven M-genome chromosomes.

Two subspecies of *Ae. comosa* – *comosa* and *heldreichii*, are karyotypically distinct and diverge from each other as a result of subspecies-specific translocation 3M-6M, which probably affects functioning of gliadin locus. Divergence of subspecies was accompanied with amplification/ elimination and re-distribution of the repeated DNA sequences.

Three FISH probes, pTa-k566, pTa-713, and oligo-45 generate clear and reproducible patterns specific for *comosa* or *heldreichii* accessions; they can serve as reliable markers for discrimination of *Ae. comosa* subspecies.

An extremely broad genetic variability of GAA_n-FISH patterns and gliadin profiles revealed in *Ae. comosa* – an endemic autogamous plant species (Van Slageren 1994), can be due to frequent occurrence of hybridization, including hybridization of *comosa* with *heldreichii* or with other neighboring wheat or *Aegilops* species.

Author contributions

EB (Badaeva E.D.) planned and performed the experiments, analyzed data, wrote the first draft of the manuscript; KV: make chromosomal preparations and participates in FISH experiments; FA: carried gliadin electrophoresis, analyzed data; CN and BM: provide and characterized materials for this study; ZP: designed oligo probe; SS: synthesized oligo-probes; DA: analyzed gliadin spectra and wrote the manuscript. All authors read and approved the submitted version of the manuscript and agree to be personally accountable for their own contributions and for ensuring that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and documented in the literature.

Acknowledgements

References

Abdolmalaki Z, Mirzaghaderi G, Mason AS, Badaeva ED (2019) Molecular cytogenetic analysis reveals evolutionary relationships between polyploid *Aegilops* species. Plant Systematics and Evolution 305(6): 459–475. http://doi.org/10.1007/s00606-019-01585-3

Badaeva E, Dedkova O, Zoshchuk S, Amosova A, Reader S, Bernard M, Zelenin A (2011) Comparative analysis of the N-genome in diploid and polyploid *Aegilops* species. Chromosome Research 19(4): 541–548. http://doi.org/10.1007/s10577-011-9211-x

Badaeva ED, Amosova AV, Goncharov NP, Macas J, Ruban AS, Grechishnikova IV, Zoshchuk SA, Houben A (2015a) A set of cytogenetic markers allows the precise identification of

all A-genome chromosomes in diploid and polyploid wheat. Cytogenetic and Genome Research 146(1): 71–79. http://doi.org/10.1159/000433458

Badaeva ED, Amosova AV, Muravenko OV, Samatadze TE, Chikida NN, Zelenin AV, Friebe B, Gill BS (2002) Genome differentiation in *Aegilops*. 3. Evolution of the D-genome cluster. Plant Systematics and Evolution 231(1–4): 163–190. http://doi.org/10.1007/s006060200018

Badaeva ED, Amosova AV, Samatadze TE, Zoshchuk SA, Shostak NG, Chikida NN, Zelenin AV, Raupp WJ, Friebe B, Gill BS (2004) Genome differentiation in *Aegilops*. 4. Evolution of the U-genome cluster. Plant Systematics and Evolution 246(1–2): 45–76. http://doi.org/10.1007/s00606-003-0072-4

Badaeva ED, Chikida NN, Filatenko AA, Zelenin AV (1999) Comparativa analysis of the M-genome chromosomes of *Aegilops comosa* and *Ae. heldreichii* using C-banding and in situ hybridization. Russian Journal of Genetics 35(6): 791–799.

Badaeva ED, Fisenko AV, Surzhikov SA, Yankovskaya AA, Chikida NN, Zoshchuk SA, Belousova MK, Dragovich AY (2019a) Genetic heterogeneity of a diploid grass *Aegilops tauschii* revealed by chromosome banding methods and electrophoretic analysis of the seed storage proteins (gliadins). Russian Journal of Genetics 55(11): 1315–1329. http://doi.org/10.1134/S1022795419110024

Badaeva ED, Friebe B, Gill BS (1996a) Genome differentiation in *Aegilops*. 1. Distribution of highly repetitive DNA sequences on chromosomes of diploid species. Genome 39(2): 293–306. http://doi.org/10.1139/g96-040

Badaeva ED, Friebe B, Gill BS (1996b) Genome differentiation in *Aegilops*. 2. Physical mapping of 5S and 18S-26S ribosomal RNA gene families in diploid species. Genome 39(6): 1150–1158. http://doi.org/10.1139/g96-145.

Badaeva ED, Friebe B, Zoshchuk SA, Zelenin AV, Gill BS (1998) Molecular cytogenetic analysis of tetraploid and hexaploid *Aegilops crassa*. Chromosome Research 6(8): 629–637. http://doi.org/10.1023/A:1009257527391

Badaeva ED, Keilwagen J, Knüpffer H, Waßermann L, Dedkova OS, Mitrofanova OP, Kovaleva ON, Liapunova OA, Pukhalskiy VA, Özkan H, Graner A, Willcox G, Kilian B (2015b) Chromosomal passports provide new insights into diffusion of emmer wheat. PLoS ONE 10: e0128556. http://doi.org/10.1371/journal.pone.0128556

Badaeva ED, Konovalov FA, Knüpffer H, Fricano A, Ruban AS, Kehel Z, Zoshchuk SA, Surzhikov SA, Neumann K, Graner A, Hammer K, Filatenko A, Bogaard A, Jones G, Özkan H, Kilian B (2022) Genetic diversity, distribution and domestication history of the neglected GGAᵗAᵗ genepool of wheat. Theoretical and Applied Genetics 135(3): 755–776. http://doi.org/10.1007/s00122-021-03912-0

Badaeva ED, Ruban AS, Aliyeva-Schnorr L, Municio C, Hesse S, Houben A (2017) In Situ Hybridization to Plant Chromosomes. In: Liehr T (Ed.) Fluorescence In Situ Hybridization (FISH) Application Guide. Springer, Berlin, 477–494. http://doi.org/10.1007/978-3-662-52959-1

Badaeva ED, Sozinova LF, Badaev NS, Muravenko OV, Zelenin AV (1990) "Chromosomal passport" of *Triticum aestivum* L. em Thell. cv. Chinese Spring and standardization of chromosomal analysis of cereals. Cereal Research Communications 18(4): 273–281.

Badaeva ED, Surzhikov SA, Agafonov AV (2019b) Molecular-cytogenetic analysis of diploid wheatgrass *Thinopyrum bessarabicum* (Savul. and Rayss) A. Löve. Comparative Cytogenetics 13(4): 389–402. http://doi.org/10.3897/CompCytogen.v13i4.36879

Boguslavsky RL, Golik OV (2003) The genus *Aegilops* L. as the genetic resourse for breeding Yuriev Institute of Plant Industry, NAAN, Kharkov, Ulrain, 129 pp.

Chennaveeraiah MS (1960) Karyomorphologic and cytotaxomonic studies in *Aegilops*. Acta Horti Gotoburgensis 23: 85–186.

Cole EW, Fullington JG, Kasarda DD (1981) Grain protein variability among species of *Triticum* and *Aegilops*: quantitative SDS-PAGE studies. Theoretical and Applied Genetics 60(1): 17–30. http://doi.org/10.1007/BF00275173

Cuadrado A, Cardoso M, Jouve N (2008a) Increasing the physical markers of wheat chromosomes using SSRs as FISH probes. Genome 51(10): 809–815. http://doi.org/10.1139/G08-065

Cuadrado A, Cardoso M, Jouve N (2008b) Physical organisation of simple sequence repeats (SSRs) in *Triticeae*: structural, functional and evolutionary implications. Cytogenetic and Genome Research 120(3–4): 210–219. https://doi.org/10.1159/000121069

Cuadrado A, Schwarzacher T, Jouve N (2000) Identification of different chromatin classes in wheat using in situ hybridization with simple sequence repeat oligonucleotides. Theoretical and Applied Genetics 101(5): 711–717. http://doi.org/10.1007/s001220051535

Dong L, Huo N, Wang Y, Deal K, Wang D, Hu T, Dvořák J, Anderson OD, Luo M-C, Gu YQ (2016) Rapid evolutionary dynamics in a 2.8-Mb chromosomal region containing multiple prolamin and resistance gene families in *Aegilops tauschii*. The Plant Journal 87(5): 495–506. http://doi.org/10.1111/tpj.13214

Dudnikov AJ (2018) Polymorphism of gliadins in *Aegilops tauschii* Coss. local populations in two primary habitats in Dagestan. Genetic Resources and Crop Evolution 65(3): 845–854. https://doi.org/10.1007/s10722-017-0575-4

Dvořák J, Zhang H-B, Kota RS, Lassner M (1989) Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. Genome 32(6): 1003–1016. https://doi.org/10.1139/g89-545

Ebrahimzadegan R, Orooji F, Ma P, Mirzaghaderi G (2021) Differentially amplified repetitive sequences among *Aegilops tauschii* subspecies and genotypes. Frontiers in Plant Science 12: 716750–716750. http://doi.org/10.3389/fpls.2021.716750

Edet OU, Gorafi YSA, Nasuda S, Tsujimoto H (2018) DArTseq-based analysis of genomic relationships among species of tribe *Triticeae*. Scientific Reports 8: 16397. http://doi.org/10.1038/s41598-018-34811-y

Eig A (1929) Monographisch-kritische Ubersicht der Gattung *Aegilops*. Vol. 55,Beihefte, Berlin, 228 pp.

Friebe B, Badaeva ED, Hammer K, Gill BS (1996) Standard karyotypes of *Aegilops uniaristata*, *Ae. mutica*, *Ae. comosa* subspecies *comosa* and *heldreichii* (Poaceae). Plant Systematics and Evolution 202: 199–210. http://doi.org/10.1007/BF00983382

Friebe B, Mukai Y, Gill BS (1992) C-banding polymorphism in several accessions of *Triticum tauschii* (*Aegilops squarrosa*). Genome 35(2): 192–199. https://doi.org/10.1139/g92-030

Friebe B, Tuleen N, Jiang J, Gill BS (1993) Standard karyotype of *Triticum longissimum* and its cytogenetic relationship with *T. aestivum*. Genome 36(4): 731–742. http://doi.org/10.1139/g93-098

Friebe BR, Tuleen NA, Gill BS (1999) Development and identification of a complete set of *Triticum aestivum* – *Aegilops geniculata* chromosome addition lines. Genome 42(3): 374–380. http://doi.org/10.1139/g99-011

Garg M, Tsujimoto H, Gupta R, Kumar A, Kaur N, Kumar R, Ch V, Sharma N, Chawla M, Sharma S, Mundey J (2016) Chromosome specific substitution lines of *Aegilops geniculata* alter parameters of bread making quality of wheat. PLoS ONE 11: e0162350. http://doi.org/10.1371/journal.pone.0162350

Georgiou A, Karataglis S, Roupakias D (1992) Inter- and intraplant C-banding polymorphism in one population of *Aegilops comosa* subsp. *comosa* var. *comosa* (Poaceae). Plant Systematics and Evolution 180(1): 105–114. http://doi.org/10.1007/BF00940400

Gill BS, Sharma C, Raupp WJ, Browder LE, Heachett JH, Harvey TL, Moseman JG, Waines JG (1985) Evaluation of *Aegilops* species for resistance to wheat powdery mildew, wheat leaf rust, Hessian fly, and greenbug. Plant Disease 69(4): 314–316.

Hammer K (1980) Zur Taxonomie und Nomenklatur der Gattung *Aegilops* L. Feddes Reports 91: 225–258. https://doi.org/10.1002/fedr.19800910404

Huo N, Zhang S, Zhu T, Dong L, Wang Y, Mohr T, Hu T, Liu Z, Dvorak J, Luo M-C, Wang D, Lee J-Y, Altenbach S, Gu YQ (2018) Gene duplication and evolution dynamics in the homeologous regions harboring multiple prolamin and resistance gene families in hexaploid wheat. Frontiers in Plant Science 9: 673. http://doi.org/10.3389/fpls.2018.00673

Iqbal N, Reader SM, Caligari PDS, Miller TE (2000) Characterization of *Aegilops uniaristata* chromosomes by comparative DNA marker analysis and repetitive DNA sequence in situ hybridization. Theoretical and Applied Genetics 101(8): 1173–1179. https://doi.org/10.1007/s001220051594

Jiang M, Xaio ZQ, Fu SL, Tang ZX (2017) FISH karyotype of 85 common wheat cultivars/lines displayed by ND-FISH using oligonucleotide probes. Cereal Research Communications 45(4): 549–563. http://doi.org/10.1556/0806.45.2017.049

Karataglis S (1975) Karyotype analysis on some diploid native Greek *Aegilops* species. Cytologia 28(1): 99–110. http://doi.org/10.1080/00087114.1975.10796601

Keskin Şan S, Özbek Ö, Eser V, Göçmen Taşkin B (2015) Polymorphism in seed endosperm proteins (gliadins and glutenins) of Turkish cultivated einkorn wheat [*Triticum monococcum* ssp. *monococcum*] landraces. Cereal Research Communications 43(1): 108–122. http://doi.org/10.1556/crc.2014.0028

Kibbe WA (2007) OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Research 35(suppl.2): W43–W46. https://doi.org/10.1093/nar/gkm234

Kihara H (1940) Verwandtschaft der *Aegilops*-Arten im Lichte der Genomanalyse. Ein Überblick. Der Zuechter 12(3): 49–62. https://doi.org/10.1007/BF01812302

Kihara H (1947) The genus *Aegilops* classified on the basis of genome analysis. Seiken Ziho 3: 7–25.

Kihara H (1954) Considerations on the evolution and distribution of *Aegilops* species based on the analyser-method. Cytologia 19(4): 336–357. https://doi.org/10.1508/cytologia.19.336

Kilian B, Mammen K, Millet E, Sharma R, Graner A, Salamini F, Hammer K, Ozkan H (2011) *Aegilops*. In: Kole C (Ed.) Wild crop relatives: Genomics and breeding resources Cereals. Springer, Berlin Heidelberg, 1–76. http://doi.org/10.1007/978-3-642-14228-4

Kimber G, Pignone D, Sallee PJ (1983) The relationships of the M and Mᵘ genomes of *Triticum*. Canadian Journal of Genetics and Cytology 25(5): 509–512. https://doi.org/10.1139/g83-076

Kimber G, Sallee PJ, Feiner MM (1988) The interspecific and evolutionary relationships of *Triticum ovatum*. Genome 30(2): 218–221. https://doi.org/10.1139/g88-037

Kimber G, Tsunewaki K (1988) Genome symbols and plasma types in the wheat group. In: Miller TE, Koebner RMD (Eds) Proceedings of the 7th International Wheat Genetics Symposium, 13–19 July 1988. Bath Press, Avon, Cambridge, England, 1209–1210.

Kroupin PY, Badaeva ED, Sokolova VM, Chikida NN, Belousova MK, Surzhikov SA, Nikitina EA, Kocheshkova AA, Ulyanov DS, Ermolaev AS, Khuat TML, Razumova OV, Yurkina AI, Karlov GI, Divashuk MG (2022) *Aegilops crassa* Boiss. repeatome characterized using low-coverage NGS as a source of new FISH markers: Application in phylogenetic studies of the Triticeae. Frontiers in Plant Science 13: 980764. http://doi.org/10.3389/fpls.2022.980764

Lang T, Li G, Yu Z, Ma J, Chen Q, Yang E, Yang Z (2019) Genome-wide distribution of novel Ta-3A1 mini-satellite repeats and its use for chromosome identification in wheat and related species. Agronomy 9(2): 60. https://doi.org/10.3390/agronomy9020060

Li D, Li T, Wu Y, Zhang X, Zhu W, Wang Y, Zeng J, Xu L, Fan X, Sha L, Zhang H, Zhou Y, Kang H (2018) FISH-based markers enable identification of chromosomes derived from tetraploid *Thinopyrum elongatum* in hybrid lines. Frontiers in Plant Science 9: 526. http://doi.org/10.3389/fpls.2018.00526

Li G, Zhang T, Yu Z, Wang H, Yang E, Yang Z (2021) An efficient Oligo-FISH painting system for revealing chromosome rearrangements and polyploidization in Triticeae. The Plant Journal 105(4): 978–993. http://doi.org/10.1111/tpj.15081

Linc G, Gaál E, Molnár I, Icsó D, Badaeva E, Molnár-Láng M (2017) Molecular cytogenetic (FISH) and genome analysis of diploid wheatgrasses and their phylogenetic relationship. PLoS ONE 12: e0173623. http://doi.org/10.1371/journal.pone.0173623

Liu C, Gong W, Han R, Guo J, Li G, Li H, Song J, Liu A, Cao X, Zhai S, Cheng D, Li G, Zhao Z, Yang Z, Liu J, Reader SM (2019) Characterization, identification and evaluation of a set of wheat-*Aegilops comosa* chromosome lines. Scientific Reports 9: 4773. http://doi.org/10.1038/s41598-019-41219-9

Luo X, Tinker N, Zhou Y, Liu J, Wan W, Chen L (2018) Chromosomal distributions of oligo-Am1 and (TTG)$_6$ trinucleotide and their utilization in genome association analysis of sixteen *Avena* species. Genetic Resources and Crop Evolution 65(6): 1625–1635. http://doi.org/10.1007/s10722-018-0639-0

Luo X, Tinker NA, Zhou Y, Wight CP, Liu J, Wan W, Chen L, Peng Y (2017) Genomic relationships among sixteen species of *Avena* based on (ACT)$_6$ trinucleotide repeat FISH. Genome 61(1): 63–70. http://doi.org/10.1139/gen-2017-0132

Majka M, Kwiatek MT, Majka J, Wisniewska H (2017) *Aegilops tauschii* accessions with geographically diverse origin show differences in chromosome organization and polymorphism of molecular markers linked to leaf rust and powdery mildew resistance genes. Frontiers in Plant Science 8: 1149. http://doi.org/10.3389/fpls.2017.01149

Masci S, D'Ovidio R, Lafiandra D, Tanzarella OA, Porceddu E (1992) Electrophoretic and molecular analysis of alpha-gliadins in *Aegilops* species (Poaceae) belonging to the D-genome cluster and in their putative progenitors. Plant Systematics and Evolution 179(1–2): 115–128. http://doi.org/10.1007/bf00938024

Medouri A, Bellil I, Khelifi D (2015) The genetic diversity of gliadins in *Aegilops geniculata* from Algeria. Czech Journal of Genetics & Plant Breeding 51(1): 9–15. http://doi.org/10.17221/158/2014-CJGPB

Metakovsky E, Melnik V, Rodriguez-Quijano M, Upelniek V, Carrillo JM (2018) A catalog of gliadin alleles: Polymorphism of 20th-century common wheat germplasm. The Crop Journal 6(6): 628–641. https://doi.org/10.1016/j.cj.2018.02.003

Metakovsky EV, Kudryavtsev AM, Iakobashvili ZA, Novoselskaya AY (1989) Analysis of phylogenetic relations of durum, carthlicum and common wheats by means of comparison of alleles of gliadin-coding loci. Theoretical and Applied Genetics 77(6): 881–887. https://doi.org/10.1007/BF00268342

Metakovsky EV, Novoselskaya AY (1991) Gliadin allele identification in common wheat 1. Methodological aspects of the analysis of gliadin patterns by one-dimensional polyacrylamide gel electrophoresis. Journal of Genetics and Breeding 45(4): 317–324.

Mitrenina EYu, Erst AS, Badaeva ED, Alexeeva SS, Artyomov GN (2021) Molecular-cytogenetic study of *Erantus* (Ranunculaceae). Botanical Probelems of Southern Siberia and Mongolia 20(1): 305–308. https://doi.org/10.14258/pbssm.2021060

Molnár I, Kubaláková M, Šimková H, Cseh A, Molnár-Láng M, Doležel J (2011) Chromosome isolation by flow sorting in *Aegilops umbellulata* and *Ae. comosa* and their allotetraploid hybrids *Ae. biuncialis* and *Ae. geniculata*. PLoS ONE 6: e27708. http://doi.org/10.1371/journal.pone.0027708

Molnár I, Molnár-Láng M (2010) Visualization of U and M genome chromosomes by multicolour genomic in situ hybridization in *Aegilops biuncialis* and *Triticum aestivum-Ae. biuncialis* amphiploids. Acta Agronomica Hungarica 58(3): 195–202. https://doi.org/10.1556/AAgr.58.2010.3.1

Molnár I, Vrána J, Burešová V, Cápal P, Farkas A, Darkó É, Cseh A, Kubaláková M, Molnár-Láng M, Doležel J (2016) Dissecting the U, M, S and C genomes of wild relatives of bread wheat (*Aegilops* spp.) into chromosomes and exploring their synteny with wheat. The Plant Journal 88(3): 452–467. http://doi.org/10.1111/tpj.13266

Nasuda S, Friebe B, Busch W, Kynast RG, Gill BS (1998) Structural rearrangement in chromosome 2M of *Aegilops comosa* has prevented the utilization of the Compair and related wheat-*Ae. comosa* translocations in wheat improvement Theoretical and Applied Genetics 96(6–7): 780–785. https://doi.org/10.1007/s001220050802

Novoselskaya-Dragovich AY, Yankovskaya AA, Badaeva ED (2018) Alien introgressions and chromosomal rearrangements do not affect the activity of gliadin-coding genes in hybrid lines of *Triticum aestivum* L. × *Aegilops columnaris* Zhuk. Vavilov Journal of Genetics and Breeding 22(5): 507–514. http://doi.org/10.18699/VJ18.388

Pedersen C, Langridge P (1997) Identification of the entire chromosome complement of bread wheat by two-colour FISH. Genome 40(5): 589–593. https://doi.org/10.1139/g97-077

Pomortsev AA, Martynov SP, Kovaleva ON, Lyalina EV (2011) Polymorphism of hordein-coding loci in barley (*Hordeum vulgare* L.) populations of Iran and Central Asian countries. Russian Journal of Genetics 47(11): 1372–1390. http://doi.org/10.1134/S1022795411110147

Resta P, Zhang HB, Dubkovsky J, Dvořák J (1996) The origin of the genomes of *Triticum biunciale*, *T. ovatum*, *T. neglectum*, *T. columnare*, and *T. rectum* based on variation in repeated nucleotide sequences. American Journal of Botany 83(12): 1556–1565. https://doi.org/10.1002/j.1537-2197.1996.tb12813.x

Riley R, Chapman V, Johnson ROY (1968) Introduction of yellow rust resistance of *Aegilops comosa* into wheat by genetically induced homoeologous recombination. Nature 217(5126): 383–384. http://doi.org/10.1038/217383a0

Said M, Cápal P, Farkas A, Gaál E, Ivanizs L, Friebe B, Doležel J, Molnár I (2022) Flow karyotyping of wheat-*Aegilops* additions facilitate dissecting the genomes of *Ae. biuncialis* and *Ae. geniculata* into individual chromosomes. Frontiers in Plant Science 13: 1017958. http://doi.org/10.3389/fpls.2022.1017958

Said M, Holušová K, Farkas A, Ivanizs L, Gaál E, Cápal P, Abrouk M, Martis-Thiele MM, Kalapos B, Bartoš J, Friebe B, Doležel J, Molnár I (2021) Development of DNA markers from physically mapped loci in *Aegilops comosa* and *Aegilops umbellulata* using single-gene FISH and chromosome sequences. Frontiers in Plant Science 12: 689031. http://doi.org/10.3389/fpls.2021.689031

Shewry PR, Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. Journal of Experimental Botany 53(370): 947–958. http://doi.org/10.1093/jexbot/53.370.947

Song Z, Dai S, Bao T, Zuo Y, Xiang Q, Li J, Liu G, Yan Z (2020) Analysis of structural genomic diversity in *Aegilops umbellulata*, *Ae. markgrafii*, *Ae. comosa*, and *Ae. uniaristata* by Fluorescence in situ hybridization karyotyping. Frontiers in Plant Science 11: 710. http://doi.org/10.3389/fpls.2020.00710

Tang S, Qiu L, Xiao Z, Fu S, Tang Z (2016) New oligonucleotide probes for ND-FISH analysis to identify barley chromosomes and to investigate polymorphisms of wheat chromosomes. Genes 7: 118. http://doi.org/10.3390/genes7120118

Tang S, Tang Z, Qiu L, Yang Z, Li G, Lang T, Zhu W, Zhang J, Fu S (2018a) Developing new oligo probes to distinguish specific chromosomal segments and the A, B, D genomes of wheat (*Triticum aestivum* L.) using ND-FISH. Frontiers in Plant Science 9: 1104. http://doi.org/10.3389/fpls.2018.01104

Tang S, Tang Z, Qiu L, Yang Z, Li G, Lang T, Zhu W, Zhang J, Fu S (2018b) Developing new oligo probes to distinguish specific chromosomal segments and the A, B, D Genomes of wheat (*Triticum aestivum* L.) using ND-FISH. Frontiers in Plant Science 9: 1104. http://doi.org/10.3389/fpls.2018.01104

Tang Z, Yang Z, Fu S (2014) Oligonucleotides replacing the roles of repetitive sequences pAs1, pSc119.2, pTa-535, pTa71, CCS1, and pAWRC.1 for FISH analysis. Journal of Applied Genetics 55(3): 313–318. http://doi.org/10.1007/s13353-014-0215-z

Teoh SB, Hutchinson J (1983) Interspecific variation in C-banded chromosomes of diploid *Aegilops* species. Theoretical and Applied Genetics 65(1): 31–40. http://doi.org/10.1007/BF00276259

Teoh SB, Miller TE, Reader SM (1983) Intraspecific variation in C-banded chromosomes of *Aegilops comosa* and *Ae. speltoides*. Theoretical and Applied Genetics 65(4): 343–348. http://doi.org/10.1007/BF00276575

Tsunewaki K (1996) Plasmon analysis as the counterpart of genome analysis. In: Jauhar PP (Ed.) Methods of Genome Analysis in Plant: Their Merits and Piffals. CRC Press, Boca Ration, 271–299.

Van Slageren MW (1994) Wild Wheats: a monograph of *Aegilops* L. and *Amblyopyrum* (Jaub. et Spach) Eig (*Poaceae*). Wageningen Agricultural University, Wageningen and ICARDA, Aleppo, Syria, Wageningen, 514 pp.

Woychik JH, Boundy JA, Dimler RJ (1961) Starch gel electrophoresis of wheat gluten proteins with concentrated urea. Archives of Biochemistry and Biophysics 94(3): 477–482. https://doi.org/10.1016/0003-9861(61)90075-3.

Weng Y (1995) Development of *Aegilops comosa* addition lines in common wheat (*Triticum aestivum* L.) I. Effection of wheat anther culture to development of *Aegilops comosa* addition lines in common wheat. Acta Agronomica Sinica 21(1): 39–44.

Xu X, Monneveux P, Damania AB, Zaharieva M (1996) Evaluation for salt tolerance in genetic resources of *Triticum* and *Aegilops* species. Bulletin des Ressources Genetiques Vegetales (CIRP/FAO); Noticiario de Recursos Geneticos Vegetales (CIRF/FAO): 11–16.

Yan YM, Hsam SLK, Yu JZ, Jiang Y, Zeller FJ (2003) Genetic polymorphisms at *Gli-D-t* gliadin loci in *Aegilops tauschii* as revealed by acid polyacrylamide gel and capillary electrophoresis. Plant Breeding 122(2): 120–124. http://doi.org/10.1046/j.1439-0523.2003.00824.x

Yu Z, Wang H, Xu Y, Li Y, Lang T, Yang Z, Li G (2019) Characterization of chromosomal rearrangement in new wheat – *Thinopyrum intermedium* addition lines carrying *Thinopyrum* – specific grain hardness genes. Agronomy 9(1): 18. http://doi.org/10.3390/agronomy9010018

Zhang W, Tang Z, Luo J, Guangrong L, Yang Z, Yang M, Yang E, Fu S (2022) Location of tandem repeats on wheat chromosome 5B and the breakpoint on the 5BS arm in wheat translocation T7BS.7BL-5BS using single-copy FISH analysis. Plants (Basel, Switzerland) 11(18) 2394: http://doi.org/10.3390/plants11182394

Zhao L, Ning S, Yi Y, Zhang L, Yuan Z, Wang J, Zheng Y, Hao M, Liu D (2018) Fluorescence in situ hybridization karyotyping reveals the presence of two distinct genomes in the taxon *Aegilops tauschii*. Bmc Genomics 19: 3. http://doi.org/10.1186/s12864-017-4384-0

Zhukovsky PM (1928) A critical-systematical survey of the species of the genus *Aegilops* L. Bul Appl Bot Genet Pl Breed 18: 417–609. [In Russian]

Zuo Y, Xiang Q, Dai S, Song Z, Bao T, Hao M, Zhang L, Liu G, Li J, Liu D, Wei Y, Zheng Y, Yan Z (2020) Development and characterization of *Triticum turgidum* – *Aegilops comosa* and *T. turgidum* – *Ae. markgrafii* amphidiploids. Genome 63(5): 263–273. http://doi.org/10.1139/gen-2019-0215

ORCID

Ekaterina D. Badaeva https://orcid.org/0000-0001-7101-9639
Violetta V. Kotseruba https://orcid.org/0000-0003-1872-2223
Andnrey V. Fisenko https://orcid.org/0000-0001-9063-1145
Nadezhda N. Chikida https://orcid.org/0000-0002-9698-263X
Maria Kh. Belousova https://orcid.org/0000-0003-0980-3531
Peter M. Zhurbenko https://orcid.org/0000-0002-2102-4568
Sergei A. Surzhikov https://orcid.org/0000-0002-6043-1182
Alexandra Yu. Dragovich https://orcid.org/0000-0002-9731-0106

Supplementary material 1

Variation of hybridization patterns of pTa794 and o-18S or pTa71 rDNA probes

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: figure (TIF-file)

Explanation note: Variation of hybridization patterns of pTa794 (red) and o-18S (a, h, m, v–y) or pTa71 (b–g, i–l, n–u) rDNA probes (green) on chromosomes of following accessions of *Ae. comosa* (c01–c17 – subsp. *comosa*; h01–h14 – subsp. *heldreichii*): c01 – K-3819; c02 – K-3820; c03 – K-3781; c04 – K-3920; c05 – K-3810; c06 – AE 1254; c07 – AE 1257; c08 – AE 1258; c09 – AE 1259; c10 – AE 1376; c11 – AE 1377; c12 – AE 1378; c13 – K-3308; c14 – K-3787; c15 – K-3857; c16 – K-3809; c17 – K-3780; h01 – K-3804; h02 – K-1601; h03 – K-3806; h04 – K-2432; h05 – K-3911; h06 – K-3897; h07 – K-3919; h08 – K-4498; h09 – K-3914; h10 – K-3809; h11 – K-2272; h12 – AE 783; h13 – K-669; h14 – K-3824. Green arrows point to inactivated major NORs. Position of minor 45S rDNA loci specific for either *comosa* or *heldreichii* group are underlines with green lines. The 5S rDNA sites are shown with pink lines. Translocated 1M-6M chromosomes are arrowed.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl1

Supplementary material 2

Distribution of (GAA)$_{10}$ microsatellite probe on chromosomes of different accessions of *Ae. comosa* subsp. *comosa*

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: figure (TIF-file)

Explanation note: Distribution of (GAA)$_{10}$ microsatellite probe (green) on chromosomes of different accessions of *Ae. comosa* subsp. *comosa*: a – K-3810; b – K-3781; c – K-3787; d, e – K-3308; f – K-3309; g – AE 1256; h, i – AE 1254; j – AE 1257; k – AE 1258; l – AE 1259; m – AE 1378; n – K-3810; o – AE 1376; p – AE 1378; q – AE 1377; r – K-3857; s – K-3920; t – K-3820; u – K-3909; v, w – K-3809, x – K-3819. Translocated 1M-6M chromosomes are indicated. Position of (GTT)$_n$ sites (red) on chromosome 4M of AE 1256 (g) is shown with red arrows.

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl2

Supplementary material 3

Distribution of (GAA)$_{10}$ microsatellite probe on chromosomes of different accessions of *Ae. comosa* subsp. *heldreichii*

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: figure (TIF-file)

Explanation note: Distribution of (GAA)$_{10}$ microsatellite probe (green) on chromosomes of different accessions of *Ae. comosa* subsp. *heldreichii*: a – AE 117; b – K-3811; c – AE 783; d-f – K-1601; g – K-3919; h – K-3804; i – K-669; j – K-4873; k, l – K-2272; m – K-2432; n, o – K-3914; p – K-3897; q – K-4498. Translocated 1M^h-6M^h chromosomes are indicated. Lane (e) presents karyotype of hybrid plant of K-1601, where "c" indicates homologous chromosomes of "*comosa*" type and "h" – homologous chromosome of *heldreichii* type. Localization of (GTT)$_{10}$ probe (red) is shown with red arrows for accession AE 117 (a).

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl3

Supplementary material 4

Distribution of oligo-42, oligo-44, oligo-45, (AC)$_{20}$, and pTa-k566 probes on chromosomes of *Ae. comosa* subsp. *comosa* and subsp. *heldreichii*

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: figure (TIF-file)

Explanation note: Distribution of oligo-42, oligo-44, oligo-45, (AC)$_{20}$, and pTa-k566 probes on chromosomes of *Ae. comosa* subsp. *comosa* (a-h) and subsp. *heldreichii* (i-p): a – K-3824; b – AE 1377; c – AE 1257; d – AE 1378; e – K-3820; f – AE 1258; g – K-3781; h – K-3819; i – K-3914; j – K-4873; k – K-3806; l – AE 783; m – K-3811; n – K-2432; o – K-669; p – K-1601. Probe names are shown on the top; probe color corresponds to signal color. Subspecies specific sites are underlined. Pink arrow points to heteromorphic signal.

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl4

Supplementary material 5

Distribution of pTa-713 probe on chromosomes of different accessions of *Ae. comosa* subsp. *comosa* from Greece and Turkey

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: figure (TIF-file)

Explanation note: Distribution of pTa-713 probe (red) on chromosomes of different accessions of *Ae. comosa* subsp. *comosa* from Greece (a–o) and Turkey (p–t): a – AE 1258; b – AE 1254; c – AE 1256; d – AE 1257; e – AE 115; f – AE 1259; g – AE 1376; h – AE 1377; i – K-3810; j – K-3820; k – K-3819; l – AE 1378; m – K-3809; n – K-3808; o – K-3857; p – K-3309; q – K-3780; r – K-3781; s – K-3308; t – K-3787. Localization of (GAA)$_{10}$ probe (green) is shown for accession AE 1258 (a). Positions of uncommon pTa-713 sites are arrowed.

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl5

Supplementary material 6

Distribution of pTa-713 probe on chromosomes of different accessions of *Ae. comosa* subsp. *heldreichii*

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: figure (TIF-file)

Explanation note: Distribution of pTa-713 probe (red) on chromosomes of different accessions of *Ae. comosa* subsp. *heldreichii*: a – K-3414; b – K-3919; c – K-3897; d – K-1601; e – K-3811; f – K-2432; g – K-3804; h – K-4498; i – AE 783; j – K-4873; k – K-2272; l – AE 117; m –K-669. Translocated $1M^h$-$6M^h$ chromosomes are indicated. Localization of $(GAA)_{10}$ probe (green) is shown for accession K-3914 (a).

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl6

Supplementary material 7

Distribution of pSc119.2 and pAs1 or pTa-535 (t) probes on chromosomes of *Ae. comosa* subsp. *comosa* and subsp. *heldreichii*

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: figure (TIF-file)

Explanation note: Distribution of pSc119.2 (green) and pAs1 (a-s) or pTa-535 (t) (red) probes on chromosomes of *Ae. comosa* subsp. *comosa* (a-j) and subsp. *heldreichii* (k-t): a – AE 115; b – AE 1254; c – AE 1256; d – AE 1257; e – AE 1258; f – AE 1259; g – AE 1377; h – K-3309; i – AE 1378; j – K-3857; k – K-3914; l – K-2432; m, n – K-1601; o – AE 117; p – K-4873; q – K-783; r – K-3897; s – K-3811; t – K-669. Translocated $1M^h$-$6M^h$ chromosomes are arrowed.

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl7

Supplementary material 8

Distribution of pTa794 and oligo-pTa71 or o-18S probes on chromosomes of *Ae. comosa* subsp. *heldreichii* and *Ae. crassa*

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: figure (TIF-file)

Explanation note: Distribution of pTa794 (red, a-c) and oligo-pTa71 (b) or o-18S (green, a, c) probes on chromosomes of *Ae. comosa* subsp. *heldreichii* (a, b) and *Ae. crassa* 6x, IG 131680 (c). a – c: Position of major NORs visualized with o-18S probe are shown with yellow arrows; position of major and minor NORs visualized with o-pTa71 probe are shown with red arrows. d – alignment of different variants of 18S rDNA fragments identified in *Ae. tauschii* genome using blast of NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra).

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl8

Supplementary material 9

List of *Ae. comosa* accessions and their origin

Authors: Ekaterina D. Badaeva, Violetta V. Kotseruba, Andnrey V. Fisenko, Nadezhda N. Chikida, Maria Kh. Belousova, Peter M. Zhurbenko, Sergei A. Surzhikov, Alexandra Yu. Dragovich

Data type: table (.docx file)

Link: https://doi.org/10.3897/compcytogen.17.101008.suppl9

More hidden diversity in a cryptic species complex: a new subspecies of *Leptidea sinapis* (Lepidoptera, Pieridae) from Northern Iran

Vazrick Nazari[1], Vladimir A. Lukhtanov[2], Alireza Naderi[3],
Zdenek Faltýnek Fric[4], Vlad Dincă[5,6], Roger Vila[7]

1 *Department of Biology, University of Padova, Padova, Italy* **2** *Department of Karyosystematics, Zoological Institute of Russian Academy of Science, Universitetskaya nab. 1, 199034 St. Petersburg, Russia* **3** *National Natural History Museum & Genetic Resources, Tehran, Iran* **4** *Department of Biodiversity and Conservation Biology, Institute of Entomology, Biology Centre of the Czech Academy of Sciences, České Budějovice, Czech Republic* **5** *Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland* **6** *Research Institute of the University of Bucharest (ICUB), University of Bucharest, Bucharest, Romania* **7** *Institut de Biologia Evolutiva (CSIC – Universitat Pompeu Fabra), Barcelona, Spain*

Corresponding author: Vazrick Nazari (nvazrick@yahoo.com)

Abstract

A new subspecies of *Leptidea sinapis* from Northern Iran, discovered by means of DNA barcoding, is described as *Leptidea sinapis tabarestana* **ssp. nov.** The new subspecies is allopatric with respect to other populations of *L. sinapis* and is genetically distinct, appearing as a well-supported sister clade to all other populations in COI-based phylogenetic reconstructions. Details on karyotype, genitalia, ecology and behaviour for the new subspecies are given and a biogeographical speciation scenario is proposed.

Keywords

allopatry, butterflies, DNA barcoding, Palearctic, taxonomy, Wood White

Introduction

The cryptic diversity within the *Leptidea sinapis* (Linnaeus, 1758) complex progressively came to light in recent history (Réal 1988) with the discovery of differences in genitalic morphology (Lorković 1993) and allozyme markers (Martin et al. 2003) between *L. sinapis* and *L. reali* Reissinger, 1989. It is considered one of the first documented cases of cryptic species in Europe. Since then, numerous studies have revealed a plethora of new information on the mechanisms of speciation within this species complex (Mazel 2005; Bolshakov 2006; Friberg et al. 2008), including the presence of an additional widespread hidden taxon, *L. juvernica* Williams, 1946 (e.g. Dincă et al. 2011, 2013, 2021; Lukhtanov et al. 2011; Šíchová et al. 2015, 2016; Vodă et al. 2015; Shtinkov et al. 2016; Talla et al. 2017, 2019a, b; Leal et al. 2018; Platania et al. 2020; Yoshido et al. 2020; Näsvall et al. 2021). Despite the explosion of interest in this group, many regions of Eurasia where *Leptidea* species occur are still not well sampled or studied. The new subspecies described in this paper was discovered accidentally in the course of a genetic investigation in order to determine whether any of the populations of *L. sinapis* in Iran belong to the related cryptic species *L. juvernica*.

Materials and methods

Fourteen Iranian specimens from various disjunct populations in NW and N Iran were selected ex. coll. A. Naderi (Tehran) and W. ten Hagen (Germany) and their legs were submitted for DNA barcoding. Samples were processed in the Center for Biodiversity Genomics in Guelph, Ontario, Canada using standard protocols and LepF/LepR primers, supplemented by failure-tracking with mini-primers (mLepF and mLepR) (Hajibabaei et al. 2006). Eleven additional samples from Javaherdeh (VLU396-VLU405, RVcoll10C196) sequenced in 2012 were later added to the dataset. The majority of these sequences were full length barcodes (658 bp). An additional specimen from Javaherdeh included later in our analysis (MR ZF 449) was isolated using the Geneaid Blood and Tissue kit and sequenced in the Czech Republic using RON-HCO primers, and thus only partially overlaps (420 bp) with the standard barcode region. Thirty-six new barcode sequences were submitted to GenBank (Accessions OQ359842–OQ359877). In addition to the sequences pertaining to the new taxon, a selection of 80 other samples from previous studies (Lukhtanov et al. 2011; Dincă et al. 2013; Shtinkov et al. 2016) representing various haplotypes of *L. sinapis* and several other species of *Leptidea* was used to conduct the analyses in this study (Suppl. material 1). All records are publicly available in the BOLD dataset "DS-SINIRAN" (https://doi.org/10.5883/DS-SINIRAN).

A Maximum Likelihood (ML) tree was generated with PHYML online (Guindon and Gascuel 2003) using the AIC criterion and 100 bootstrap replicates. The best-fit model selected by PHYML for the combined dataset (GTR + Γ + I) was further corroborated by IQ-TREE (Nguyen et al. 2015), and parameters from this model were used to conduct a Bayesian analysis in MRBAYES 3.2.6 (Ronquist et al. 2012). The MCMC analysis was allowed to run for 10,000,000 generations until stationary was

reached. Convergence of parameters after the exclusion of the burnin phase was tested using TRACER 1.7.1 (Rambaut et al. 2018). Trees were edited using FIGTREE 1.4.4 (Rambaut 2018). Genetic distances were calculated using the Maximum Composite Likelihood model in MEGA 11.0.8 (Tamura et al. 2021). A haplotype diagram only including *L. sinapis*, *L. juvernica* and *L. reali* was constructed in TCS 1.21 (Clement et al. 2000), with a 95% confidence limit for parsimony. Shorter barcode fragments or those with ambiguous bases were excluded from haplotype analyses.

Male genitalia were examined following maceration in 10% potassium hydroxide (KOH) for 15 minutes at 95 °C, dissection and cleaning under a stereomicroscope and storage in tubes with glycerol. Male genitalia were photographed in a thin layer of 30% ethanol (without being pressed under a cover slip), using a Carl Zeiss Stemi 2000-C stereomicroscope equipped with a CMEX PRO-5 DC.5000p digital camera (RV) or a Leica DFC450 digital camera (ZFF). Care was taken to arrange the measured structures parallel to the focal plane of the stereomicroscope in order to minimize the measurement error. Measurements were performed based on digital photographs using the AxioVision software (Carl Zeiss MicroImaging GmbH). Eight specimens were analysed and the dataset was combined with data from Dincă et al. (2011) (135 specimens). We measured three elements of the male genitalia: phallus length (PL), saccus length (SL) and vinculum width (VW), known to be the most informative for differentiating *Leptidea* species (e.g. Dincă et al. 2011; Shtinkov et al. 2016) (Suppl. material 2). Bivariate scatterplots were generated using VW as a size variable (Shtinkov et al. 2016).

Chromosome preparations were made for ten adult males representing the population from Javaherdeh (field codes VLU396-VLU405) and were processed as previously described (Vershinina and Lukhtanov 2010). Briefly, gonads were removed from the abdomen and placed into freshly prepared fixative (3:1; 96% ethanol and glacial acetic acid) directly after capturing the butterfly in the field. Testes were stored in the fixative for 3–36 months at +4 °C. Then the gonads were stained in 2% acetic orcein for 30–60 days at +18–20 °C. Metaphase II (MII) and mitotic plates were examined using the original two-phase method of chromosome analysis (Lukhtanov et al. 2020a). Abbreviation "ca" (circa) means that the count was made with approximation due to overlapping of some chromosomes or due to difficulties in distinguishing between chromosome bivalents and trivalents. Images were edited in open source software GIMP 2.10.32 (The GIMP Development Team 2019) and Inkscape X11 (Inkscape Project 2020). Map was created using Simplemappr (Shorthouse 2010).

Results

None of the barcoded Iranian specimens belonged to *L. juvernica.* Specimens from the Iranian province of East Azerbaijan (Arasbaran) showed several haplotypes identical to those of the common and widespread Eurasian *L. sinapis*; however, samples collected across the Alborz mountains from Talesh to NE Iran represented a unique and well-supported COI clade that appeared as sister to a weakly-supported clade containing all other *L. sinapis* (Figs 1, 2). A comparison of average uncorrected pairwise
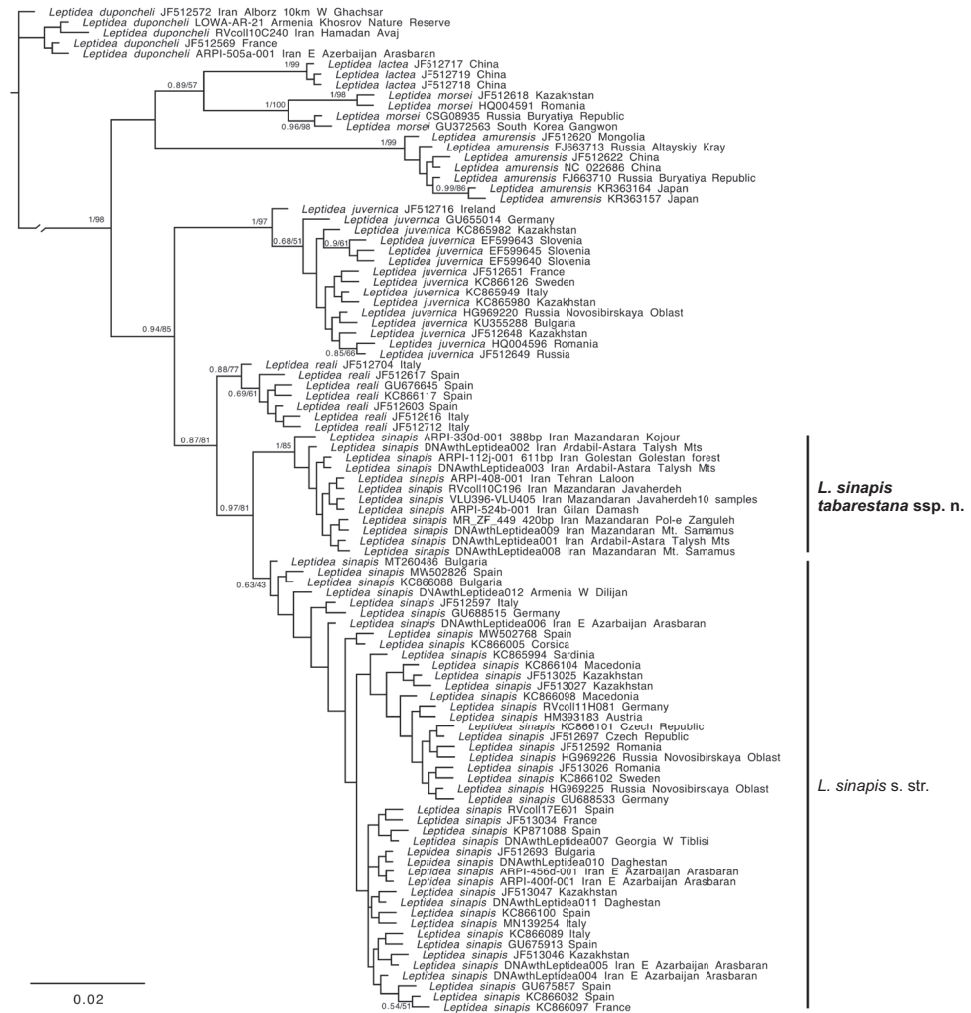
Figure 1. Bayesian phylogeny of *Leptidea* COI barcodes. Node support values (Bayesian Posterior Probebilities / ML bootstrap) are shown only for supported nodes. All sequences are 658 bp in length unless indicated otherwise.

distances between this new lineage and other *Leptidea* species showed that it is indeed genetically closer to *L. sinapis* (average: 0.74%; range: 0.42%–1.76%) and further from all the other *Leptidea* (Table 1).

The genitalia of the eight specimens analysed belonging to the above-mentioned COI lineage showed broad overlap with other specimens of *L. sinapis* and a certain degree of variability, despite their fairly restricted geographic origin (Fig. 3). Based on the three characters measured (PL, SL, VW), the male genitalia also indicated a close similarity to *L. sinapis*, with respect to which we did not notice any significant differences.
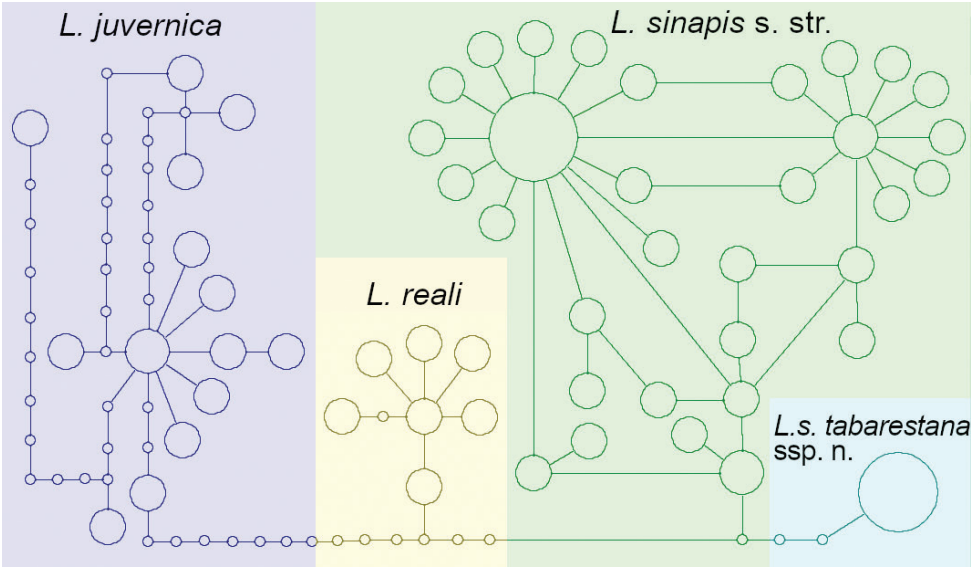
Figure 2. TCS haplotype network for *L. sinapis*, *L. reali* and *L. juvernica*.

Table 1. Average uncorrected *p*-distances (in % of the COI barcoding region) and standard deviation between *Leptidea* taxa.

	L. duponcheli	*L. lactea*	*L. morsei*	*L. amurensis*	*L. juvernica*	*L. reali*	*L. s. sinapis*	*L. s. tabarestana*
L. duponcheli (n=5)	0.27 ± 0.13							
L. lactea (n=3)	5.80 ± 0.13	0.00 ± 0.00						
L. morsei (n=4)	5.94 ± 0.26	2.33 ± 0.26	0.71 ± 0.44					
L. amurensis (n=7)	7.40 ± 0.14	4.23 ± 0.09	3.75 ± 0.13	0.26 ± 0.16				
L. juvernica (n=15)	6.29 ± 0.20	2.51 ± 0.16	3.39 ± 0.24	3.97 ± 0.15	0.30 ± 0.13			
L. reali (n=7)	5.35 ± 0.16	2.33 ± 0.13	2.96 ± 0.25	3.79 ± 0.15	1.75 ± 0.16	0.21 ± 0.07		
L. s. sinapis (n=44)	5.72 ± 0.18	2.71 ± 0.18	3.05 ± 0.21	3.74 ± 0.21	1.97 ± 0.21	0.92 ± 0.15	0.24 ± 0.11	
L. s. tabarestana (n=21)	5.69 ± 0.18	2.76 ± 0.16	2.78 ± 0.25	4.02 ± 0.13	2.00 ± 0.17	0.96 ± 0.19	**0.74 ± 0.20**	0.02 ± 0.05

Considering the allopatric distribution of the new taxon with respect to *L. sinapis*, its similar genitalia, and the fact that the new taxon appears to be genetically closer and phylogenetically sister to the rest of *L. sinapis* specimens, here we describe it as a new subspecies of *L. sinapis*:

Leptidea sinapis ssp. *tabarestana* Nazari, Lukhtanov et Naderi, ssp. nov.
https://zoobank.org/BED12A6B-C1D3-4897-8D40-A955333D6C7C
Fig. 4a–i

Type material. *Holotype.* ♂ [white label] "330d= Mazandaran- E Kojour-/Kodir – 1000 m – 2.Jul.[20]10- / leg. A.R. Naderi"; [red label] "Holotype/ *Leptidea sinapis tabarestana* / Nazari, Lukhtanov & Naderi 2023". BOLD Sample ID: ARPI-330d-001;
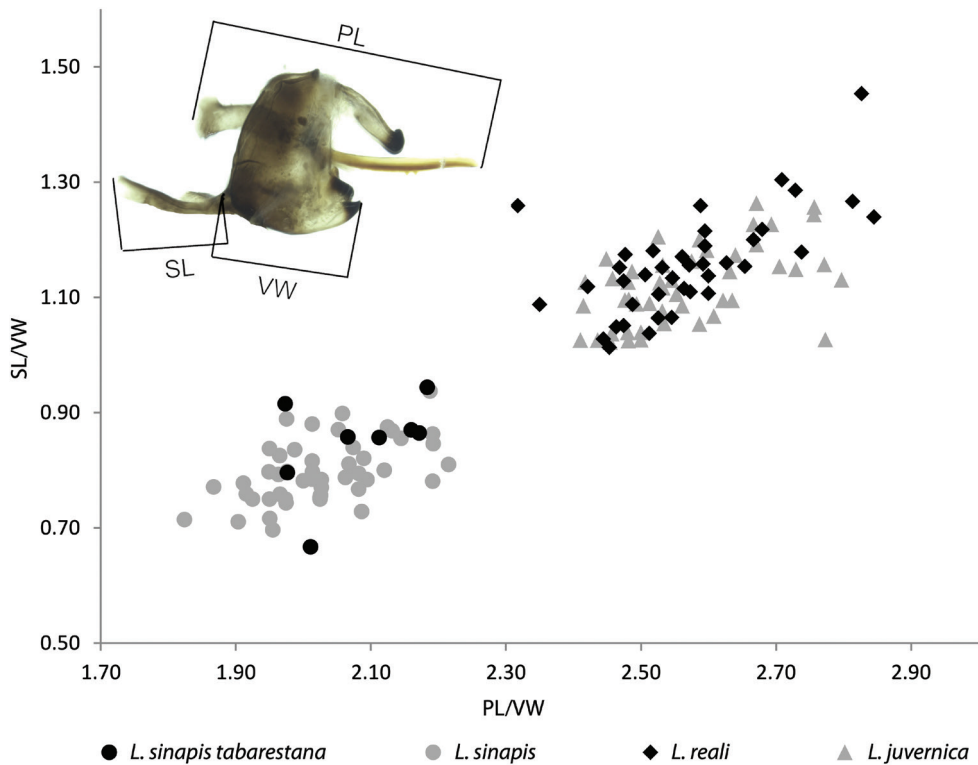
Figure 3. Bivariate scatterplot based on male genitalia morphometry (phallus length, PL; saccus length, SL), using vinculum width (VW) as a size variable. *L. s. tabarestana* ssp. nov. overlaps broadly with *L. s. sinapis* s. str., however it is distinct from *L. juvernica* and *L. reali.* Inset: Male genitalia of *L. sinapis tabarestana* ssp. nov. (specimen MR ZF 449), showing the variables measured.

Deposited in coll. National Natural History Museum & Genetic Resources of Department of Environment, Tehran, Iran.

Paratypes. **Gilan:** 1♂ Damash, 1200m, 26.III.2021, leg. et coll. A.R. Naderi (ARPI-524b-001, AR# 254); 1♀ Khoshkab, rd. Siyahkal-Deylaman, 02.VII.1990, leg. et coll. Harandi. **Ardabil/Gilan:** 2♂♂ 1♀ Paß Ardabil-Astara (Paßhöhe, W Tunnel), 1600m, 10.V.2010, W. ten Hagen. **Tehran:** 1♀ Laloon, 2000–2200 m, 30.VIII.2013, leg. et coll. A.R. Naderi (ARPI-408-001, AR# 186). **Mazandaran:** 1♂ Chalus road, Yush road, 40 km from Pole Zangooleh, 2400 m, 4.VII.1997, leg. & coll. A.R. Naderi (AR# 58); 2♂♂ Galanderoud, 1000 m, 13.VII.07, leg. & coll. A.R. Naderi; 1♂ Siahkal, 03.VII.1990, leg. et coll. Harandi; 1♂ Pol-e Zanguleh – Baladeh Rd, W of Minak, 36.2254°N, 51.58409°E, 15.V.2016, leg. & coll. Z. F. Fric, Biology Centre CAS, Institute of Entomology (IECA) (MR ZF 449); 10♂♂ Javaherdeh (Jirkooh), 36.866, 50.506, 24.VII.2011, leg. V. Lukhtanov & N. Shapoval, in Institut de Biologia Evolutiva (CSIC-UPF), Butterfly Diversity and Evolution Lab (VLU396-VLU405); 43♂♂, 12♀♀ *ibid*, in coll. Zoological Institute of Russian Academy of Sciences;
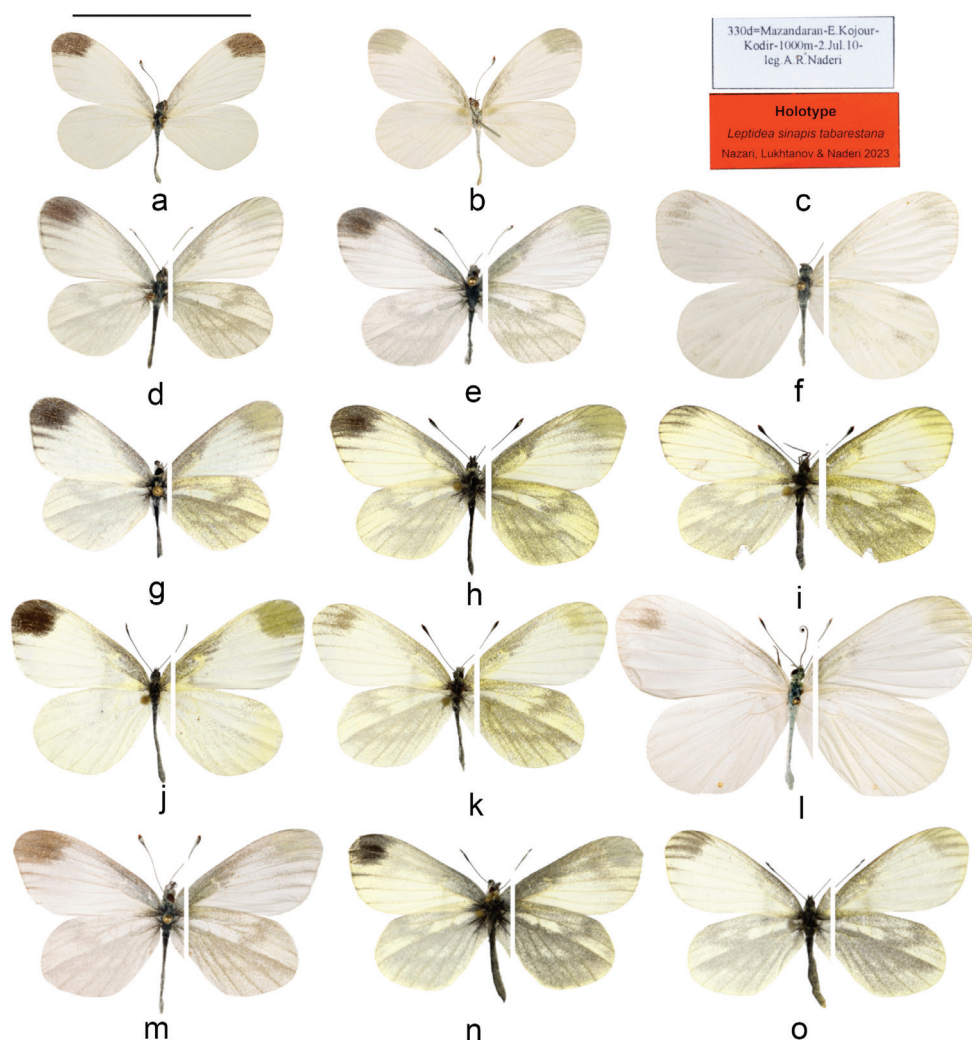
Figure 4. Adults **a–i** *L. sinapis* ssp. *tabarestana* **j–o** *L. sinapis* ssp. *sinapis* **a–c** holotype Mazandaran: Kojur (♂ ARPI-330d-001) **d** Golestan: Golestan forest (♂ ARPI-112j-001) **e** Gilan: Damash (♂ ARPI-524b-001) **f** Tehran: Laloon (♀ ARPI-408-001) **g** Mazandaran: Javaherdeh (♂ MR ZF 449) **h, i** Ardabil/Gilan: Talesh (♂♀ DNAwthLeptidea001–2) **j, k** Iran: E. Azerbaijan prov.: Kaleybar (♂ DNAwthLeptidea006, ♀-004) **l, m** Iran: E. Azerbaijan prov.: Arasbaran (♀ ARPI-479a [not barcoded], ♂ ARPI-456d-001) **n, o** Russia: Daghestan Republic (♂♀ DNAwthLeptidea010-11). Scale bar: 20 mm.

1♂, Javaherdeh (Samamus Mt.), 14.VIII.2010, leg. V.V. Tshikolovets, in Institut de Biologia Evolutiva (CSIC-UPF), Butterfly Diversity and Evolution Lab (RV-coll10C196); 1♂ Samamus Mt., 15 km S Ramsar, 1350 m, 8.VIII.2003, leg. W. ten Hagen; 1♂ Samamus Mt., S Rudbar, N Javaherdeh, 1500 m, 21.VI.2006, leg. W. ten Hagen; 1♂ Samamus Mt., 2800 m, 29.V.2009, leg. et coll. Harandi. **Golestan:** 1♂ Golestan Forest, 800–1000 m, 13.V.2001, leg. & coll. A.R. Naderi (112j, AR# 185).

Description. Male (Fig. 4a, b, d, e, g, h). Length of forewing 16–21 mm; ground colour pure white. ***First generation*** forewing upperside with a rectangular grey-black apical patch, veins v3 and v4 under this patch often covered with dark scales near the outer margin; forewing discal cell covered in grey scales that extend faintly along the costa towards the apex; a small dark patch near the base at the Inner margin. Hindwing upperside veins near the base of the wing covered with dark scales, otherwise without any other markings; the dark scales of the underneath show through. Forewing underside ground colour white with light yellowish-greenish tint at the apex, along the costa and at the discal cell except for a yellowish discoidal spot not covered in grey scales; all veins except v2 covered with dark scales at the outer half of the wing. Hindwing underside ground colour greenish-yellow covered in sparse grey scales; discal cell and space s5 lighter and covered in fewer dark scales; a faint postdiscal band broken into two sections: a costal S- shaped part and a lower postdiscal section in the form of a slightly curved streak. ***Second generation*** similar but grey scales on the underside highly reduced, sometimes completely absent.

Female (Fig. 4f, i). Length of forewing 19–23 mm; similar to male but bigger, forewing apex more rounded; apical dark patch highly reduced, sometimes absent.

Male genitalia (Fig. 3 inset). Based on the eight dissections examined, the male genitalia appear similar to that of the nominotypical *sinapis*. The three elements of the male genitalia (phallus length, saccus length and vinculum width) measured for *L. s. tabarestana* ssp. nov. (PL: 1.47±0.07, SL: 0.60±0.06, VW: 0.71±0.04, n=8) were comparable to those of the nominotypical *L. sinapis* (PL: 1.60±0.08, SL: 0.63±0.04, VW: 0.79±0.05, n=48) (Suppl. material 2).

Diagnosis. Morphologically inseparable from the nominotypical *L. sinapis*, however the new taxon is distinguishable from it only by COI barcodes. Unlike ssp. *sinapis*, which in Iran (East Azerbaijan province) is strictly limited to humid and damp forests or clearings, the new subspecies is found primarily in semi-humid or even semi-dry mountainous habitats.

Etymology. The subspecies name is a reference to "Tabarestan", the medieval name for the mountainous regions south of the Caspian coast in northern Iran and roughly corresponding to the modern-day province of Mazandaran, the type locality of *L. s. tabarestana* ssp. nov.

DNA barcode analysis. The COI barcodes of *L. s. tabarestana* ssp. nov. fall within the Barcode Identification Number (BIN) of *L. sinapis* (BOLD:AAA6298), however they form a unique and distinct cluster that is on average 0.74% (range: 0.42%–1.76%) distant from all other *L. sinapis* (Fig. 1). Uncorrected *p*-distances are smaller than those between *L. sinapis* and *L. reali* (0.92%) or between *L. sinapis* and *L. juvernica* (1.97%) (Table 1). Since the topology of ML and Bayesian trees were similar, only the Bayesian tree is shown with ML bootstrap values plotted on the supported nodes. In both trees, the *L. s. tabarestana* ssp. nov. clade appeared as sister to all other *L. sinapis* samples with strong support (Fig. 1).

Karyotype. Of the 10 specimens studied, only two samples demonstrated metaphase plates suitable for counting the number of chromosomes. Such a low proportion of adult males with dividing cells is a common phenomenon in the genus *Leptidea* and has been noted previously (Lukhtanov et al. 2011). In the sample VLU396, in mitotic cells, the diploid number of chromosomes was determined to be approximately

2n=ca 58. An exact diploid number could not be determined due to numerous over-laps or contacts of chromosomes (Fig. 5).

The MI metaphase cells were not found in the studied individuals; however, MII metaphase plates were found in the sample VLU405. The MII plates demonstrated clear traces of the phenomenon for which we previously used the term *inverted meiosis* (Lukhtanov et al. 2018; 2020a, b). In this type of meiosis, heterozygosity for chromosomal fusions/fissions leads to the very specific chromosomal structures at the MII stage, when heterozygotes retain a configuration similar to that of trivalents. Such trivalent-like structures were observed at the MII stage in the sample U405 (shown in green in Fig. 5d). The number of such trivalent-like structures reached 7, while the total number of chromosome entities was n = 29. If these elements are interpreted as trivalents, then the diploid number can be estimated as 2n = 65. If these elements are bivalents, then the diploid chromosome number is 2n = 58. Thus, the preliminary haploid number of chromosomes can be estimated as n = 29–33.

Previously, a chromosome cline was found in *L. sinapis*, within which the diploid chromosome number gradually decreases from 2n = 106 in Spain to 2n = 56 in Sweden and in eastern Kazakhstan (Lukhtanov et al. 2011, 2018). Thus, the studied population from Mazandaran, Iran has an oriental variant of karyotype, that is, with a relatively low number of chromosomes. We were not able to study the karyotype from the Iranian Talesh; however, the karyotype of the population from Yardimli in Republic of Azerbaijan's Talysh region was studied previously (Lukhtanov 1992). The latter population (Azerbaijani Talysh) demonstrated variation in the haploid chromosome number from n = 28 to n = 34 (Lukhtanov 1992), thus, similar to the Mazandaran population.

Distribution and ecology. So far, the presence of *L. s. tabarestana* ssp. nov. has been confirmed by DNA evidence only in northern Iran, in provinces of Ardabil, Gilan, Mazandaran, Tehran and Golestan (Fig. 6). Specimens from the Talysh mountains in Republic of Azerbaijan, across the border from Iranian Talesh region, show the same karyotype and possibly belong to ssp. *tabarestana*, however this remains to be further confirmed by DNA sequencing. In Turkmenistan, even though reports of *L. sinapis* from the Kopet Dagh mountains are as of yet unconfirmed (Tshikolovets 1998), these also likely belong to ssp. *tabarestana*.
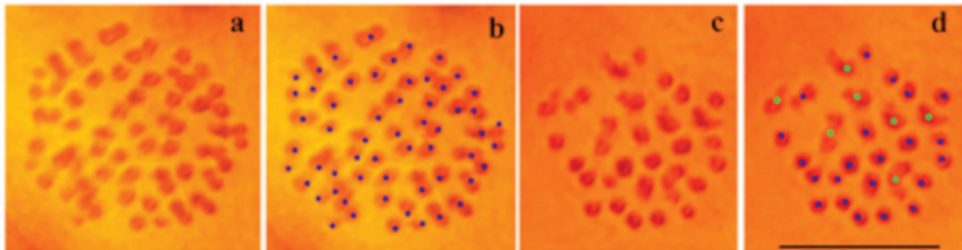
Figure 5. Karyotype of *Leptidea sinapis tabarestana* ssp. nov. **a, b** mitotic cell demonstrating ca 58 chromosomes (sample VLU396) **c, d** MII plate demonstrating 29 entities, 22 entities were interpreted as bivalents (shown by blue dots on Fig. 5d) and 7 entities were interpreted as trivalents (shown by green dots on Fig. 4d). Scale bar: 10 μm.
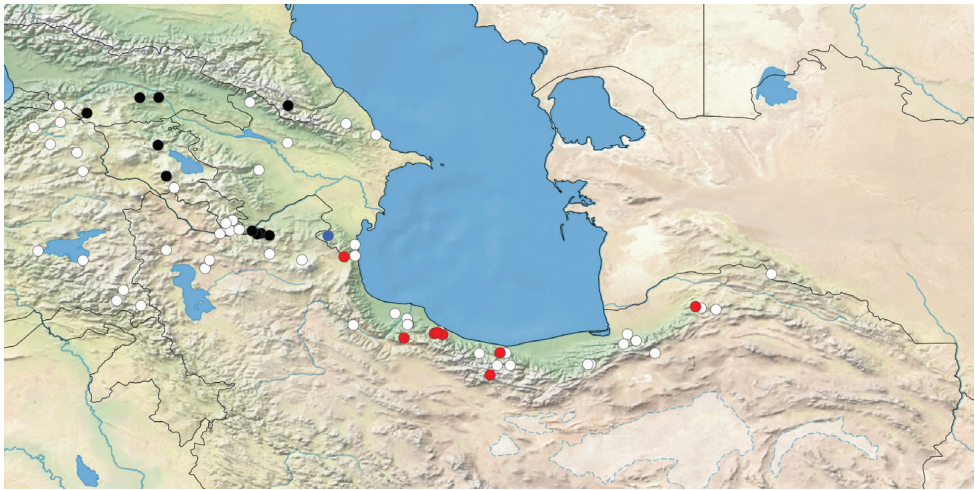
Figure 6. Distribution of *Leptidea sinapis* in E Turkey, S Caucasus and N Iran. Black dots: barcoded *L. s. sinapis*; red dots: barcoded *L. s. tabarestana* ssp. nov.; blue dot: karyotyped sample from Yardamli in Republic of Azerbaijan's Talysh region (most likely *L. s. tabarestana* ssp. nov.); white dots: non-barcoded material, data obtained from literature or personal collections.

In the Iranian Talesh mountains, *L. s. tabarestana* ssp. nov. occurs approximately 100 km from the closest population of the nominotypical *L. sinapis* in Arasbaran region. The habitat of ssp. *tabarestana* is in the Alborz forest belt, in humid meadows, forest river banks, forest clearings, and sometimes gardens at mountain steppes from 1000 to 2000 m above sea level. Adults fly mostly in undisturbed or lightly-grazed habitats with lush of green vegetation (Fig. 7). The accompanying species include *Ochlodes hyrcana* (Christoph, 1893), *Pieris napi mazandarana* Eitschberger, 1987, *Lasiommata adrastioides* (Bienert, [1870]), and *Maniola jurtina* (Linnaeus, 1758). It is normally found in two (or maybe three) generations, from April at lower altitudes to the end of September at higher altitudes. The early stages of *L. s. tabarestana* ssp. nov. are unknown, however adults are often seen near *Lathyrus* plants (AN, personal observation). Even though the larval host plant is likely among the herbaceous Fabaceae of the genera *Lathyrus*, *Vicia*, *Lotus* etc., it is as of yet unrecorded and thus it is unclear if ssp. *tabarestana* displays any host plant preferences different from the rest of populations of ssp. *sinapis*.

Discussion

Reissinger (1989) recognized twelve subspecies of *L. sinapis* across its range, including *reali* and *juvernica*, both of which were later confirmed as separate species (Dincă et al. 2011, 2013). Since then, this complex has taken a central stage in efforts to understand the mechanisms of cryptic speciation in butterflies, and thus the idea of the existence of subspecies within *L. sinapis* seems to have slowly faded away. Modern taxonomic treatments of the group (e.g., Bozano et al. 2016) regard all populations of *L. sinapis*
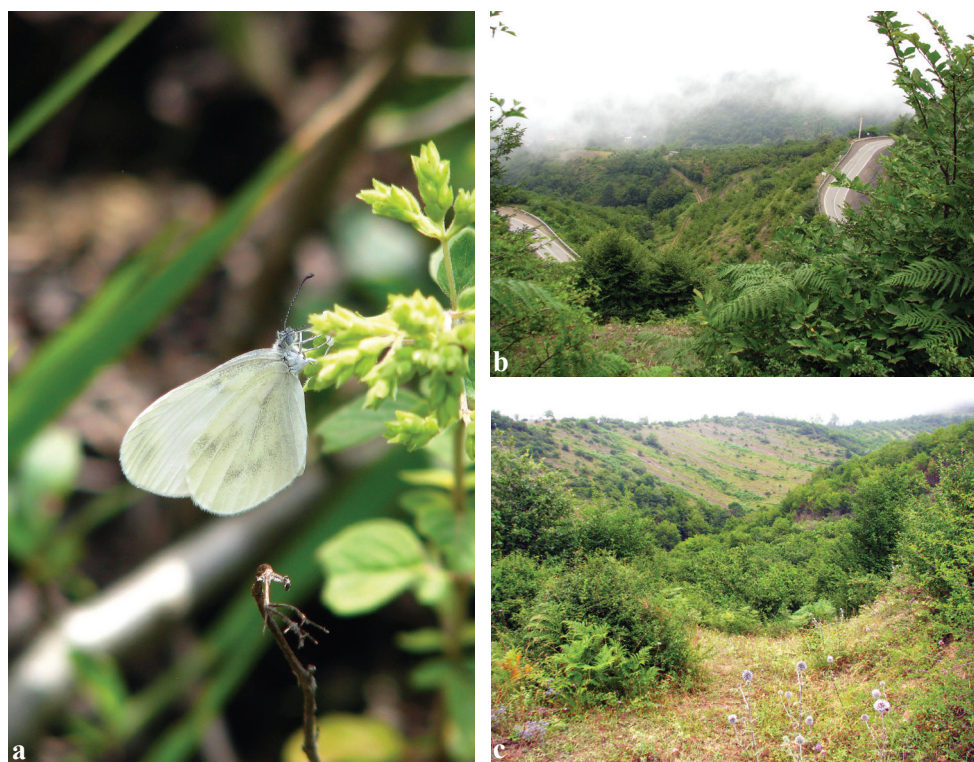
Figure 7. *Leptidea sinapis tabarestana* ssp. nov. **a** adult **b, c** habitat in Iran, Mazandaran Prov., Javaherdeh (Jirkooh), 24.VII.2011. Photos: V. Lukhtanov.

from Europe to NW Mongolia as a single entity, corresponding to the nominotypical subspecies. It occasionally flies in sympatry with closely related and extremely similar species of *Leptidea* across its range and can be separated from them only by DNA sequencing and analysis of karyotype or genitalia.

In a similar vein, in this study we found no single external morphological character or combination of characters that could reliably separate *L. s. tabarestana* ssp. nov. from the nominotypical *L. sinapis*. Individual variation in morphology observed within *L. s. tabarestana* ssp. nov. is not unexpected, as similar variation can also be seen in *L. s. sinapis*, as well as other species within the genus. In the Arasbaran mountains in NW Iran, where the nominotypical *L. sinapis* is found, individuals flying in colder slopes at high altitudes (1700–2000 m) tend to be smaller and darker, while those found in warmer forests at lower altitudes (1200–1400 m) are usually larger in size and have a lighter complexion.

Recent studies have estimated the age of the most recent common ancestor (MRCA) of *L. sinapis* at 1.5 mya, and for MRCA of *sinapis+juvernica* at 3 mya (Talla et al. 2017). The subsequent dispersal of *L. sinapis* eastward however appears to have occurred much later, either before or after the Last Glacial Maximum (LGM) (24,000 to 17,000 years ago) (Lukhtanov et al. 2011). During the Pleistocene, dense forests covered the entire northern Iran, from the northwest (Azerbaijan province) across the Alborz mountains

and extending further into the northeast (Kopet Dagh); However, since the Last Glacial Maximum (LGM; 21 kya), the Alborz mountain range has been nearly entirely isolated from all other regions surrounding it. Subsequent decline in forest cover resulted in isolated refugia in parts of southern Caucasus as well as in northern Iran (Yousefi et al. 2015; Asadi et al. 2018; Parvizi et al. 2018; Liu et al. 2019; Saberi-Pirooz et al. 2020). With the likely extinction of intervening populations, the range of many butterflies adapted to this habitat – including the ancestral *L. sinapis* – became fragmented, resulting in the geographic and genetic isolation of *L. s. tabarestana* ssp. nov.

Presence of *Wolbachia* endosymbionts affecting mtDNA in *Leptidea* has been noted previously (e.g. Solovyev et al. 2015) and we cannot rule out that this may have had an effect on our results. Further studies are needed to confirm the presence of *L. s. tabarestana* ssp. nov. in the Republic of Azerbaijan and in Turkmenistan. Potential sympatric occurrence of the two entities in the intervening areas in NW Iran needs to be investigated. If the two are found to co-occur sympatrically and synchronically without geneflow, or other new information (e.g., karyotype, nuDNA, morphology etc.) comes to light that clearly signals the two taxa to be distinct at species level, the taxon *tabarestana* may be raised as bona species.

Acknowledgements

References

Asadi A, Montgelard C, Nazarizadeh M, Moghaddasi A, Fatemizadeh F, Simonov E, Kami HG, Kaboli M (2018) Evolutionary history and postglacial colonization of an Asian pit viper (*Gloydius halys caucasicus*) into transcaucasia revealed by phylogenetic and phylogeographic analyses. Scientific Reports 9: e1224. https://doi.org/10.1038/s41598-018-37558-8

Bolshakov LV (2006) New subspecies of *Leptidea reali* Reissinger, 1989 (Lepidoptera: Pieridae) from the mountains regions of Middle Asia. Eversmannia 5: 6–10.

Bozano GC, Coutsis JG, Heřman P, Allegrucci G, Cesaroni D, Sbordoni V (2016) Guide to the Butterflies of the Palearctic Region. Pieridae Part III: Subfamily Coliadinae, Tribes Rhodocerini, Euremini, Coliadini, Genus *Catopsilia* & Subfamily Dismorphiinae. Omnes Artes, 70 pp.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657–1660. https://doi.org/10.1046/j.1365-294x.2000.01020.x

Dincă V, Lukhtanov VA, Talavera G, Vila R (2011) unexpected layers of cryptic diversity in wood white *Leptidea* butterflies. Nature Communications 2: e324. https://doi.org/10.1038/ncomms1329.

Dincă V, Wiklund C, Lukhtanov VA, Kodandaramaiah U, Noren K, Dapporto L, Wahlberg N, Vila R, Friberg M (2013) Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex. Journal of Evolutionary Biology 26: 2095–2106. https://doi.org/10.1111/jeb.12211

Dincă V, Dapporto L, Somervuo P, Vodă R, Cuvelier S, Gascoigne-Pees M, Huemer P, Mutanen M, Hebert PDN, Vila R (2021) High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity. Communications Biology 4: e315. https://doi.org/10.1038/s42003-021-01834-7

Friberg M, Bergman M, Kullberg J, Wahlberg N, Wiklund C (2008) Niche separation in space and time between two sympatric sister species–a case of ecological pleiotropy. Evolutionary Ecology 22: 1–18. https://doi.org/10.1007/s10682-007-9155-y

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 52: 696–704. https://doi.org/10.1080/10635150390235520

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006) DNA barcodes distinguish species of tropical Lepidoptera. Proceedings of the National Academy of Sciences of the United States of America 103: 968–971. https://doi.org/10.1073/pnas.0510466103

Inkscape Project (2020) Inkscape. https://inkscape.org [Accessed March 2023]

Leal L, Talla V, Källman T, Friberg M, Wiklund C, Dincă V, Vila R, Backström N (2018) Gene expression profiling across ontogenetic stages in the wood white (*Leptidea sinapis*) reveals pathways linked to butterfly diapause regulation. Molecular Ecology 27: 935–948. https://doi.org/10.1111/mec.14501

Liu J, Guo X, Chen D, Li J, Yue B, Zeng X (2019) Diversification and historical demography of the rapid racerunner (*Eremias velox*) in relation to geological history and Pleistocene climatic oscillations in arid Central Asia. Molecular Phylogenetics and Evolution 130: 244–258. https://doi.org/10.1016/j.ympev.2018.10.029

Lorković Z (1993) *Leptidea reali* Reissinger, 1989 (=*lorkovicii* Real 1988), a new European species (Lepid., Pieridae). Natura Croatica 2: 1–26.

Lukhtanov VA (1992) Karyotype evolution and systematics of higher taxa of Pieridae (Lepidoptera) of the world. Entomological Review 71(5): 57–82.

Lukhtanov VA, Dincă V, Talavera G, Vila R (2011) Unprecedented within-species chromosome number cline in the Wood White butterfly *Leptidea sinapis* and its significance for karyotype evolution and speciation. BMC Evolutionary Biology 11: e109. https://doi.org/10.1186/1471-2148-11-109

Lukhtanov VA, Dincă V, Friberg M, Šíchová J, Olofsson M, Vila R, Marec F, Wiklund C (2018) Versatility of multivalent orientation, inverted meiosis, and rescued fitness in holocentric chromosomal hybrids. Proceedings of the National Academy of Sciences of the United States of America 115(41): E9610–E9619. https://doi.org/10.1073/pnas.1802610115

Lukhtanov VA, Dantchenko AV, Khakimov FR, Sharafutdinov D, Pazhenkova EA (2020a) Karyotype evolution and flexible (conventional versus inverted) meiosis in insects with holocentric chromosomes: a case study based on *Polyommatus* butterflies. Biological Journal of the Linnean Society 130(4): 683–699. https://doi.org/10.1093/biolinnean/blaa077

Lukhtanov VA, Dincă V, Friberg M, Vila R, Wiklund C (2020b) Incomplete sterility of chromosomal hybrids: implications for karyotype evolution and homoploid hybrid speciation. Frontiers in Genetics 11: e583827. https://doi.org/10.3389/fgene.2020.583827

Mazel R (2005) Éléments de phylogénie dans le genre *Leptidea* Bilberg, 1820. Revue de L'association Roussillonnaise d'entomologie 14(3): 98–111.

Martin JF, Gilles A, Descimon H (2003) Species concepts and sibling species: the case of *Leptidea sinapis* and *Leptidea reali*. In: Boggs CL, Watt WB, Ehrlich P (Eds) Butterflies: Ecology and Evolution Taking Flight. University of Chicago Press, Chicago, 459–476.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Molecular Biology and Evolution 32: 268–274. https://doi.org/10.1093/molbev/msu300

Parvizi E, Naderloo R, Keikhosravi A, Solhjouy-Fard S, Schubart CD (2018) Multiple Pleistocene refugia and repeated phylogeographic breaks in the southern Caspian Sea region: Insights from the freshwater crab *Potamon ibericum*. Journal of Biogeography 2018: 1–12. https://doi.org/10.1111/jbi.13195

Platania L, Menchetti M, Dincă V, Corbella C, Kay-Lavelle I, Vila R, Wiemers M, Schweiger O, Dapporto L (2020) Assigning occurrence data to cryptic taxa improves climatic niche assessments: Biodecrypt, a new tool tested on European butterflies. Global Ecology and Biogeography 29: 1852–1865. https://doi.org/10.1111/geb.13154

Réal P (1988) Lépidoptères nouveaux principalement jurassiens. Mémoires du Comité de liaison pour les recherches écofaunistiques dans le Jura 4: 1–28.

Reissinger EJ (1989) Checkliste Pieridae Duponchel, 1835 (Lepidoptera) der Westpalaearktis (Europa, Nordwestafrika, Kaukasus, Kleinasien). Atalanta 20: 149–185.

Rambault A (2018) Figtree v1. 4.4. Computer program and documentation distributed by the author. http://github.com/rambaut/figtree [Accessed March 2023]

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67: 901–904. https://doi.org/10.1093/sysbio/ syy032

Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61: 539–542. https://doi.org/10.1093/sysbio/sys029

Saberi-Pirooz R, Rajabi-Maham H, Ahmadzadeh F, Kiabi BH, Javidkar M, Carretero MA (2020) Pleistocene climate fluctuations as the major driver of genetic diversity and distribution patterns of the Caspian green lizard, *Lacerta strigata* Eichwald, 1831. Ecology and Evolution 11: 6927–6940. https://doi.org/10.1002/ece3.7543

Shorthouse DP (2010) SimpleMappr, an online tool to produce publication-quality point maps. https://www.simplemappr.net [Accessed March 2023]

Shtinkov N, Kolev Z, Vila R, Dincă V (2016) The sibling species *Leptidea juvernica* and *L. sinapis* (Lepidoptera, Pieridae) in the Balkan Peninsula: ecology, genetic structure, and morphological variation. Zoology 119: 11–20. https://doi.org/10.1016/j.zool.2015.12.003

Šíchová J, Voleníková A, Dincă V, Nguyen P, Vila R, Sahara K, Marec F (2015) Dynamic karyotype evolution and unique sex determination systems in *Leptidea* wood white butterflies. BMC Evolutionary Biology (2015) 15: 89. https://doi.org/10.1186/s12862-015-0375-4

Šíchová J, Ohno M, Dinca V, Watanabe M, Sahara K, Marec F (2016) Fissions, fusions, and translocations shaped the karyotype and multiple sex chromosome constitution of the northeast-Asian wood white butterfly, *Leptidea amurensis*. Biological Journal of the Linnean Society 118: 457–471. https://doi.org/10.1111/bij.12756

Solovyev VI, Ilinsky Y, Kosterin OE (2015) Genetic integrity of four species of *Leptidea* (Pieridae, Lepidoptera) as sampled in sympatry in West Siberia. Comparative Cytogenetics 9(3): 299–324. https://doi.org/10.3897/CompCytogen.v9i3.4636

Talla V, Suh A, Kalsoom F, Dincă V, Vila R, Friberg M, Wiklund C, Backström N (2017) Rapid Increase in Genome Size as a Consequence of Transposable Element Hyperactivity in Wood-White (*Leptidea*) Butterflies. Genome Biology and Evolution 9(10): 2491–2505. https://doi.org/10.1093/gbe/evx163

Talla V, Johansson A, Dincă V, Vila R, Friberg M, Wiklund C, Backström N (2019a) Lack of gene flow: Narrow and dispersed differentiation islands in a triplet of *Leptidea* butterfly species. Molecular Ecology 28: 3756–3770. https://doi.org/10.1111/mec.15188

Talla V, Soler L, Kawakami T, Dincă V, Vila R, Friberg M, Wiklund C, Backström N (2019b) Dissecting the Effects of Selection and Mutation on Genetic Diversity in Three Wood White (*Leptidea*) Butterfly Species. Genome Biology and Evolution 11(10): 2875–2886. https://doi.org/10.1093/gbe/evz212

Tamura K, Stechrer G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis version 11. Molecular Biology and Evolution 38: 3022–3027. https://doi.org/10.1093/molbev/msab120

The GIMP Development Team (2019) GIMP. https://www.gimp.org [Accessed March 2023]

Tshikolovets VV (1998) The Butterflies of Turkmenistan. V.V. Tshikolovets publications, 237 pp.

Vershinina AO, Lukhtanov VA (2010) Geographical distribution of the cryptic species *Agrodiaetus alcestis alcestis*, *A. alcestis karacetinae* and *A. demavendi* (Lepidoptera, Lycaenidae) revealed by cytogenetic analysis. Comparative Cytogenetics 4(1): 1–11. https://doi.org/10.3897/compcytogen.v4i1.21

Vodă R, Dapporto L, Dincă V, Vila R (2015) Why do cryptic species tend not to co-occur? A case study on two cryptic pairs of butterflies. PLoS ONE 10(2): e0117802. https://doi.org/10.1371/journal. pone.0117802

Yoshido A, Šíchová J, Pospíšilová K, Nguyen P, Voleníková A, Šafář J, Provazník J, Vila R, Marec F (2020) Evolution of multiple sex-chromosomes associated with dynamic genome reshuffling in *Leptidea* wood-white butterflies. Heredity 125: 138–154. https://doi.org/10.1038/s41437-020-0325-9

Yousefi M, Ahmadi M, Nourani E, Behrooz R, Rajabizadeh M, Geniez P, Kaboli M (2015) Upward altitudinal shifts in habitat suitability of Mountain Vipers since the Last Glacial Maximum. PLoS ONE 10(9): e0138087. https://doi.org/10.1371/journal.pone.0138087

Supplementary material 1

Material examined and GenBank accessions
Authors: Vazrick Nazari, Vladimir A. Lukhtanov, Alireza Naderi, Zdenek Faltýnek Fric, Vlad Dincă, Roger Vila
Data type: excel file
Link: https://doi.org/10.3897/compcytogen.17.102830.suppl1

Supplementary material 2

Length of male genitalia components (phallus, saccus, vinculum) and ratios (PL/VW, SL/VW) among *Leptidea* specimens measured in this study
Authors: Vazrick Nazari, Vladimir A. Lukhtanov, Alireza Naderi, Zdenek Faltýnek Fric, Vlad Dincă, Roger Vila
Data type: excel file
Link: https://doi.org/10.3897/compcytogen.17.102830.suppl2

ORCID

Vazrick Nazari https://orcid.org/0000-0001-9064-8959
Vladimir A. Lukhtanov https://orcid.org/0000-0003-2856-2075
Zdenek Faltýnek Fric https://orcid.org/0000-0002-3611-8022
Vlad Dincă https://orcid.org/0000-0003-1791-2148
Roger Vila https://orcid.org/0000-0002-2447-4388

REVIEW ARTICLE

Allium cytogenetics: a critical review on the Indian taxa

Biplab Kumar Bhowmick[1], Sayantika Sarkar[2], Dipasree Roychowdhury[2],
Sayali D. Patil[3], Manoj M. Lekhak[3], Deepak Ohri[4], Satyawada Rama Rao[5],
S. R. Yadav[3], R. C. Verma[6], Manoj K. Dhar[7], S. N. Raina[8] , Sumita Jha[2]

1 *Department of Botany, Scottish Church College, 1&3, Urquhart Square, Kolkata- 700006, West Bengal, India* **2** *Department of Botany, University of Calcutta, 35 Ballygunge Circular Road, Kolkata- 700019, West Bengal, India* **3** *Angiosperm Taxonomy Laboratory, Department of Botany, Shivaji University, Kolhapur, Maharashtra- 416004, India* **4** *Amity Institute of Biotechnology, Research Cell, Amity University Uttar Pradesh, Lucknow Campus, Lucknow- 226028, Uttar Pradesh, India* **5** *Department of Biotechnology and Bioinformatics, North-Eastern Hill University, Shillong, Meghalaya- 793022, India* **6** *School of Studies in Botany, Vikram University, Ujjain, Madhya Pradesh 456010, India* **7** *Genome Research Laboratory, School of Biotechnology, University of Jammu, Jammu, Jammu and Kashmir- 180006, India* **8** *Amity Institute of Biotechnology, Amity University, Sector 125, Noida, Uttar Pradesh- 201313, India*

Corresponding author: Sumita Jha (sumitajha.cu@gmail.com)

Citation: Bhowmick BK, Sarkar S, Roychowdhury D, Patil SD, Lekhak MM, Ohri D, Rama Rao S, Yadav SR, Verma RC, Dhar MK, Raina SN, Jha S (2023) *Allium* cytogenetics: a critical review on the Indian taxa. Comparative Cytogenetics 17: 129–156. https://doi.org/10.3897/compcytogen.17.98903

Abstract

The genus *Allium* Linnaeus, 1753 (tribe Allieae) contains about 800 species worldwide of which almost 38 species are reported in India, including the globally important crops (onion, garlic, leek, shallot) and many wild species. A satisfactory chromosomal catalogue of *Allium* species is missing which has been considered in the review for the species occurring in India. The most prominent base number is x=8, with few records of x=7, 10, 11. The genome size has sufficient clues for divergence, ranging from 7.8 pg/1C to 30.0 pg/1C in diploid and 15.16 pg/1C to 41.78 pg/1C in polyploid species. Although the karyotypes are seemingly dominated by metacentrics, substantial variation in nucleolus organizing regions (NORs) is noteworthy. The chromosomal rearrangement between *A. cepa* Linnaeus, 1753 and its allied species has paved way to appreciate genomic evolution within *Allium*. The presence of a unique telomere sequence and its conservation in *Allium* sets this genus apart from all other Amaryllids and supports monophyletic origin. Any cytogenetic investigation regarding NOR variability, telomere sequence and genome size in the Indian species becomes the most promising field to decipher chromosome evolution against the background of species diversity and evolution, especially in the Indian subcontinent.

Introduction

The genus *Allium* Linnaeus, 1753 is considered a wonder crop of global importance, catering to the agriculture, condiment, pharmaceutical, nutraceutical and cosmetic sectors of economy owing to the presence of numerous species with tremendous significance. Among several herb species, an onion (*A. cepa* Linnaeus, 1753) that is valued throughout the continent attracts a lot of attention of the economic sectors mentioned above, followed by garlics, leeks and shallots having limited uses. Onion is the second of the five main world vegetables species (after tomato) whose worldwide production accounted for 9% of the total (42–45%) increase in production of vegetables between 2000–2019 (https://www.fao.org/3/cb4477en/online/cb4477en.html#chapter-2_1).

Allium, previously referred to Liliaceae, is now a member of Amaryllidaceae sensu Angiosperm Phylogeny Group or APG III (Haston et al. 2009). This large genus (about 800 species, Costa et al. 2020) was divided into 15 subgenera and 56 sections (Friesen et al. 2006). At present, *Allium* has its primary evolution centre across the Irano-Turanian phytochorion while secondary centres of diversity include Mediterranean basin and western North America (Friesen et al. 2006). The taxonomy and evolution of this diverse genus has been accepted as difficult.

Cytogenetics, being the only elementary discipline of genetics, focuses on genome structure, function and evolution. The evolutionary history of organisms is inscribed in the chromosomes, the physically visible form of genome. The very fundamental parameters such as chromosome count reports, when combined with molecular cytogenetic and phylogenetic data (Islam-Faridi et al. 2020; Senderowicz et al. 2021), or genome size estimates, can elucidate trends of evolution in context of ploidy changes. Molecular cytogenetic approaches, in line with the parameters mentioned already, can accelerate the understanding of the evolutionary questions (Borowska-Zuchowska et al. 2022; Nath et al. 2022). A general correlation between evolutionary trends and chromosomal features has been shown in many plant families (Van-Lume et al. 2017; Carta et al. 2020; Bhowmick and Jha 2022; Nath et al. 2022). Recently, a broad concurrence between karyology and geographical distribution has been shown in three Allioidae tribes, with respect to the diversification of Allieae to Northern Hemisphere from the Indian tectonic plate around 30 million years ago (Costa et al. 2020).

India is the world's second-largest producer of onion after China, with a production rate of 16360 kg/ Ha (2020–2021) (https://eands.dacnet.nic.in/). After onion, *A. sativum* Linnaeus, 1753 (garlic) is the second largest species of *Allium* contributing significantly to agro-economical development of the country (https://eands.dacnet.nic.in/). Among the other species, *A. schoenoprasum* Linnaeus, 1753 and *A. roylei*

Stearn, 1947 exhibited resistance qualities (Nanda et al. 2016) and promise adoption of advanced breeding. Keeping in mind the significance of *Allium* and the complications in taxonomy and evolution, a comprehensive summary of cytogenetic characters has been presented for Indian species of *Allium*.

Data compilation

Distribution of taxa, chromosome counts, ploidy, karyotypes and molecular cytogenetic reports have been compiled from original publications, chromosome atlases and databases e.g. Database on Genome-Related Information of Indian Plants or d-GRIP (http://indianpcd.com/; Jha et al. 2019), Index to Plant Chromosome Numbers or IPCN (http://www.tropicos.org/project/ipcn, Goldblatt and Lowry 2011), Chromosome Counts Database or CCDB (http://ccdb.tau.ac.il/, Rice et al. 2015), The Plant DNA C-values database (https://cvalues.science.kew.org/, Pellicer and Leitch 2020) and Plant rDNA Database (www.plantrdnadatabase.com, Vitales et al. 2017). In case of synonyms, the present taxonomic designations are retained with appropriate references.

Cytogenetic catalogue of *Allium* species in India

There are 35–40 species of *Allium* currently reported from India (ca. 38 species) (d-GRIP, Pandey et al. 2021, 2022). The species of *Allium* in India belong to nine subgenera namely, *Cepa* (5 species), *Allium* (5 species), *Amerallium* (4 species), *Reticulatobulbosa* (3 species), *Polyprason* (3 species), *Anguinum* (2 species), *Butomissa* (2 species), *Melanocrommyum* (1 species) and *Rhizirideum* (2 species) (Friesen et al. 2006). Majority of the *Allium* species prefer temperate mixed forests or rocky slopes ranging 1200–5480 meters of the western Himalayas (e.g. *A. atropurpureum* Waldst. et Kit., 1800, *A. atrosanguineum* Schrenk, 1842, *A. auriculatum* Kunth, 1843, *A. caesioides* Wendelbo, 1969, *A. carolinianum* Redouté, 1804, *A. consanguineum* Kunth, 1843, *A. fedtschenkoanum* Regel, 1875, *A. griffithianum* Boiss., 1859, *A. loratum* Baker, 1874, *A. oreoprasum* Schrenk, 1842, *A. roylei*, *A. schoenoprasum* and *A. schrenkii* Regel, 1875). There are few species endemic to Kashmir and Uttarakhand (e.g. *A. gilgiticum* F.T. Wang et Tang, 1937 which is also endangered, *A. stracheyi* Baker, 1874 and *A. negianum* A. Pandey, K.M. Rai, Malav et S. Rajkumar, 2021) (Pandey et al. 2021). Rest of the species occupy the temperate habitats of north-eastern hill region (e.g. *A. fasciculatum* Rendle, 1906, *A. hookeri* Thwaites, 1864, *A. macranthum* Baker, 1874, *A. platyspathum* Schrenk, 1841, *A. prattii* C.H. Wright, 1903, *A. rhabdotum* Stearn, 1960, *A. sikkimense* Baker, 1874) while some wild or semi-wild species (*A. przewalskianum* Regel, 1875, *A. tuberosum* Rottler et Sprengel, 1825, *A. victorialis* Linnaeus, 1753, *A. wallichii* Kunth, 1843) occur in the western and eastern Himalayan regions.

Chromosome counts

The chromosome counts and karyotype details are known perhaps in 33 and 25 species, respectively (Table 1, Fig. 1). The prominent base number (x) is 8, irrespective of the subgenera, sections or the distribution pattern. Some western Himalayan species which are still not assigned to any of the subgenera (e.g. *A. atropurpureum*, *A. caesioides*, *A. consanguineum*, *A. ascalonicum* Linnaeus, 1756, *A. blandum* Wall., 1832, *A. hypsistum* Stearn, 1960) and endemic *A. stracheyi* have x=8. Divergent numbers such as x=7, 10 and 11 are found in the Indian species of the subgenus *Amerallium* (Table 1) which also justify their inclusion in a separate subgenus (Peruzzi et al. 2017). Chromosome number has not been studied in the newly discovered *A. negianum* of *Rhizirideum*, sect. *Eduardia* (Pandey et al. 2021), which together with its close relative *A. przewalskianum* of sect. *Caespitosoprason* (Pandey et al. 2021) not studied from the territory of India, needs to be investigated. Similarly, *A. loratum*, *A. auriculatum*, *A. rhabdotum* and an endemic *A. gilgiticum* still are not assigned to any of the subgenera, and any cytological information is also missing. The meiotic studies in some species have shown various configurations like multivalents or univalents and occasional irregularities as in *A. chinense* G. Don, 1827 (Gohil and Koul 1973, 1981), *A. hookeri* (Sharma et al. 2011), *A. roylei* (Sharma and Gohil 2003, 2011a; Kohli and Gohil 2011), *A. rubellum* M. Bieb., 1808 (Khoshoo and Sharma 1959; Koul et al. 1971) and *A. tuberosum* (Gohil and Koul 1983; Sharma and Gohil 2004, 2013a, b). In case of tetraploid *A. ampeloprasum* Linnaeus, 1753 (as *A. porrum* Linnaeus, 1753 in many studies), 16 bivalents were recorded regularly with complete absence of any multivalent (Koul and Gohil 1970b; Ved Brat and Dhingra 1973; Gohil and Koul 1977; Pandita and Mehra 1981a; Stack and Roelofs 1996). In this species, some peculiar features like appearance of bivalents in metaphase I instead of quadrivalents, localized chiasmata at pericentromeric regions have been reported (Levan 1940; Koul and Gohil 1970b; Stack and Roelofs 1996). Considering the incidence of vivipary and hybridization in *A. cepa* (Singh et al. 1967; Langer and Koul 1983; Puizina and Papea 1996), thorough meiotic analysis of the agriculturally important species (*A. cepa*, *A. sativum*, etc.) would be a significant aspect of future revision.

Ploidy and genome size

The greatest variation in ploidy has been observed in *A. tuberosum* (subgenus *Butomissa*), *A. przewalskianum* (subgenus *Rhizirideum*), *A. chinense* G. Don, 1827 (subgenus *Cepa*) and *A. rubellum*, *A. ampeloprasum*, *A. griffithianum* (subgenus *Allium*) (Table 1). Polyploidy is reported in almost all subgenera and species. However, Peruzzi et al. (2017) reported absence of polyploidy in subgenus *Anguinum* and emphasized on correlation between chromosome size and ploidy to infer the trend of evolution. Any such correlation for Indian taxa is not possible at this stage due to lack of data for all the species.

Table 1. Chromosome numbers, ploidy and nuclear genome sizes in Indian species of *Allium* of Amaryllidaceae (Tribe Allieae, Subfamily Allioideae, sensu APG IV 2016).

Subgenus/ section	Species (syn.)	Chromosome number			Ploidy	4C DNA value in diploid/ polyploid nuclei (pg)	Genome size in diploid/ polyploid (pg)		References
		Basic (x)	Gametic (n)	Zygotic (2n)			1C	1Cx	
Amerallium/ Bromatorrhizad	**A. fasciculatum Rendle** (*A. gageanum*)	10[a]	–	20[b], 40[c]	Diploid[d], Tetraploid[e]	–	–	–	a,b,d,(Xu et al. 1998; Li et al. 2017), b(Huang et al. 1995), c-e(Dutta et al. 2015)
Amerallium/ Bromatorrhizad*	**A. bookeri Thwaites** (*A. toongi*)	–	–	22[a], 33[b]–44[c]	–	63.24 (diploid, Feulgen cytophotometry)[d]	15.81 (diploid)[d]	15.81[d]	a(Sen 1974a; Tang et al. 2005; Sharma et al. 2011), a,b,c(Huang et al. 1995), a-c(Phuong et al. 2010), a,d(Ohri et al. 1998; Ohri and Pistrick 2001)
Amerallium/ Bromatorrhizad	**A. macranthum Baker** (*A. oviflorum* Regel, *A. simethis* H.Lev.)	–	14[a]	14[b], 28[c]	–	–	–	–	a(Levan 1934), b-c(Huang et al. 1995; Tang et al. 2005)
Amerallium/ Bromatorrhizad*†	**A. wallichii Kunth.** (*A. bulleyanum* Diels, *A. caeruleum* Wall.)	7[a]	-	14[b], 28[c], 32[d]	Diploid[e], Tetraploid[f]	64.98 (diploid, Feulgen Cytophotometry)[g], 121.79 (tetraploid, Feulgen Cyrophotometry)[h], 119.13 (tetraploid, Feulgen microdensitometry)[i]	16.24 (diploid)[g], 30.45 (tetraploid)[h]	16.24[g], 15.22[h]	a,b,c,e,f,(Huang et al. 1995), c-f,(Labani and Elkington 1987), d(Ved Brat 1965), a,b,c,e,f,g,h(Ohri et al. 1998), a,b,c,e,g,i(Ohri and Pistrick 2001)
Anguinum/ Anguinum*	**A. prattii C.H.Wright** (*A. camiflolum* H. Lev., *A. ellipticum* Wall et Kunth)	8[a]	16[b]	16[c], 32[d]	Diploid[e], Tetraploid[f]	–	–	–	a-c(Tang et al. 2005), b(Kurosawa 1966), c,d,e(Chunying et al. 2000)
Anguinum/ Anguinum	**A. victorialis L.** (*A. anguinum* Bubani, *A. reticulatum* St.-Lag.)	8[a]	8[b]	16, 32[c], 36[e]	Diploid[f], Tetraploid[g]	81.00 (diploid)[h], 86.42 (diploid, Feulgen microdensitometry)[i], 162.02 to 167.10 (Tetraploid, Feulgen cytophotometry)[j]	20.25 (diploid)[h], 21.60 (diploid), 40.5–41.78(tetraploid)[j]	20.25[h], 21.60, 20.25–20.89	a, h(Pandita and Mehra 1981a), c-j(Pandita and Mehra 1981b), a-c(Mehra and Sachdeva 1976; Lu et al. 2017), c-f,(Labani and Elkington 1987), d,g(Ohri et al. 1998), d,g,j(Ohri and Pistrick 2001), e(Sen 1973a), h(Vakhtina et al. 1977)
Melanocrommyum/ Breviacule #	**Allium chitralicum** Wang & Tang (*A. badakhshanicum*, *A. pauli*)	–	–	16[a], 32[b]	–	–	34.35 (tetraploid, flow cytometry)[a*]	17.17 (tetraploid, flow cytometry)[a*]	a(Pedersen and Wendelbo 1966), b,c,a*(Gurushidze et al. 2012)
Butomissal/ Butomissa*	**A. tuberosum Rottler ex Spreng.** (*A. chinense* Maxim., *A. clarkei* Hook.f.)	8[a]	8[b], 16[c], 32[d]	16[e], 32[f], 24[g], 31, 33[h], 48[i], 61–64[j], 62[k], 64[l]	Tetraploid[m], Hexaploid[n], Octaploid[o], Autotetraploid[p], Autopolyploid[q]	66.80 (tetraploid)[r], 121 (tetraploid)[s], 109.36 (tetraploid, Feulgen cyrophotometry)[t], 121.47–123.25 (tetraploid, Feulgen Cytophotometry)[u]	30.36–30.62 (tetraploid)[u]	15.18–15.31[u]	a,c,f,m,r(Pandita and Mehra 1981a), a,f,m(Talukder and Sen 2000; Kumar and Thonger 2018), b(Li et al. 1985), c-f(Sharma and Gohil 2004), c-p(Ohri et al. 1998; Ohri and Gohil 2001), f,m,u(Ohri 1963), d(Koul 1963), s(Yang et al. 1998), s(Sharma and Gohil 2013b; f,i,n,q(Sharma and Gohil 2013a), g(Huang et al. 1985), h(Gohil and Koul 1973; Gohil and Kaul 1981), i(Gohil and Kaul 1979; Ohri 1990), k(Seo 1977), l(Kojima et al. 1991), o(Gohil and Kaul 1979), n(Dutta and Bandyopadhyay 2014), o(Nanushyan and Polyakov 1989), j(Walters 1992), t(Talukder and Sen 1999)

Subgenus/ section	Species (syn.)	Chromosome number			Ploidy	4C DNA value in diploid/ polyploid nuclei (pg)	Genome size in diploid/ polyploid (pg)		References
		Basic (x)	Gametic (n)	Zygotic (2n)			1C	1Cx	
Butomissa/ Austromontana[*][1]	**A. oreoprasum Schrenk**	–	–	16[a], 48[b]	–	–	–	–	[a](Gohil and Koul 1973; Gohil and Kaul 1981), [b](Ved Brat 1965)
Rhiziridium/ Caespitosoprason[*]	**A. przewalskianum Regel** (*A. jacquemontii* var. *parviflorum* (Ledeb.) Aswal, *A. junceum* Jacquem. et Baker)	8[a]	–	16[b], 32[c], 64[d]	Diploid[e], Tetraploid[f], Octaploid[d], Autopolyploid[b]	–	–	–	[a, b, c](Tang et al. 2005), [c](Gohil and Kaul 1981), [d, g](Xue et al. 2000), [h](Ao 2008)
Allium/ Allium[*]	**A. ampeloprasum L.** (*A. adscendens*, *A. porrum* var *ampeloprasum*)	8[a]	–	16[b], 24[c], 32[d], 40[e], 56[f]	Diploid[g], polyploid[h/i] autotetraploid[i]	48.20 (tetraploid, feulgen cytophotometry), 100.54 (cytometry)[k], 119.64/ 121.15 (tetraploid, feulgen cytophotometry)[l], 119.80 (tetraploid, feulgen cytophotometry)[m]	16.7 (diploid, flow cytometry)[n*], 25.35–27.45 (tetraploid, flow cytometry)[m,n*]	16.7 (diploid, flow cytometry)[n*], 12.67–13.73 (tetraploid, flow cytometry)[m,n*]	[a, h, d, g, h, i, n*](Ricroch et al. 2005), [a, d, h.](Pandita and Mehra 1981a), [c, c](IPCN), [d, h.](Maragheh et al. 2019), [d, h, i, k](Arumuganathan and Earle 1991), [d, h, i,](Labani and Elkington 1987), [d, h, m](Ohri et al. 1998; Ohri and Pistrick 2001), [f, h](von Bothmer 1975), [i](Ranjekar et al. 1978)
Allium/ Allium[*]	**A. sativum L.** (*A. arenarium* Sadler et Rchb., *A. controversum* Schrad. et Willd.)	8[a]	8[b]	16[c], 12[d]	Diploid[e]	63.00 (diploid)[i], 64.90 (diploid, Feulgen Cytophotometry)[g], 65.40 (diploid)[h], 66.40–69.00 (diploid)[h], 68.20, 71.40[m], 73.59–91.80 (diploid, Feulgen Cyrophotometry)[i], 120 (diploid, Feulgen Cyrophotometry)[k]	15.75 (diploid)[i], 16.23 (diploid)[g], 16.35 (diploid)[h], 16.6–17.25 (diploid)[h], 16.6–17.25 (diploid)[i], 17.05, 17.85[m], 18.40–22.95 (diploid)[k], 30.0 (diploid)[k]	15.75[i], 16.23[i], 16.35[h], 16.6–17.25[i], 17.05, 17.85[m], 18.40–22.95[k], 30.0[k]	[a, c, c](Kumar and Thonger 2018), [a, b](Koul and Gohil 1970a), [c, e](Bacelar et al. 2021), [c, e, g](Ohri et al. 1998; Ohri and Pistrick 2001), [d](Sato and Kawamura 1981), [f](Ranjekar et al. 1978), [i](Murin 1976), [i](Chakravarty and Sen 1992), (Talukder and Sen 1999), l (Walters 1992), m(Olszewska and Osiecka 1982)
Allium/ Avulsea[*]	**A. griffithianum Boiss.** (*A. babri*, *A.jacquemontii* var. *grandiflorum*)	8[a]	16[b]	16[c], 32[d]	Diploid[e], Tetraploid[f], Autotetraploid[g]	41.15 (diploid, Feulgen cytophotometry)[h]	10.29 (diploid)[h]	10.29[h]	[a, b, d,](Pandita and Mehra 1981a), [c, e,](Ohri et al. 1998; Ohri and Pistrick 2001), [f, g](Pandita and Mehra 1981b)
Allium/ Avulsea[*]	**A. rubellum M. Bieb.** (*A. albanum* Grossh., *A. leptophyllum* Wall.)	–	16[a]	16[b], 24[c],	Diploid[d], Triploid[c], Tetraploid[f], Numerical hybrid[g], Autopolyploid[b]	–	–	–	[a, f](Koul et al. 1971), [b, d](Abdali and Miri 2020), [c](Gohil and Koul 1973), [c, h](Khoshoo and Sharma 1959), [g](Ved Brat 1967)
Allium/ Caerulea[*]	**A. jacquemontii Kunth**	8[a]	8[b]	16[c]	Diploid[d]	–	–	–	[a, b](Pandita and Mehra 1981a), [c](Gohil and Kaul 1981), [c, d](Pandita and Mehra 1981b)
Reticulatobulbosa/ Reticulatobulbosa[*]	**A. humile Kunth** (*A. govanianum*, *A. nivale*)	8[a]	8[b]	–	Diploid[c]	–	–	–	[a, b, c](Pandita and Mehra 1981a), [b](Mehra and Sachdeva 1975), [c](Pandita and Mehra 1981b)

| Subgenus/ section	Species (syn.)	Chromosome number			Ploidy	4C DNA value in diploid/ polyploid nuclei (pg)	Genome size in diploid/ polyploid (pg)		References
		Basic (x)	Gametic (n)	Zygotic (2n)			1C	1Cx	
Reticulatobulbosa/ *Reticulatobulbosa*	**A. schrenkii Regel** (*A. bogdoicola* Regel)	–	–	32[a]	–	–	–	–	[a](Friesen 1985)
Reticulatobulbosa/ *Sikkimensia**	**A. sikkimense Baker** (*A. kansuense* Regel, *A. tibeticum* Rendle)	–	–	16[a], 32[b]	–	–	–	–	[a](Mehra and Pandita 1979), [b](Gu et al. 1993)
Polyprason/ *Falcatifolia**	**A. carolinianum DC.** (*A. aitchisonii*, *A. obtusifolium*)	8[a]	16[b]	16[c], 32[d]	Diploid[f], Tetraploid[f]	52.90 (diploid, Feulgen cytophotometry)[g]	13.23 (diploid)[g]	13.23[g]	[a](Tang et al. 2005), [b](Kumari and Saggoo 2016), [c, e, g](Ohri et al. 1998; Ohri and Pistrick 2001), [d](Gohil and Kaul 1981), [f](Oyuntsetseg et al. 2013), [d, f](Pandita and Mehra 1981b; Dutta et al. 2015)
Polyprason/ *Oreiprason**	**A. roylei Stearn** (*A. lilacinum* Royle et Regel. *A. rubens* Baker)	8[a]	8[b]	16[c]	Diploid[d]	63.00 (diploid)[e], 70.03 (diploid, Feulgen microdensitometry)[f]	15.75 (diploid)[e], 17.51 (diploid)[f]	15.75[e], 17.51[f]	[a, b, c, d](Kohli and Gohil 2011), [b, c, d](Sharma and Gohil 2011a; Kohli and Kaul 2013), [c, d, e](Labani and Elkington 1987), [e](Walters 1992)
Polyprason/ *Falcatifolia*†	**A. platyspathum Schrenk** (*A. platyspathum* subsp. *platyspathum*)	–	–	16[a]	–	–	–	–	[a](Friesen1986; Zakirova and Nafanailova 1988)
Cepa/ *Cepa**	**A. cepa L.** (*A. cepa* var. *aggregatum*, *A. cepa* var. *anglicum*)	8[a]	6[b], 8[c]	14[d], 16[e], 24[f]	Diploid[g], Triploid[h]	65.4 (diploid, flow cytometry), 66.40–69.00 (diploid, Feulgen cytophotometry), 67–71.61 (diploid, Feulgen cytophotometry)[k], 67.5 (diploid, flow cytometry)[l]	16.35 (diploid), 16.60–17.25 (diploid), 16.75–17.90 (diploid)[k], 16.87 (diploid)[m*], 16.2 (diploid)[m*], 17.18–17.32 (diploid)[n*]	16.35[j], 16.60–17.25, 16.75–17.90[k], 16.87, 16.2[m*], 17.18–17.32[n*]	[a, e, g](Mancia et al. 2015), [a, m''](Ricroch et al. 2005), [b](Wang and Zheng 1987), [c](Ved Brat and Dhingra 1973; Gohil and Kaul 1980a; Talukder and Sen 2000; Sharma and Gohil 2011b), [e](Rees et al. 1979; Sato 1981; Joshi and Ranjekar 1982; Cortes et al. 1983; Schubert and Wobus 1985; Fuchs et al. 1995; Johnson and Zhatay 1996; Kim et al. 2002), [e](Van't Hof 1965), [e, f, g, h](Puizina and Papea 1996), [e, g](Narayan 1988; Ahirwar and Verma 2015), [c, i](Arumuganathan and Earle 1991), [c, i](Ulrich et al. 1988), [c, i](Chakravarty and Sen 1992), [n''](Baranyi and Grielhuber 1999), [k](Talukder and Sen 1999)
Cepa/ *Annuloprason**	**A. atrosanguineum Kar. et Kir.** (*A. monadelphum*)	8[a]	–	16[b], 32[d]	diploid[d]	–	–	–	[a, b](Tang et al. 2005), [b, d](Ved Brat 1965), [c](Zhukova 1967)
Cepa/ *Annuloprason**	**A. fedschenkoanum Regel.** (*A. atrosanguineum* var. *fedschenkoanum*)	8[a]	8[b]	16[c]	Diploid[d]	–	–	–	[a, b, d](Pandita and Mehra 1981a), [c, d](Pandita and Mehra 1981b)
Cepa/ *Sacculiferum**	**A. chinense G. Don.** (*A. bakeri*, *A. bodinieri*)	8[a]	–	16[b], 24[c], 32[d]	Triploid[e], Tetraploid[f], Segmental allotetraploid[g]	130.86 (tetraploid, Feulgen cytophotometry)[h]	32.7 (tetraploid)[h]	16.35[h]	[a, d, f, i](Ohri et al. 1998), [a, d, f, g, h](Ohri and Pistrick 2001), [b](Katayama 1928), [c, d, e, f](Wufeng et al. 1993), [d](Ohri et al. 1998; Ogura et al. 1999), [d](Dutta and Bandyopadhyay 2014), [g](Gohil and Koul 1981)

Subgenus/ section	Species (syn.)	Chromosome number			Ploidy	4C DNA value in diploid/ polyploid nuclei (pg)	Genome size in diploid/ polyploid (pg)		References
		Basic (x)	Gametic (n)	Zygotic (2n)			1C	1Cx	
*Cepa Schoenoprasum**	**A. schoenoprasum** L. (A. acutum Spreng., A. alpinum (DC.) Hegetschw.)	8[a]	8[b]	14[c], 16[d], 24[e], 32[f], 48[g]	Diploid[h]	31.20 (diploid, 79)[i], 33.20 (diploid)[j], 33.80 (diploid)[k], 34.90 (diploid)[l] 37.73(diploid, Feulgen Cytophotometry)[m], 60.66 (tetraploid)[n]	7.8 (diploid), 8.3 (diploid)[j], 8.45 (diploid)[k], 8.72 (diploid)[l], 9.43 (diploid)[m], 15.16 (tetraploid)[n]	7.8[i], 8.3; 8.45[j], 8.72; 9.43[m], 7.58[g]	[a,b,h](Pandita and Mehra 1981a), [c](Ohri 1990), [d](Dutta and Bandyopadhyay 2014), [d,h,m](Ohri et al. 1998; Ohri and Pistrick 2001), [e](Kurosawa 1979), [f](El-Gadi and Elkington 1977), [g](Pogosian 1997), [h](Pandita and Mehra 1981b), [i](Ranjekar et al. 1978), [j](Anderson et al. 1985), [k](Jones and Rees 1968), [l](Nanushyan and Polyakov 1989), [n](Labani and Elkington 1987)
–	**A. ascalonicum** L. (A. carneum, A. fissile)	8[a]	8[b]	16[c]	Diploid[d]	66.32–68.67 (diploid, Feulgen cytophotometry)[c]	16.58–17.16 (diploid)[c]	8.29–8.28[c]	[a](Darlington and Wylie 1955), [b,c](Cortes et al. 1983), [d,c](Talukder and Sen 1999)
–	**A. atropurpureum** Waldst. et Kit. (A. nigrum var. atropurpureum)	8[a]	8[b]	16[c], 32[d]	diploid[e], tetraploid[f]	112.81 (tetraploid, Feulgen cytophotometry)[g], 113.66 (diploid, Feulgen cytophotometry)[h]	28.2 (tetraploid)[g], 28.45 (diploid)[h]	14.1[g] 28.45[h]	[a,b,c,e](Koul 1966; Pandita and Mehra 1981a), [c,h](Labani and Elkington 1987), [d,f,g](Ohri et al. 1998; Ohri and Pistrick 2001), [c-h](Gurushidze et al. 2012)
–	**A. blandum** Wall.	–	16[a]	32[b]	Tetraploid[c]	–	–	–	[a,b,c](Mehra and Sachdeva 1976; dGRIP)
–	**A. caesioides** Wendelbo (A. kachrooi)	–	8[a]	16[b]	Diploid[c]	–	–	–	[a,b,c](dGRIP), [a,c](Pandita and Mehra 1981a), [b](Gohil and Kaul 1980a)
–	**A. consanguineum** Kunth	8[a]	8[b]	16[c]	Diploid[d]	–	–	–	[a,b,d](Pandita and Mehra 1981a), [a,b](Gohil and Koul 1971), [c,d](Gohil and Kaul 1980b)
–	**A. hypsistum** Stearn	–	–	32[a]	–	–	–	–	[a](dGRIP)
–	**A. stracheyi** Baker (A. longistaminum Royle)	8[a]	8[b]	16[c], 14[d], 32[e], 48[f]	Diploid[g]	–	–	–	[a,b,g](Pandita and Mehra 1981a), [c](Pandita and Mehra 1981b), [d](Shopova 1966), [d,e](Sen 1974a)

*(Friesen et al. 2006), # (Fritsch et al. 2010), ! (Li et al. 2010), superscripts with the same letters correspond to references from which data are obtained, 1C and 1Cx genome sizes have been calculated from 4C DNA values published in references, '' indicate 1C and 1Cx genome sizes that have been determined following 2C DNA values in corresponding references.
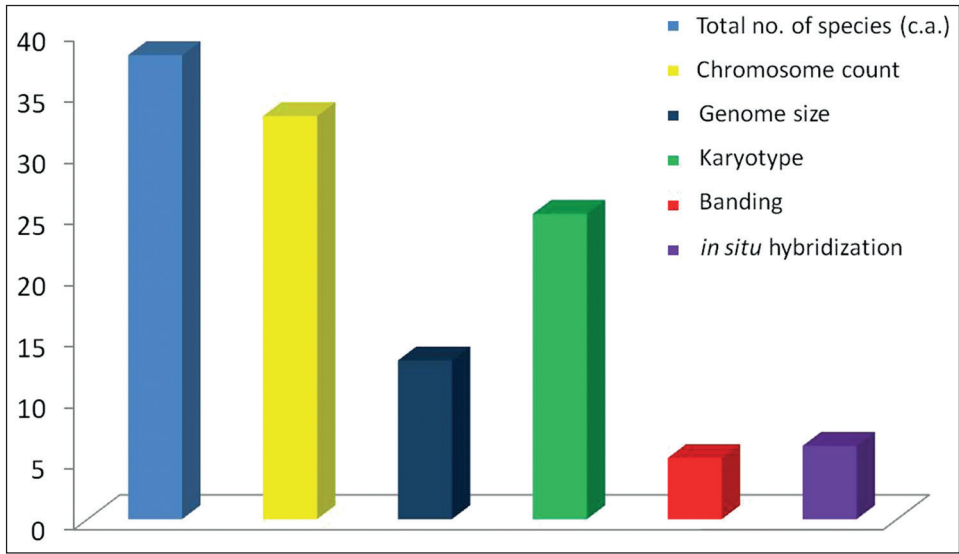
Figure 1. Bar graph showing statistics of cytological reports in the species of *Allium* in India.

Among the diploid species, the range of genome size (Table 1) is from 7.8 pg/1C in *A. schoenoprasum* (subgenus *Cepa*) to 30.0 pg/1C in *A. sativum* Linnaeus, 1753 (subgenus *Allium*). Among the polyploid taxa, the range of genome size (Table 1) is 15.16 pg/1C in *A. schoenoprasum*, 34.35 pg/1C (*A. chitralicum* F.T. Wang et Tang, 1937) to 40.5–41.78 pg/1C in *A. victorialis*. Thus, the lowest values of genome size for the entire array of *Allium* species in India is represented by diploid and polyploid species of *A. schoenoprasum* (subgenus *Cepa*).

The genome size evolution of *Allium* species has been envisaged in relation to growth pattern (dormancy), habitat preference and evolutionary history of the subgenera and sections (Ohri et al. 1998). The authors suggested an overall lack of correlation between genome size and chromosome numbers, although continuity in variation was particularly evident in few species. The present review has showed a 2.25-fold (diploid) or 2.43-fold (tetraploid) difference in genome size in the species occurring in India, although the base number (x) is predominantly 8.

Karyotype features

The karyotype features are known in 8 subgenera and 14 sections of *Allium* species occurring in India (Fig. 1). The majority of species are characterized by metacentric chromosomes except for subgenus *Amerallium* with predominantly submetacentric chromosomes (Table 2). One pair of chromosomes with subterminal constriction has been the characteristic of some species such as *A. cepa* (Sato1981), *A. blandum*, *A. stracheyi* and *A. victoralis* (Mehra and Sachdeva 1976).

Table 2. Karyotype features and molecular chromosomal landmarks in species of *Allium* (Amaryllidaceae, Subfamily Allioideae, Tribe Allieae, sensu APG IV 2016) occurring in India.

Subgenera/ sections	Species	Karyotype — Chromosome morphology	SAT or NORs/ 2n	Heterochromatin banding (Giemsa/ Fluorochrome/others)	rDNA/ telomeric/ other signals — No. of signals/2n	rDNA/ telomeric/ other signals — Features	References
Amerallium/ Bromatorrhiza‡	*A. fasciculatum* Rendle	Majorly submetacentric, few telocentric and metacentric[a]	4[b]	–	–	–	[a, b](Xu et al. 1998; Dutta et al. 2015; Li et al. 2017)
*Amerallium/ Bromatorrhiza**	*A. hookeri* Thwaites	Majorly submetacentric, few metacentric[a]	2[b]	–	–	–	[a, b](Sharma et al. 2011)
Amerallium/ Bromatorrhiza‡	*A. wallichii* Kunth.	Majority submetacentric[a]	2[b]	–	–	–	[a, b](Huang et al. 1995)
*Anguinum/ Anguinum**	*A. prattii* C.H.Wright	Majority metacentric[a]	2[b]/ 4[c]	–	–	–	[a, b](Tang et al. 2005), [a, b, c](Chunying et al. 2000)
Anguinum/ Anguinum‡	*A. victoridis* L.	Majority metacentric[a] or sub–metacentric[b]	2[c]	–	–	–	[a](Pandita and Mehra 1981b), [b, c](Mehra and Sachdeva 1976)
*Butomissa/ Butomissa**	*A. tuberosum* Rottler et Spreng.	Majority metacentric[a] or submetacentric[b,c]	3[d]/ 4[e]/ 6[f]	–	5S:4–6[g]	5S: proximal and intercalary[h]	[a](Kumar and Thonger 2018), [a–c](Talukder and Sen 2000), [a, c, d, e](Sharma and Gohil 2013b), [a, b](Sharma and Gohil 2013a), [g–h](Do and Seo 2000)
*Rhizirideum/ Caespitosoprason**	*Allium przewalskianum* Regel	Majority metacentric chromosomes[a]	2[b]	–	–	–	[a, b](Tang et al. 2005)
*Allium/ Allium**	*A. ampeloprasum* L.	Majorly metacentric, few sub–metacentric[a], few subacrocentric[b]	8[c]	Interstitial C– bands colocalized to silver stained regions in 8 active NORs[d], 8 CMA3⁺/DAPI⁻ bands colocalized to silver stained regions and 35S rDNA sites in NORs[e]	35S:8, 5S: 13 (polymorphic)[f]	35S: interstitial (4) and pericentromeric (4) in short arms[g], 5S: interstitial/ pericentromeric, non–colocalized to 35S except in one chromosome of 8th pair where it flanks 35S site[h]	[a, c, e, f, g, h](Maragheh et al. 2019), [b, c](Koul and Gohil 1970b), [b, c, d](Stack and Roelofs 1996)
*Allium/ Allium**	*A. sativum* L.	Majority metacentric[a]	2[b]/ 4[c]/ 6[d]/ 4–8[e]	C–Bands: nucleolar[f], telomeric and interstitial[g], centromeric (2 pairs)[h]; N–bands: nucleolar (4)[i]; Active NORs (AgNORs)2[j], occasionally4; CMA⁺/DAPI⁻ bands: 4–6[k]	5S: 4[l], 6[m]; 45S and 5S rDNA localized[n]; telomeric signals in all chromosomes[o]; numerous satellite signals[p]	telomeric signals distal[q], satellite signals sub– telomeric and interstitial[r]	[a](Kumar and Thonger 2018), [a–c](Talukder and Sen 2000), [a–k](Bacelar et al. 2021), [k](Koul and Gohil 1970a), [d, f, k](Yuzbasioglu 2004), [c–g–h](Cortes et al. 1983), [l](Cortes and Escalza 1986; Wajahatullah and Vahidy 1990), [l](Do and Seo 2000), [m](Lee et al. 1999; Son et al. 2012), [n](Adams et al. 2001), [o, p, q](Peška et al. 2019)
*Allium/ Arulsea**	*A. griffithianum* Boiss.	Majority metacentric[a]	–	–	–	–	[a](Pandita and Mehra 1981b),
*Allium/ Arulsea**	*Allium rubellum* M. Bieb.	Majority metacentric to sub–metacentric[a]	2[b]/ 6[c]/ 8[d]	–	–	–	[a, b](Abdali and Miri 2020), [a, c](Koul et al. 1971), [a, d](Khoshoo and Sharma 1959)
Allium/ Caeruleat	*A. jacquemontii* Kunth	Majority metacentric[a]	–	–	–	–	[a](Pandita and Mehra 1981b)

Subgenera/ sections	Species	Karyotype		Heterochromatin banding (Giemsa/ Fluorchrome/others)	rDNA/ telomeric/ other signals		References
		Chromosome morphology	SAT or NORs/ 2n		No. of signals/2n	Features	
Reticulatobulbosa/ Reticulatobulbosa#	A. humile Kunth	Majority metacentric[a]	–	–	–	–	[a](Pandita and Mehra 1981b)
Polyprason/ Oreiprason*	A. roylei Stearn	Majority metacentric or sub–metacentric[b]	2[c]	–	–	tyr-FISH mapping of bulb allinase gene[d]	[a,b,c](Sharma and Gohil 2008); [c](Kohli and Gohil 2011), [d](Khrustaleva et al. 2019)
Polyprason/ Falcatifolia*	A. carolinianum DC.	Majorly metacentric, few sub–meta– or sub–telocentric[a]	2[b]	–	–	–	[a,b](Pandita and Mehra 1981b; Tang et al. 2005; Dutta et al. 2015;)
Cepa/ Cepa*	A. cepa L.	Majority metacentric, few submetacentric[a]	1[b], 1–2[c], 1–4[d], 2[e], 2–4[f]	C-Bands: telomeric[g], intercalary[h], distal[i], centromeric and at satellites[j]; heterochromatic CMA/DAPI/ AMD bands at NORs and telomeres[k]	18S–5.8S–25S rDNA loci: 2–4[l], 45S rDNA loci: 3[n], 4[p], 4–5[o], 5[q]; 5S rDNA loci: 2[s], 4[t]	Variable rDNA sites[l]; distal 45S rDNA loci colocalized with telomeric tandem repeat[l] and non–colocalized to 5S loci[s]; 5S loci proximal and distal[h] or interstitial[u]; tyrFISH (with allinase, CHS-B and EST markers) reveal chromosome evolution[v]	[a](Fiskesjo 1975; Sulistyaningsih et al. 2002; Ahirwar and Verma 2014), [a,d,e](Sato 1981), [a,c](Talukder and Sen 2000), [a,f](Battaglia 1957), [b](Bozzini 1964), [c,g,h,i](Puizina and Papea 1996), [c,d](Kim et al. 2002), [c,o,q,s,u,w](Mancia et al. 2015), [f](Ghosh and Roy 1977), [l](Tanaka and Taniguchi 1975), [j](Schubert and Wobus 1985), [m](Fu et al. 2019), [n](Do et al. 2001), [p](Ricroch et al. 1992), [t](Shibata and Hizume 2002), [t](Fajkus et al. 2016), [v](Shibata and Hizume 2002), [s](Khrustaleva et al. 2019)
Cepa/ Annuloprason*	A. atrosanguineum Kar. et Kir.	majorly metacentric[a]	2[b]	–	–	–	[a,b](Tang et al. 2005)
Cepa/ Annuloprason*	A. fedschenkoanum Regel.	Majority metacentric chromosomes[a]	–	–	–	–	[a](Pandita and Mehra 1981b)
Cepa/ Sacculiferum*	A. chinense G. Don.	Majority sub– metacentric[a] or submetacentric[b]	2–4[c]	–	–	–	[a](Ogura et al. 1999), [a,c](Sen 1973b), [b,c](Gohil and Koul 1981)
Cepa/ Schoenoprasum*	A. schoenoprasum L.	Majority metacentric[a]	1–6[b]	C-bands[c]	5S: 4[d]	5S: interstitial in chromosome 6[e]; tyr–FISH of allinase reveal chromosome evolution[f]	[a](Pandita and Mehra 1981b; Cai and Chinnappa 1987), [a,b](Dutta and Bandyopadhyay 2014), [a,c](Tardif and Morisset 1992), [d,e](Shibata and Hizume 2002), [f](Khrustaleva et al. 2019)
–	A. ascalonicum L.	metacentric to sub–metacentric[a]	2[b]	Distal C bands in all chromosomes[c]	–	–	[a,b](Darlington and Haque 1955; Talukder and Sen 2000), [b](Darlington and Wylie 1955), [b,c](Cortes et al. 1983), [c](Seo and Kim 1975)
–	A. atropurpureum Waldst. et Kit.	Majority nearly metacentric and few submetacentric[a]	–	–	–	–	[a](Pandita and Mehra 1981b)
–	A. blandum Wall.	metacentric[a]	–	–	–	–	[a](Mehra and Sachdeva 1976; Pandita and Mehra 1981b)
–	A. consanguineum Kunth	Majority metacentric or sub–metacentric chromosomes[a]	2 (interstitial)[b]	–	–	–	[a](Mehra and Sachdeva 1976), [a,b](Pandita and Mehra 1981b)
–	A. stracheyi Baker	Majority metacentric or sub–metacentric[b]	–	–	–	–	[a](Pandita and Mehra 1981b), [b](Mehra and Sachdeva 1976)

*(Friesen et al. 2006), # (Fritsch et al. 2010), ! (Li et al. 2010), superscripts with the same letters correspond to references from which data are obtained.

The predominance of metacentric chromosomes and symmetric nature of karyotypes is in accordance with earlier studies (Peruzzi et al. 2017). However, few species show a tendency for asymmetry (*A. atrosanguineum*, *A. carolinianum*, *A. griffithianum*, *A. fasciculatum*) and some fall into 2A (*A. chinense*, *A. przewalskianum*) or 2B category (*A. schoenoprasum*, *A. tuberosum*) in Stebbins' index.

Presence of B-chromosomes has been reported in 97 species of *Allium* (Vujošević et al. 2013) belonging mostly to *Allium*, *Cepa* and *Rhizirideum* subgenera (Peruzzi et al. 2017). Among the species found in India, Sharma and Iyengar (1961) first reported the occurrence of B-chromosomes (2–10 in number) in diploid population of *A. stracheyi* and not in the polyploid populations. The B-chromosomes were found to occur in pollen mother cells as well as in pollen grains of *A. stracheyi* (Sen 1974c). However, Mehra and Sachdeva (1976) reported 2n=16 in *A. stracheyi* collected from the Valley of Flowers with no B-chromosome. One or two B-chromosome(s) were reported in *A. ascalonicum* (Bartolo et al. 1984), *A. ampeloprasum*, (subgenus *Allium*) (Khazanehdari and Jones 1996), *A. prattii* (subgenus *Anguinum*) (Chunying et al. 2000), *A. przewalskianum* (subgenus *Rhizirideum*) (Ao 2008; Xie-Kui et al. 2008) while many B-chromosomes (1–10) were recorded in *A. schoenoprasum* (Halkka 1985; Cai and Chinnappa 1987; Tardif and Morisset 1992) and in *A. stracheyi* (subgenus *Cepa*) (Sharma and Aiyangar 1961; Shopova 1966; Pandita and Mehra 1981b).

Nucleolus organizer regions or NORs are significant markers for chromosome identification. Among the species considered presently, NORs/ satellite-bearing chromosomes often show infra-specific or cultivar-specific differences particularly in *A. cepa*, *A. sativum* and *A. tuberosum* (Table 2).

In case of subgenus *Allium*, eight active NORs have been shown in *A. ampeloprasum* by C- banding, CMA3+/DAPI- banding, AgNOR staining and FISH (Table 2). In *A. sativum* secondary constrictions were observed in two to even six chromosomes by C and N banding (Ghosh and Roy 1977; Roy 1978; Cortes et al. 1983), in addition to showing population specific differences (Roy 1978). NORs were also confirmed in four chromosomes by N banding (Cortes and Escalza 1986; Wajahatullah and Vahidy 1990). Recently, two pairs of chromosomes with secondary constrictions were reported in some Brazilian accessions of *A. sativum* of which one pair was suggested to contain intercalary NOR (Bacelar et al. 2021). CMA banding method was used to show the infraspecific heterochromatin variability of nucleolar (proximal) and non-nucleolar (distal and proximal) CMA bands in the Brazilian garlic accessions for their identification. This study remains to be done in case of Indian cultivars.

Allium cepa varieties with different ploidy levels (e.g. *A. cepa* var. *viviparum*, then supposed to be a hybrid between *A. cepa* and *A. fistulosum* Linnaeus, 1753) (Singh et al. 1967; Langer and Koul 1983; Puizina and Papea 1996) show variable number of satellited chromosomes (Bozzini 1964; Singh et al. 1967; Koul and Gohil 1971; Langer and Koul 1983; Puizina and Papea 1996). Many of the conventional staining and C-banding studies showed the presence of two satellite chromosomes in *A. cepa* (Ved Brat and Dhingra 1973; Fiskesjo 1975; Bhattacharyya 1976; Talukder and Sen 2000). Application of differential staining with sequence specific fluorochromes elucidated two NORs in *A. cepa* (Kim et al. 2002). However, reports

claiming variable numbers of NORs (Battaglia 1957; Sato 1981; Puizina and Papea 1996) could not be ruled out. With the application of silver staining, 1–4 active NORs in the satellite region were observed (Sato 1981) while variable number of NORs (2–5) was elucidated by 45S-rDNA hybridization (Table 2). The 45S rDNA sites are distally located and found to co-occur with telomeric tandem repeats (18S). The 5S rDNA loci are reported to range from 2–4 and do not co-occur with the 45S rDNA site.

One interesting feature is that satellites occur mostly in the short arms except for some cases in the subgenera *Allium* and *Amerallium* (Peruzzi et al. 2017). The same phenomenon has been found to exist in case of *A. cepa*, *A. sativum* and *A. ampelopra- sum* (Kim et al. 2002; Maragheh et al. 2019; Bacelar et al. 2021). However, the locali- zation of satellites in the species of *Amerallium* and other subgenera of Indian occur- rence opens interesting scope of future study. The major difference between subgenera *Allium* and *Cepa* lie in the localization of the NORs rather than numbers of rDNA loci. The NORs are interstitial in *Allium* and distal in *Cepa* (Fig. 2) as confirmed by heterochromatic CMA banding, Ag-NOR staining as well as rDNA FISH (Kim et al. 2002; Fajkus et al. 2016; Maragheh et al. 2019; Bacelar et al. 2021).
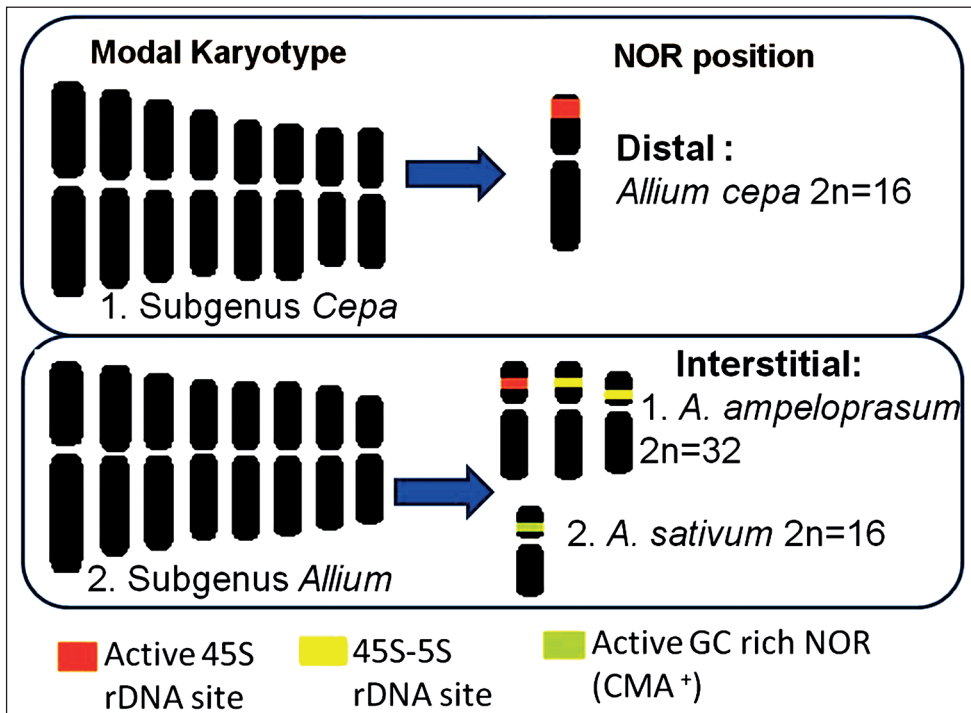
Figure 2. Diagram showing NOR landmarks based on globally published reports in the three species of the genus *Allium* occurring in India. The modal karyotypes for subgenera are adopted and modified after Peruzzi et al. (2017). Diagrams showing NORs are modified after the published reports on *A. am- peloprasum* (as *A. porrum* in Maragheh et al. 2019), *A. sativum* (Bacelar et al. 2021) and *A. cepa* (Fajkus et al. 2016).

Chromosome specialization in *A. cepa*

Telomeres and rDNA loci are the two especially variable features of *A. cepa* chromosomes. Many authors have previously argued that genomic rearrangements are responsible for positional variations of 45S rDNA loci in *A. cepa* (Ricroch et al. 1992; Do et al. 2001; Mancia et al. 2015). The rDNA sequences have been found to contain Copia-like retroelements in *A. cernuum* Roth, 1798 that were dispersed via homogenization mechanisms (Fajkus et al. 2016). The rDNA loci in *A. cepa* have been observed to co-occur with telomeric repeats although telomeres evolved independently of rDNA sequences (Fajkus et al. 2016).

The plant telomere was once thought to be composed of *Arabidopsis* Heynhold, 1842 prototype TTTAGGG repeats (Richards and Ausubel 1988). Exception to this was observed in Asparagales, where an 80 million years old mutation gave rise to human type (TTAGGG) repeat in the family Iridaceae (Adams et al. 2001; Weiss and Scherthan 2002; Sýkorová et al. 2003) and subfamily Allioideae (Sýkorová et al. 2006). The genus *Allium* is different from all other subfamilies of Amaryllidaceae and also other plant groups in terms of the unique telomere sequence. The telomeric sequence (TTATGGGCTCGG)$_n$ surfaced long back (Fuchs et al. 1995) and is neither *Arabidopsis* nor human type. The sequence has been found to be conserved in *Allium*, probing for monophyletic origin of this genus (Fajkus et al. 2016). The telomeres of land plants, including the unique ones like that of Amaryllids, have received less attention (Peska and Garcia 2020). For example, telomeric repeat in *Arabidopsis thaliana* (Linnaeus, 1753) Heynhold, 1842 is a Pol III transcribed lncRNA (Fajkus et al. 2019). Hence, the *Allium* and non-*Allium* taxa of Amaryllidaceae provide excellent scope for studying telomere evolution in eukaryotes.

Recent updates on cytogenetic relationships

A robust phylogenetic analysis supported by genome size and karyotype parameters was found to elucidate the evolution of Gilliesieae of Allioideae (Pellicer et al. 2017). The phylogenetic background of the genus *Allium* has paved way for refinement of classification, inter-species relationships and cytogeographic evolution (Friesen et al. 2006; Gurushidze et al. 2007, 2008; Fritsch et al. 2010; Li et al. 2010, 2017; Abugalieva et al. 2017; Herden et al. 2016; Huo et al. 2019; Costa et al. 2020). Global sampling of 207 species of *Allium* (Allieae) highlighted the ancestral number (x=8) and the reasons behind symmetric karyotype evolution (Peruzzi et al. 2017).

The utility of cytogenetic mapping remains unparallel to investigate synteny comparison between phylogenetically related species that has been employed to interpret chromosome evolution in *Allium* crop species from Russia (Khrustaleva et al. 2019). The presence of flavonoids and sulphur-containing compounds are responsible for the onion's characteristic flavour and the enzyme alliinase is part of the biosynthesis

(Lancaster and Collin 1981). Recent techniques like ultra-sensitive tyramide-FISH (tyr-FISH) and SteamDrop protocol have facilitated the physical detection of the alliinase as well as chalcone synthase genes along with expressed sequence tag (EST) markers. The bulb alliinase gene was located on the long arm of chromosome 4 in *A. cepa* and *A. schoenoprasum* while the same gene was found in the short arm of chromosome 4 in the related (*A. fistulosum*, *A. altaicum* Pallas, 1773, *A. oschaninii* O. Fedtschenko, 1906, and *A. pskemense* B. Fedtschenko, 1905) and phylogenetically distant species (*A. roylei* and *A. nutans* Linnaeus, 1753) (Khrustaleva et al. 2019). Khrustaleva et al. (2019) proposed a pericentric inversion model for rearrangements in chromosome 4 in line with divergence of *A. cepa* and *A. fistulosum*, responsible for breaking collinearity of the genes controlling flavour and bulb colour. This particular report focussed on genomic kinship and genomic rearrangement among the closely related *Allium* species. Also, the practical benefit of molecular cytogenetic mapping becomes apparent in terms of suitably utilizing the genomic resources for onion breeding. These studies would also help to address genomic relationships among *A. cepa*, *A. schoenoprasum* and *A. roylei*, occurring in India.

Summary and future prospects

Considering the impact of cytogenetic investigation in *Allium* phylogeny at a global scale, it is unfortunate to notice the lack of attention in an Indian context in spite of species abundance. Although *A. cepa* has often been regarded as the common material for cytogenetic analysis and the popular '*Allium cepa* test' (Pathiratne et al. 2015; Bonciu et al. 2018), systematic chromosome analysis is still missing in Indian *A. cepa* as well as other species. The present dataset and existing references are not exhaustive but furnish the prerequisite to search for further chromosomal landmarks (NORs, genome size etc) and complement future phylogenetic studies or cyto-geographical evolution of *Allium*, involving the unexplored wild and endemic species in the subcontinent. The crops, onion and garlic, have been admired from ancient time in global cuisines and Indian culinary practices (c.a. 5000 years ago) and continue to be tremendously important in agriculture and pharmaceutical industries (Rana et al. 2011; Nile and Park 2013). The cultivation of onions is challenged by a number of biotic threats which are the direct or indirect manifestation of the current climatic adversity (Le et al. 2021). Identification of wild relatives of the crop having high resistance is germane to address available genomic sources (Dempewolf et al. 2014), which is necessary for *Allium* crop species of India (Gedam et al. 2021). Interesting discoveries on the 'neodomesticate' western Himalaya taxon *A. negianum* (Pandey et al. 2021) along with other endemic less-known species (*A. stracheyi*, *A. roylei*, *A. wallichii* and *A. przewaliskianum*) are assets of Indian repository in line with global assemblages. The genomic attributes of Indian *Allium* germplasm as outlined in this review, could help strategic upgradation of cultivation practices.

Author contribution

Acknowledgements

References

Abdali S, Miri SM (2020) Chromosome counts for six species of *Allium* (Amaryllidaceae) from Iran. The Iranian Journal of Botany 26(2): 179–187.

Abugalieva S, Volkova L, Genievskaya Y, Ivaschenko A, Kotukhov Y, Sakauova G, Turuspekov Y (2017) Taxonomic assessment of *Allium* species from Kazakhstan based on ITS and *matK* markers. BMC Plant Biology 17(2): 51–60. https://doi.org/10.1186/s12870-017-1194-0

Adams SP, Leitch IJ, Bennett MD, Leitch AR (2001) *Aloe* L. – a second plant family without (TTTAGGG) telomeres. Chromosoma 109: 201–205. https://doi.org/10.1007/s004120050429

Ahirwar R, Verma RC (2015) Colchicine induced asynaptic chromosomal behavior at meiosis in *Allium cepa* L. The Nucleus 58: 47–51. https://doi.org/10.1007/s13237-015-0133-4

Anderson LK, Stack SM, Fox MH, Chuanshan Z (1985) The relationship between genome size and synaptonemal complex length in higher plants. Experimental Cell Research 156(2): 367–378. https://doi.org/10.1016/0014-4827(85)90544-0

Angiosperm Phylogeny Group (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Botanical Journal of the Linnean Society 181: 1–20. https://doi.org/10.1111/boj.12385

Ao C (2008) Chromosome numbers and karyotypes of *Allium przewalskianum* populations. Acta Biologica Cracoviensia Series Botanica 50: 43–49.

Arumuganathan K, Earle ED (1991) Estimation of nuclear DNA content of plants by flow cytometry. Plant Molecular Biology Reporter 9: 229–241. https://doi.org/10.1007/BF02672073

Bacelar PAA, Feitoza LL, Valente SES, Gomes RLF, Martins LV, Almeida PM, Silva VB, Lopes ACA, Carvalh R, Peron AP (2021) Variations in heterochromatin content reveal important polymorphisms for studies of genetic improvement in garlic (*Allium sativum* L.). Brazilian Journal of Biology 83: e243514. https://doi.org/10.1590/1519-6984.243514

Baranyi M, Greilhuber J (1999) Genome size in *Allium*: in quest of reproducible data. Annals of Botany 83(6): 687–695. https://doi.org/10.1006/anbo.1999.0871

Bartolo G, Brullo S, Pavone P, Terrasi MC (1984) Cytotaxonomical notes on some Liliaceae of N. Cyrenaica. Webbia 38: 601–622. https://doi.org/10.1080/00837792.1984.10670329

Battaglia E (1957) *Allium ascalonicum* L., *A. fistolosum* L., *A. cepa* L.: Analisi Cariotipica. Caryologia 10: 1–28. https://doi.org/10.1080/00087114.1957.10797610

Bhattacharyya R (1976) A tetraploid *Allium cepa* from Bangladesh. Cytologia 41: 513–521. https://doi.org/10.1508/cytologia.41.513

Bhowmick BK, Jha S (2022) A critical review on cytogenetics of Cucurbitaceae with updates on Indian taxa. Comparative Cytogenetics 16: 93–125. https://doi.org/10.3897/compcytogen.v16.i2.79033

Bonciu E, Firbas P, Fontanetti CS, Wusheng J, Karaismailoğlu MC, Liu D, Menicucci F, Pesnya DS, Popescu A, Romanovsky AV, Schiff S (2018) An evaluation for the standardization of the *Allium cepa* test as cytotoxicity and genotoxicity assay. Caryologia 71(3): 191–209. https://doi.org/10.1080/00087114.2018.1503496

Borowska-Zuchowska N, Senderowicz M, Trunova D, Kolano B (2022) Tracing the evolution of the angiosperm genome from the cytogenetic point of view. Plants 11(6): e784. https://doi.org/10.3390/plants11060784

Bozzini A (1964) On the karyotype of a viviparous onion, known as *Allium cepa* L. var *viviparum* (Metzg.) Alef. Caryologia 17: 459–470. https://doi.org/10.1080/00087114.1964.10796142

Cai Q, Chinnappa CC (1987) Giemsa C-banded karyotypes of seven North American species of *Allium*. American Journal of Botany 74: 1087–1092. https://doi.org/10.2307/2443949

Carta A, Bedini G, Peruzzi L (2020) A deep dive into the ancestral chromosome number and genome size of flowering plants. New Phytologist 228(3): 1097–1106. https://doi.org/10.1111/nph.16668

Chakravarty B, Sen S (1992) DNA and protein contents in different varieties of *Allium cepa* and *Allium sativum*. Allium Improvement Newsletter 1: 61–66.

Chunying X, Jiemei X, Jianquan L (2000) Karyotype studies of *Allium prattii* among 4 populations in southern Qinghai. Acta Botanica Boreali-Occidentalia Sinica 20(2): 288–293.

Cortes F, Escalza P (1986) Analysis of different banding patterns and late replicating regions in chromosomes of *Allium cepa*, *A. sativum* and *A. nigrum*. Genetica 71: 39–46. https://doi.org/10.1007/BF00123231

Cortes F, Gonzalez-Gil G, Hazen MJ (1983) C-banding and sister chromatid exchanges in three species of the genus *Allium* (*A. cepa*, *A. ascalonicum* and *A. sativum*). Caryologia 36: 203–210. https://doi.org/10.1080/00087114.1983.10797661

Costa L, Jimenez H, Carvalho R, Carvalho-Sobrinho J, Escobar I, Souza G (2020) Divide to conquer: evolutionary history of Allioideae tribes (Amaryllidaceae) is linked to distinct trends of karyotype evolution. Frontiers in Plant Science 11: e320. https://doi.org/10.3389/fpls.2020.00320

Darlington CD, Haque A (1955) The timing of mitosis and meiosis in *Allium ascalonicum*: A problem of differentiation. Heredity 9: 117–127. https://doi.org/10.1038/hdy.1955.6

Darlington CD, Wylie AP (1955) Chromosome Atlas of Flowering Plants. George Allen & Unwin Ltd., London.

Dempewolf H, Eastwood RJ, Guarino L, Khoury CK, Muller JV, Toll J (2014) Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. Agroecology and Sustainable Food Systems 38(4): 369–377. https://doi.org/10.1080/21683565.2013.870629

dGRIP (2022) Database for genome related information in Indian plants. http://indianpcd.com/

Do GS, Seo BB (2000) Phylogenetic relationships among *Allium* subg. *Rhizirideum* species based on the molecular variation of 5S rRNA genes. Korean Journal of Biological Sciences 4(1): 77–85. https://doi.org/10.1080/12265071.2000.9647527

Do GS, Seo BB, Yamamoto M, Suzuki G, Mukai Y (2001) Identification and chromosomal location of tandemly repeated DNA sequences in *Allium cepa*. Genes & Genetic Systems 76: 53–60. https://doi.org/10.1266/ggs.76.53

Dutta M, Bandyopadhyay M (2014) Comparative karyomorphological studies of three edible locally important species of *Allium* from India. The Nucleus 57(1): 25–31. https://doi.org/10.1007/s13237-014-0106-z

Dutta M, Negi KS, Bandyopadhyay M (2015) Novel cytogenetic resources of wild *Allium* (Amaryllidaceae) from India. The Nucleus 58(1): 15–21. https://doi.org/10.1007/s13237-015-0130-7

El-Gadi A, Elkington TT (1977) Numerical taxonomic studies on species in *Allium* subgenus *Rhizirideum*. New Phytologist 79: 183–201. https://doi.org/10.1111/j.1469-8137.1977.tb02195.x

Fajkus P, Peška V, Sitová Z, Fulnečková J, Dvořáčková M, Gogela R, Sýkorová E, Hapala J, Fajkus J (2016) *Allium* telomeres unmasked: the unusual telomeric sequence (CTCGGT-TATGGG)*n* is synthesized by telomerase. The Plant Journal 85(3): 337–347. https://doi.org/10.1111/tpj.13115

Fajkus P, Peška V, Závodník M, Fojtová M, Fulnečková J, Dobias Š, Kilar A, Dvořáčková M, Zachová D, Nečasová I, Sims J, Sýkorová E, Fajkus J (2019) Telomerase RNAs in land plants. Nucleic Acids Research 47(18): 9842–9856. https://doi.org/10.1093/nar/gkz695

Fiskesjo G (1975) Chromosomal relationships between three species of *Allium* as revealed by C-banding. Hereditas 81: 23–32. https://doi.org/10.1111/j.1601-5223.1975.tb01010.x

Friesen N, Fritsch RM, Blattner FR (2006) Phylogeny and new intrageneric classification of *Allium* (Alliaceae) based on nuclear ribosomal DNA ITS sequences. Aliso: A Journal of Systematic and Floristic Botany 22(1): 372–395. https://doi.org/10.5642/aliso.20062201.31

Friesen NV (1985) Chromosome numbers in the representatives of the family Alliaceae from Siberia. Botanicheskii Zhurnal SSSR 70(7): 1001–1002.

Friesen NV (1986) Chromosome numbers of the representatives of the family Alliaceae from Siberia. Botanicheskii Zhurnal 71: 113–115.

Fritsch RM, Blattner FR, Gurushidz M (2010) New classification of *Allium* L. subg. *Melanocrommyum* (Webb & Berthel.) Rouy (Alliaceae) based on molecular and morphological characters. Phyton (Horn) 49(2): 145–220.

Fu J, Zhang H, Guo F, Ma L, Wu J, Yue M, Zheng X, Qiu Z, Li L (2019) Identification and characterization of abundant repetitive sequences in *Allium cepa*. Scientific Reports 9: e16756. https://doi.org/10.1038/s41598-019-52995-9

Fuchs J, Brandes A, Schubert I (1995) Telomere sequence localization and karyotype evolution in higher plants. Plant Systematics and Evolution 196: 227–241. https://doi.org/10.1007/BF00982962

Gedam PA, Thangasamy A, Shirsat DV, Ghosh S, Bhagat KP, Sogam OA, Gupta AJ, Mahajan V, Soumia PS, Salunkhe VN, Khade YP, Gawande SJ, Hanjagi PS, Shiv Ramakrishnan R, Singh M (2021) Screening of onion (*Allium cepa* L.) genotypes for drought tolerance using physiological and yield based indices through multivariate analysis. Frontiers in Plant Science 12: e122. https://doi.org/10.3389/fpls.2021.600371

Ghosh S, Roy SC (1977) Orientation of interphase chromosomes as detected by Giemsa C-bands. Chromosoma 61: 49–55. https://doi.org/10.1007/BF00292679

Gohil RN, Kaul R (1979) Seed progeny studies in *Alliums*. I. Numerical variants in the progeny of tetraploid *Allium tuberosum* Rottl. ex Spreng. Beiträge zur Biologie der Pflanzen 54: 304–309.

Gohil RN, Kaul R (1980a) Studies on male and female meiosis in Indian *Allium* I. Four diploid species. Chromosoma 77: 123–127. https://doi.org/10.1007/BF00329538

Gohil RN, Kaul R (1980b) An interesting variation in the development of the female gametophyte of *Allium consanguineum*. Caryologia 33(2): 295–297. https://doi.org/10.1080/00087114.1980.10796843

Gohil RN, Kaul R (1981) Studies on male and female meiosis in Indian *Allium* II, Autotetraploid *Allium tuberosum*. Chromosoma 82: 735–739. https://doi.org/10.1007/BF00285778

Gohil RN, Koul AK (1971) Desynapsis in some diploid and polyploidy species of *Allium*. Canadian Journal of Genetics and Cytology 13: 723–728. https://doi.org/10.1139/g71-104

Gohil RN, Koul AK (1973) Some adaptive genetic-evolutionary processes accompanying polyploidy in the Indian *Alliums*. Botaniska Notiser 126: 426–432.

Gohil RN, Koul AK (1977) The cause of multivalent suppression in *Allium ampeloprasum* L. Beiträge zur Biologie der Pflanzen 53: 473–478.

Gohil RN, Koul AK (1981) Cytology of the tetraploid *Allium chinense* G. Don. Caryologia 34(1): 73–81. https://doi.org/10.1080/00087114.1981.10796874

Gohil RN, Koul AK (1983) Seed progeny studies in *Alliums*. II. Male meiosis in the progeny plants of tetraploid *Allium tuberosum* Rottl. ex Spreng. Cytologia 48: 109–118. https://doi.org/10.1508/cytologia.48.109

Goldblatt P, Lowry PP (2011) The Index to Plant Chromosome Numbers (IPCN): three decades of publication by the Missouri Botanical Garden come to an end. Annals of the Missouri Botanical Garden 98(2): 226–227. https://doi.org/10.3417/2011027

Gu Z, Wang L, Sun H, Wu S (1993) A cytological study of some plants from Qinghai-Xizang Plateau. Acta Botanica Yunnanica 15: 377–384.

Gurushidze M, Fritsch RM, Blattner FR (2008) Phylogenetic analysis of *Allium* subg. *Melanocrommyum* infers cryptic species and demands a new sectional classification. Molecular Phylogenetics and Evolution 49(3): 997–1007. https://doi.org/10.1016/j.ympev.2008.09.003

Gurushidze M, Fuchs J, Blattner FR (2012) The evolution of genome size variation in drumstick onions (*Allium* subgenus *Melanocrommyum*). Systematic Botany 37(1): 96–104. https://doi.org/10.1600/036364412X616675

Gurushidze M, Mashayekhi S, Blattner FR, Friesen N, Fritsch RM (2007) Phylogenetic relationships of wild and cultivated species of *Allium* section *Cepa* inferred by nuclear rDNA ITS sequence analysis. Plant Systematics and Evolution 269(3): 259–269. https://doi.org/10.1007/s00606-007-0596-0

Halkka L (1985) Chromosome counts of Finnish vascular plants. Annales Botanici Fennici 22: 315–317.

Haston E, Richardson JE, Stevens PF, Chase MW, Harris DJ (2009) The Linear Angiosperm Phylogeny Group (LAPG) III: a linear sequence of the families in APG III. Botanical Journal of the Linnean Society 161(2): 128–131. https://doi.org/10.1111/j.1095-8339.2009.01000.x

Herden T, Hanelt P, Friesen N (2016) Phylogeny of *Allium* L. subgenus *Anguinum* (G. Don. ex WDJ Koch) N. Friesen (Amaryllidaceae). Molecular Phylogenetics and Evolution 95: 79–93. https://doi.org/10.1016/j.ympev.2015.11.004

Huang R, Wei R, Yan Y (1985) Discovery of spontaneous triploid of *Allium tuberosum*. Journal of Wuhan Botanical Research 3: 429–431.

Huang R, Xu J, Hong Y (1995) A study on karyotypes and their evolutionary trends in *Allium* sect. *Bromatorrhiza* Ekberg (Liliaceae). Cathaya 7: 133–145.

Huo Y, Gao L, Liu B, Yang Y, Kong S, Sun Y, Yang Y, Wu X (2019) Complete chloroplast genome sequences of four *Allium* species: comparative and phylogenetic analyses. Scientific Reports 9(1): 1–14. https://doi.org/10.1038/s41598-019-48708-x

Islam-Faridi N, Sakhnokho HF, Nelson CD (2020) New chromosome number and cyto-molecular characterization of the African Baobab (*Adansonia digitata* L.) – "The Tree of Life". Scientific Reports 10: e13174. https://doi.org/10.1038/s41598-020-68697-6

Jha S, Raina SN, Ohri D, Verma RC, Dhar MK, Lekhak MM, Yadav SR, Mahadev N, Satyawada RR (2019) A new online database on genome-related information of Indian plants. Plant Systematics and Evolution 305(9): 837–843. https://doi.org/10.1007/s00606-019-01602-5

Johnson MAT, Zhatay N (1996) Cytology of *Allium* sect. *Allium*. In: Mathew B (Ed.) A Review of *Allium* sect. *Allium*. Kew Royal Botanic Gardens, Kew, 17–31.

Jones RN, Rees H (1968) Nuclear DNA variation in *Allium*. Heredity 23: 591–605. https://doi.org/10.1038/hdy.1968.76

Joshi CP, Ranjekar PK (1982) Visualization and distribution of heterochromatin in interphase nuclei of several plant species as revealed by a new giemsa banding technique. Cytologia 47: 471–480. https://doi.org/10.1508/cytologia.47.471

Katayama Y (1928) The chromosome number in *Phaseolus* and *Allium* and observation on the size of stomata in different species of *Triticum*. Jour. Sci. Agric. Soc. Tokyo 303: 52–54.

Khazanehdari KA, Jones GH (1996) Meiotic synapsis of the *Allium porrum* B chromosome: evidence for a derived isochromosome origin. Genome 39(6): 1199–1204. https://doi.org/10.1139/g96-151

Khoshoo TN, Sharma VB (1959) Cytology of the autotriploid *Allium rubellum*. Chromosoma 10: 136–143. https://doi.org/10.1007/BF00396567

Khrustaleva L, Kudryavtseva N, Romanov D, Ermolaev A, Kirov I (2019) Comparative Tyramide-FISH mapping of the genes controlling flavor and bulb color in *Allium* species revealed an altered gene order. Scientific Reports 9(1): 1–11. https://doi.org/10.1038/s41598-019-48564-9

Kim ES, Punina EO, Rodionov AV (2002) Chromosome CPD (PI/DAPI) and CMA/DAPI-banding patterns in *Allium cepa* L. Russian Journal of Genetics 38(4): 489–496. https://doi.org/10.1023/A:1015250219322

Kohli B, Gohil RN (2011) Is *Allium roylei* Stearn still evolving through multiple interchanges? The Nucleus 54(1): 19–23. https://doi.org/10.1007/s13237-011-0018-0

Kohli B, Kaul V (2013) Sterility in *Allium roylei* Stearn – A lesser explored taxon. International Journal of Pharma and Bio Sciences 4(1): 741–746.

Kojima A, Hinata K, Noda S (1991) An improvement of squash method for cytological study of female meiosis in *Allium tuberosum*, Liliaceae. Chromosome Information Service (CIS) 50: 5–7.

Koul AK (1963) A spontaneously occurring reciprocal translocation heterozygote of *Allium cepa*. The Journal of Indian Botanical Society 42: 416–419.

Koul AK (1966) Structural hybridity in *Allium atropurpureum* Waldst. & Kit. Journal of Cytology and Genetics 1: 87–89.

Koul AK, Gohil RN (1970a) Causes averting sexual reproduction in *Allium sativum* L. Cytologia 35: 197–202. https://doi.org/10.1508/cytologia.35.197

Koul AK, Gohil RN (1970b) Cytology of tetraploid *Allium ampeloprasum* with chiasma localization. Chromosoma 29: 12–19. https://doi.org/10.1007/BF01183658

Koul AK, Gohil RN (1971) Further studies on natural triploidy in viviparous onion. Cytologia 36(2): 253–261. https://doi.org/10.1508/cytologia.36.253

Koul AK, Sharma MC, Gohil RN (1971) Cytology of the tetraploid *Allium rubellum* Bieb. Caryologia 24: 149–155. https://doi.org/10.1080/00087114.1971.10796422

Kumar S, Thonger T (2018) Karyomorphology of five *Allium* species from Nagaland, North-Eastern Region of India. Jordan Journal of Biological Sciences 11(1): 9–15.

Kumari K, Saggoo MIS (2016) Male meiosis and morphometric analysis of ethnobotanically important *Allium carolinianum* DC. from Kinnaur district of Himachal Pradesh, India. Asian Journal of Pharmaceutical and Clinical Research 9(4): 396–398.

Kurosawa S (1966) Cytological studies on some eastern Himalayan plants. In: Hara H (Ed.) The Flora of Eastern Himalaya. University of Tokyo, Japan, 658–690.

Kurosawa S (1979) Notes on chromosome numbers of Spermatophytes II. Journal of Japanese Botany 54: 155–160.

Labani RM, Elkington TT (1987) Nuclear DNA variation in the genus *Allium* L. (Liliaceae). Heredity 59: 119–128. https://doi.org/10.1038/hdy.1987.103

Lancaster JE, Collin HA (1981) Presence of alliinase in isolated vacuoles and of alkyl cysteine sulphoxides in the cytoplasm of bulbs of onion (*Allium cepa*). Plant Science Letters 22(2): 169–176. https://doi.org/10.1016/0304-4211(81)90139-5

Langer A, Koul AK (1983) Studies on nucleolus and nucleolar chromosomes in angiosperms VII. Nature of nucleolar chromosome polymorphism in *Allium cepa* var. *viviparum* (Metzg.) Alef. Cytologia 48: 323–332. https://doi.org/10.1508/cytologia.48.323

Le D, Audenaert K, Haesaert G (2021) *Fusarium* basal rot: profile of an increasingly important disease in *Allium* spp. Tropical Plant Pathology 46: 241–253. https://doi.org/10.1007/s40858-021-00421-9

Lee SH, Do GS, Seo BB (1999) Chromosomal localization of 5S rRNA gene loci and the implications for relationships within the *Allium* complex. Chromosome Research 7: 89–93. https://doi.org/10.1023/a:1009222411001

Levan A (1934) Cytological studies in *Allium*, V *Allium macranthum*. Hereditas 18: 349–359. https://doi.org/10.1111/j.1601-5223.1934.tb02619.x

Levan A (1940) Meiosis of *Allium porrum*. A tetraploid species with chiasma localisation. Hereditas 26: 454–462. https://doi.org/10.1111/j.1601-5223.1940.tb03248.x

Li MJ, Guo XL, Li J, Zhou SD, Liu Q, He XJ (2017) Cytotaxonomy of *Allium* (Amaryllidaceae) subgenera *Cyathophora* and *Amerallium* sect. *Bromatorrhiza*. Phytotaxa 331(2): 185–198. https://doi.org/10.11646/phytotaxa.331.2.3

Li QQ, Zhou SD, He XJ, Yu Y, Zhang YC, Wei XQ (2010) Phylogeny and biogeography of *Allium* (Amaryllidaceae: Allieae) based on nuclear ribosomal internal transcribed spacer and chloroplast rps16 sequences, focusing on the inclusion of species endemic to China. Annals of Botany 106: 709–733. https://doi.org/10.1093/aob/mcq177

Li R, Liu L, Wang X (1985) Karyotype analysis on the different cultivars of *Allium tuberosum* Rottle. Chinese Bulletin of Botany 3(5): 43–46.

Lu Y, Deng Y, Lu L, He X (2017) Karyotypes of nineteen populations of four species in *Allium* subgenus *Anguinum*. Guangxi Zhiwu/Guihaia 37(7): 811–821.

Mancia FH, Sohn SH, Ahn YK, Kim DS, Kim JS, Kwon YS, Kim CW, Lee TH, Hwang YJ (2015) Distribution of various types of repetitive DNAs in *Allium cepa* L. based on dual color FISH. Horticulture, Environment and Biotechnology 56(6): 793–799. https://doi.org/10.1007/s13580-015-1100-3

Maragheh FP, Janus D, Senderowicz M, Haliloglu K, Kolano B (2019) Karyotype analysis of eight cultivated *Allium* species. Journal of Applied Genetics 60(1): 1–11. https://doi.org/10.1007/s13353-018-0474-1

Mehra PN, Pandita TK (1979) IOPB chromosome number reports LXIV. Taxon 28: 391–408. https://doi.org/10.1002/j.1996-8175.1979.tb00530.x

Mehra PN, Sachdeva SK (1975) IOPB chromosome number reports XLIX. Taxon 24: 501–516. https://doi.org/10.1002/j.1996-8175.1975.tb00341.x

Mehra PN, Sachdeva SK (1976) Cytological observations on some West Himalayan monocots. III. Alliaceae. Cytologia 41: 23–30. https://doi.org/10.1508/cytologia.41.23

Murin A (1976) Polyploidy and mitotic cycle. The Nucleus 19: 192–195.

Nanda S, Chand SK, Mandal P, Tripathy P, Joshi RK (2016) Identification of novel source of resistance and differential response of *Allium* genotypes to purple blotch pathogen, *Alternaria porri* (Ellis) Ciferri. The Plant Pathology Journal 32(6): 519–527. https://doi.org/10.5423/PPJ.OA.02.2016.0034

Nanushyan ER, Polyakov VJ (1989) Zavisimost mezhdu kolichestvom DNK, tolshchinoy mitoticheskikh khromosom i obyemom pyltsevykh zeren u nekotorykh vidov roda *Allium* L. Biologicheskiye Nauki (Moskva) 8: 50–56.

Narayan RKJ (1988) Constraints upon the organization and evolution of chromosomes in *Allium*. Theoretical and Applied Genetics 75: 319–329. https://doi.org/10.1007/BF00303971

Nath S, Sarkar S, Patil SD, Saha PS, Lekhak MM, Ray S, Rama Rao S, Yadav SR, Verma RC, Dhar MK, Raina SN, Jha S (2022) Cytogenetic diversity in Scilloideae (Asparagaceae): a comprehensive recollection and exploration of karyo-evolutionary trends. The Botanical Review. https://doi.org/10.1007/s12229-022-09279-1

Nile SH, Park SW (2013) Total phenolics, antioxidant and xanthine oxidase inhibitory activity of three colored onions (*Allium cepa* L.). Frontiers in Life Science 7(3–4): 224–228. https://doi.org/10.1080/21553769.2014.901926

Ogura H, Kondo K, Morimoto M, Aizawa T, Chen Z, Hong D (1999) A karyological study of *Allium grayi* Regel and *A. chinense* G. Don in Sichuan Province, China. Chromosome Science 3: 119–122.

Ohri D, Fritsch RM, Hanelt P (1998) Evolution of genome size in *Allium* (Alliaceae). Plant Systematics and Evolution 210: 57–86. https://doi.org/10.1007/BF00984728

Ohri D, Pistrick K (2001) Phenology and genome size variation in *Allium* L.- a tight correlation? Plant Biology (Stuttgart) 3: 654–660. https://doi.org/10.1055/s-2001-19362

Ohri M (1990) Studies on the factor of existence in *Allium chinense* guessed from elimination of the constitution of chromosomes. Journal of the Faculty of Agriculture, Shinshu University 27: 49–90.

Olszewska MJ, Osiecka R (1982) The relationship between 2C DNA content, life cycle type, systematic position, and the level of DNA endoreplication in nuclei of parenchyma cells during growth and differentiation of roots in some monocotyledonous species. Biochemie und Physiologie der Pflanzen 177(4–5): 319–336. https://doi.org/10.1016/S0015-3796(82)80026-7

Oyuntsetseg B, Friesen N, Darikhand D (2013) *Allium carolinianum* DC., A new species to the outer Mongolia. Turczaninowia 16(2): 88–90.

Pandey A, Malav PK, Rai KM, Ahlawat SP (2022) Genus *Allium* L. of the Indian Region: A field guide for germplasm collection and identification. ICAR-National Bureau of Plant Genetic Resources, New Delhi.

Pandey A, Malav PK, Rai MK, Ahlawat SP (2021) 'Neodomesticates' of the Himalayan *Allium* spices (*Allium* species) in Uttarakhand, India and studies on eco-geography and morphology. Genetic Resources and Crop Evolution 68(5): 2167–2179. https://doi.org/10.1007/s10722-021-01164-x

Pandita TK, Mehra PN (1981a) Cytology of *Alliums* of Kashmir Himalayas, III. Male Meiosis. The Nucleus 24(3): 147–151.

Pandita TK, Mehra PN (1981b) Cytology of *Alliums* of Kashmir Himalayas, I. Wild species. The Nucleus 24(1, 2): 5–10.

Pathiratne A, Hemachandra CK, De Silva N (2015) Efficacy of *Allium cepa* test system for screening cytotoxicity and genotoxicity of industrial effluents originated from different industrial activities. Environmental Monitoring and Assessment 187(12): 1–12. https://doi.org/10.1007/s10661-015-4954-z

Pedersen K, Wendelbo P (1966) Chromosome numbers of some SW Asian *Allium* species. Blyttia 24: 307–313.

Pellicer J, Hidalgo O, Walker J, Chase MW, Christenhusz MJ, Shackelford G, Leitch IJ, Fay MF (2017) Genome size dynamics in tribe Gilliesieae (Amaryllidaceae, subfamily Allioideae) in the context of polyploidy and unusual incidence of Robertsonian translocations. Botanical Journal of the Linnean Society 184(1): 16–31. https://doi.org/10.1093/botlinnean/box016

Pellicer J, Leitch IJ (2020) The Plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. New Phytologist 226(2): 301–305. https://doi.org/10.1111/nph.16261

Peruzzi L, Carta A, Altinordu F (2017) Chromosome diversity and evolution in *Allium* (Allioideae, Amaryllidaceae). Plant Biosystems – An International Journal Dealing with all Aspects of Plant Biology 151(2): 212–220. https://doi.org/10.1080/11263504.2016.1149123

Peska V, Garcia S (2020) Origin, diversity, and evolution of telomere sequences in plants. Frontiers in Plant Science 11: e117. https://doi.org/10.3389/fpls.2020.00117

Peska V, Mandakova T, Ihradska V, Fajkus J (2019) Comparative dissection of three giant genomes: *Allium cepa*, *Allium sativum*, and *Allium ursinum*. International Journal of Molecular Sciences 20(3): e733. https://doi.org/10.3390/ijms20030733

Phuong PTM, Tashiro Y (2010) Study on diversity and chromosome numbers of edible *Allium* crops in Vietnam. Journal of Science and Development 8: 138–144.

Pogosian AI (1997) Chromosome numbers in some species of monocotyledons from the Transcaucasia. Botanicheskii Zhurnal (Moscow & Leningrad) 82(6): 117–118.

Puizina J, Papea D (1996) Cytogenetical evidences for hybrid structure and origin of diploid and triploid shallots (*Allium cepa* var. *viviparum*, Liliaceae) from Dalmatia (Croatia). Plant Systematics and Evolution 199: 203–215. https://doi.org/10.1007/BF00984905

Rana S, Pal R, Vaiphei K, Sharma S, Ola R (2011) Garlic in health and disease. Nutrition Research Reviews 24(1): 60–71. https://doi.org/10.1017/S0954422410000338

Ranjekar PK, Pallotta D, Lafontaine JG (1978) Analysis of plant genomes V. Comparative study of molecular properties of DNAs of seven *Allium* species. Biochemical Genetics 16: 957–970. https://doi.org/10.1007/BF00483747

Rees H, Narayan RKJ, Hutchinson J (1979) DNA variation associated with the evolution of flowering species. The Nucleus 22: 1–5.

Rice A, Glick L, Abadi S, Einhorn M, Kopelman NM, Salman-Minkov A, Mayzel J, Chay O, Mayrose I (2015) The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. New Phytologist 206(1): 19–26. https://doi.org/10.1111/nph.13191

Richards EJ, Ausubel FM (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53(1): 127–136. https://doi.org/10.1016/0092-8674(88)90494-1

Ricroch A, Peffley EB, Baker RJ (1992) Chromosomal location of rDNA in *Allium*: *in situ* hybridization using biotin-and fluorescein-labelled probe. Theoretical and Applied Genetics 83(4): 413–418. https://doi.org/10.1007/BF00226528

Ricroch A, Yockteng R, Brown SC, Nadot S (2005) Evolution of genome size across some cultivated *Allium* species. Genome 48: 511–520. https://doi.org/10.1139/g05-017

Roy SC (1978) Polymorphism in giemsa banding patterns in *Allium sativum*. Cytologia 43: 97–100. https://doi.org/10.1508/cytologia.43.97

Sato S (1981) Cytological studies on the satellited chromosomes of *Allium cepa*. Caryologia 34(4): 431–440. https://doi.org/10.1080/00087114.1981.10796911

Sato S, Kawamura S (1981) Cytological studies on the nucleolus and the NOR-carrying segments of *Allium sativum*. Cytologia 46: 781–790. https://doi.org/10.1508/cytologia.46.781

Schubert I, Wobus U (1985) In situ hybridization confirms jumping nucleolus organizing regions in *Allium*. Chromosoma 92: 143–148. https://doi.org/10.1007/BF00328466

Sen S (1973a) Structural hybridity intra- and interspecific level in Liliales. Folia Biologica (Cracow) 21: 183–197.

Sen S (1973b) Polysomaty and its significance in Liliales. Cytologia 38: 737–751. https://doi.org/10.1508/cytologia.38.737

Sen S (1974a) Cryptic structural changes in the evolution of cultivated *Alliums*. Indian Journal of Heredity 8: 41–50.

Sen S (1974b) Floral biology, meiosis, pollen cytology and cause of seed setting in *Allium tuberosum* Rottl. Caryologia 27(1): 7–16. https://doi.org/10.1080/00087114.1974.10796557

Sen S (1974c) Nature and behaviour of B chromosomes in *Allium stracheyii* Baker and *Urginea indica* Kunth. Cytologia 39: 245–251. https://doi.org/10.1508/cytologia.39.245

Senderowicz M, Nowak T, Rojek-Jelonek M, Bisaga M, Papp L, Weiss-Schneeweiss H, Kolano B (2021) Descending dysploidy and bidirectional changes in genome size accompanied *Crepis* (Asteraceae) evolution. Genes 12(9): e1436. https://doi.org/10.3390/genes12091436

Seo B (1977) Cytogenetic studies of some tetraploids in *Allium*. Korean Journal of Botany 20: 71–76.

Seo BB, Kim JH (1975) Karyotypic analyses based on heterochromatin distribution in *Allium fistulosum* and *Allium ascalonicum*. Korean Journal of Botany 18: 92–100.

Sharma AK, Aiyangar HR (1961) Occurrence of B-chromosomes in diploid *Allium stracheyi* Baker and their elimination in polyploids. Chromosoma 12(1): 310–317. https://doi.org/10.1007/BF00328926

Sharma G, Gohil RN (2003) Cytology of *Allium roylei* Stearn. 1. Meiosis in a population with complex interchanges. Cytologia 68: 115–119. https://doi.org/10.1508/cytologia.68.115

Sharma G, Gohil RN (2004) Chromosomal chimeras in the male track of *Allium tuberosum* Rottl. ex Spreng. Caryologia 57(2): 158–162. https://doi.org/10.1080/00087114.2004.10589386

Sharma G, Gohil RN (2008) Intrapopulation karyotypic variability in *Allium roylei* Stearn – a threatened species. Botanical Journal of the Linnean Society 158(2): 242–248. https://doi.org/10.1111/j.1095-8339.2008.00862.x

Sharma G, Gohil RN (2011a) Occurrence of differential meiotic associations and additional chromosomes in the embryo-sac mother cells of *Allium roylei* Stearn. Journal of Genetics 90: 45–49. https://doi.org/10.1007/s12041-011-0031-8

Sharma G, Gohil RN (2011b) Occurrence of multivalents and additional chromosomes in the pollen mother cells of *Allium cepa* L. The Nucleus 54(3): 137–140. https://doi.org/10.1007/s13237-011-0042-0

Sharma G, Gohil RN (2013a) Origin and cytology of a novel cytotype of *Allium tuberosum* Rottl. ex Spreng. (2n=48). Genetic Resources and Crop Evolution 60: 503–511. https://doi.org/10.1007/s10722-012-9852-4

Sharma G, Gohil RN (2013b) Double hypoploid of *Allium tuberosum* Rottl. ex Spreng. (2n=4x=30): its origin and cytology. Genetic Resources and Crop Evolution 60: 2283–2292. https://doi.org/10.1007/s10722-013-9995-y

Sharma G, Gohil RN, Kaul V (2011) Cytological status of *Allium hookeri* Thwaites (2n=22). Genetic Resources and Crop Evolution 58: 1041–1050. https://doi.org/10.1007/s10722-010-9641-x

Shibata F, Hizume M (2002) Evolution of 5S rDNA units and their chromosomal localization in *Allium cepa* and *Allium schoenoprasum* revealed by microdissection and FISH. Theoretical and Applied Genetics 105(2–3): 167–172. https://doi.org/10.1007/s00122-002-0950-0

Shopova M (1966) The nature and behaviour of supernumerary chromosomes in the *Rhizirideum* group of the genus *Allium*. Chromosoma 19: 149–158. https://doi.org/10.1007/BF00293680

Singh F, Ved Brat S, Khoshoo TN (1967) Natural triploidy in viviparous onions. Cytologia 32: 403–407. https://doi.org/10.1508/cytologia.32.403

Son JH, Park KC, Lee SI, Jeon EJ, Kim HH, Kim NS (2012) Sequence variation and comparison of the 5S rRNA sequences in *Allium* species and their chromosomal distribution in four *Allium* species. Journal of Plant Biology 55: 15–25. https://doi.org/10.1007/s12374-011-9185-4

Stack SM, Roelofs D (1996) Localized chiasmata and meiotic nodules in the tetraploid onion *Allium porrum*. Genome 39: 770–783. https://doi.org/10.1139/g96-097

Sulistyaningsih E, Yamastai K, Tashiro Y (2002) Genetic characteristics of the Indonesian white shallot. Journal of the Japanese Society for Horticultural Science 71: 504–508. https://doi.org/10.2503/jjshs.71.504

Sýkorová E, Leitch AR, Fajkus J (2006) Asparagales telomerases which synthesize the human type of telomeres. Plant Molecular Biology 60(5): 633–646. https://doi.org/10.1007/s11103-005-5091-9

Sýkorová E, Lim KY, Kunická Z, Chase MW, Bennett MD, Fajkus J, Leitch AR (2003) Telomere variability in the monocotyledonous plant order Asparagales. Proceedings of the Royal Society of London, Series B, Biological Sciences 270(1527): 1893–1904. https://doi.org/10.1098/rspb.2003.2446

Talukder K, Sen S (1999) *In situ* cytophotometric estimation of nuclear DNA in different cultivars of *Allium* species. Cytobios 99(390): 57–65.

Talukder K, Sen S (2000) Chromosome characteristics in some *Allium* sp. and assessment of their interrelationship. The Nucleus 43(1, 2): 46–57.

Tanaka R, Taniguchi K (1975) A banding method for plant chromosomes. Japanese Journal of Genetics 50: 163–167. https://doi.org/10.1266/jjg.50.163

Tang H, Lihua M, Shiqimg A, Jianquaan L (2005) Origin of the Qinghai-Tibetan Plateau endemic *Milula* (Liliaceae): further insights from karyological comparisons with *Allium*. Caryologia 58(4): 320–331. https://doi.org/10.1080/00087114.2005.10589470

Tardif B, Morisset P (1992) Relation between numbers of B-chromosomes and C-bands in *Allium schoenoprasum* L. Cytologia 57: 349–352. https://doi.org/10.1508/cytologia.57.349

Ulrich I, Fritz B, Ulrich W (1988) Application of DNA fluorochromes for flow cytometric DNA analysis of plant protoplasts. Plant Science 55: 151–158. https://doi.org/10.1016/0168-9452(88)90171-9

Vakhtina LI, Zakirova RO, Vakhtin YB (1977) Interspecific differences in DNA contents and taxonomically valid characters in *Allium* L. (Liliaceae). Botanicheskii Zhurnal (Moscow & Leningrad) 62: 667–684.

Van-Lume B, Esposito T, Diniz-Filho JAF, Gagnon E, Lewis GP, Souza G (2017) Heterochromatic and cytomolecular diversification in the *Caesalpinia* group (Leguminosae): relationships between phylogenetic and cytogeographical data. Perspectives in Plant Ecology, Evolution and Systematics 29: 51–63. https://doi.org/10.1016/j.ppees.2017.11.004

Van't Hof J (1965) Relationships between mitotic cycle duration, S period duration and the average rate of DNA synthesis in the root meristem cells of several plants. Experimental Cell Research 39: 48–58. https://doi.org/10.1016/0014-4827(65)90006-6

Ved Brat S (1965) Genetic systems in *Allium*. I. Chromosome variation. Chromosoma 16: 486–499. https://doi.org/10.1007/BF00343176

Ved Brat S (1967) Genetic systems in *Allium* IV. Balance in hybrids. Heredity 22: 387–396. https://doi.org/10.1038/hdy.1967.48

Ved Brat S, Dhingra B (1973) Genetic systems in *Allium* V. Breakdown of classical system in *Allium cepa*. The Nucleus 16: 11–19.

Vitales D, D'Ambrosio U, Galvez F, Kovařík A, Garcia S (2017) Third release of the plant rDNA database with updated content and information on telomere composition and sequenced plant genomes. Plant Systematics and Evolution 303(8): 1115–1121. https://doi.org/10.1007/s00606-017-1440-9

von Bothmer R (1975) The *Allium ampeloprasum* complex on Crete. Mitteilungen (aus) der Botanischen Staatssammlung München 12: 267–288.

Vujošević M, Jovanović V, Blagojević J (2013) Polyploidy and B chromosomes in *Allium flavum* from Serbia. Archives of Biological Sciences 65(1): 23–32. https://doi.org/10.2298/ABS1301023V

Wajahatullah MK, Vahidy AA (1990) Karyotyping and localization of nucleolar organizer regions in garlic, *Allium sativum* L. Cytologia 55: 501–504. https://doi.org/10.1508/cytologia.55.501

Walters ZW (1992) Rapid nuclear DNA content estimation for *Allium* spp. using flowcytometry. Allium Improvement Newsletter 2: 4–6.

Wang C, Zheng G (1987) The relationship between the intercellular chromatin migration of pollen mother cells and the changes of chromosome numbers during the genesis of male gametes in *Allium cepa*. Acta Botanica Sinica 29: 247–252.

Weiss H, Scherthan H (2002) *Aloe* spp. – plants with vertebrate-like telomeric sequences. Chromosome Research 10(2): 155–164. https://doi.org/10.1023/A:1014905319557

Wufeng D, Jinqiao X, Xin X (1993) Studies on karyotypes of four Chinese scallions (*Allium chinensis* G. Don). Journal of Wuhan Botanical Research 11: 199–203.

Xie-Kui C, Ao C, Zhang Q, Chen L, Liu J (2008) Diploid and tetraploid distribution of *Allium przewalskianum* Regel. (Liliaceae) in the Qinghai-Tibetan Plateau and adjacent regions. Caryologia 61(2): 192–200. https://doi.org/10.1080/00087114.2008.10589629

Xu J, Yang L, He X, Xue P (1998) A study on karyotype differentiation of *Allium fasciculatum* (Liliaceae). Acta Phytotaxonomica Sinica 36(4): 346–352.

Xue CY, Xu JM, Liu JQ (2000) Karyotypes of nine populations of *Allium przewalskianum* from Qinghai. Acta Botanica Yunnanica 22: 148–154.

Yang L, Xu J, Zhang X, Wan H (1998) Karyotypical studies of six species on the genus *Allium*. Acta Phytotaxonomica Sinica 36(1): 36–46.

Yuzbasioglu D (2004) Karyotyping, C- and NOR banding of *Allium sativum* L. (Liliaceae) cultivated in Turkey. Pakistan Journal of Botany 36(2): 343–349.

Zakirova RO, Nafanailova II (1988) Chromosome numbers in some species of the Kazakhstan flora. Botanicheskii Zhurnal (Moscow & Leningrad) 73: 1493–1494.

Zhukova PG (1967) Karyology of some plants cultivated in the Arctic-Alpine Botanical Garden. In: Avrorin NA (Ed.) Plantarum in Zonam Polarem Transportatio II. Leningrad, 139–149.

ORCID

Biplab Kumar Bhowmick https://orcid.org/0000-0001-6029-1098
Sayantika Sarkar https://orcid.org/0000-0002-8738-9500
Dipasree Roychowdhury https://orcid.org/0000-0001-7537-4056
Sayali D. Patil https://orcid.org/0000-0000-0000-00000
Manoj M. Lekhak https://orcid.org/0000-0001-5753-2225
Deepak Ohri https://orcid.org/0000-0001-6327-4330
Satyawada Rama Rao https://orcid.org/0000-0003-0309-720X
S. R. Yadav https://orcid.org/0000-0001-6728-5483
R. C. Verma https://orcid.org/0000-0000-0000-00000
Manoj K. Dhar https://orcid.org/0000-0002-8777-6244
S. N. Raina https://orcid.org/0000-0002-4916-3359
Sumita Jha https://orcid.org/0000-0002-1375-2768

SHORT COMMUNICATION

COMPARATIVE
Cytogenetics
A peer-reviewed open-access journal
International Journal of Plant & Animal Cytogenetics,
Karyosystematics, and Molecular Systematics

Karyotype of *Sabanejewia bulgarica* (Drensky, 1928) (Teleostei, Cobitidae) from the Danube Delta, Romania

Eva Hnátková[1], Zuzana Majtánová[2],
Vendula Bohlen Šlechtová[2], Joerg Bohlen[2], Petr Ráb[2]

1 *Department of Mathematics, Faculty of Engineering, Czech University of Life Sciences, 165 00 Prague, Kamýcká 129, Czech Republic* **2** *Laboratory of Fish Genetics, Institute of Animal Physiology and Genetics, Academy of Sciences of Czech Republic, 277 21 Liběchov, Czech Republic*

Corresponding author: Eva Hnátková (hnatkova@tf.czu.cz)

Citation: Hnátková E, Majtánová Z, Bohlen Šlechtová V, Bohlen J, Ráb P (2023) Karyotype of *Sabanejewia bulgarica* (Drensky, 1928) (Teleostei, Cobitidae) from the Danube Delta, Romania. Comparative Cytogenetics 17: 157–162. https://doi.org/10.3897/compcytogen.17.103152

Abstract

The karyotype of the freshwater fish *Sabanejewia bulgarica* (Drensky, 1928), from the Danube Delta, was studied by conventional Giemsa staining and the C-banding technique. The diploid chromosome number was 2n = 50. The karyotype contained 2 pairs of metacentric (the first one was much larger than the second one), 6 pairs of submetacentric and 17 pairs of subtelocentric to acrocentric chromosomes. Pericentromeric blocks of heterochromatin were revealed in most of the chromosome pairs. The karyotype phenotype of *S. bulgarica* was the same as found for *S. balcanica* from Northern Carpathian Mountains.

Keywords

C-heterochromatin, Chromosome number, cobitoid loaches, karyotype structure

Introduction

Freshwater fishes of the genus *Sabanejewia* Vladykov, 1929 are small (max 15 cm TL), have a benthic lifestyle in river habitats and can be distinguished from all other genera of Cobitidae by a specific sexual dimorphism (males with lateral body swellings) (Nalbant 1994; Kottelat and Freyhof 2007; Šlechtová et al. 2008). Their distribution spans from northern Italy to the Aral Sea basin including tributaries of the Black Sea, Caspian Sea,

Baltic Sea, Aral Sea and Mediterranean Sea (Kottelat and Freyhof 2007). The highest taxonomic diversity was long time assumed to be in the Danube basin, but phylogenetic studies (Perdices and Doadrio 2001; Šlechtová et al. 2008; Vasil'eva et al. 2022) have shown that several phenotypes within this area are very closely related (*Sabanejewia balcanica* (Karaman, 1922) species complex), so that the exact species composition is still not known.

Sabanejewia bulgarica (Drensky, 1928) was originally described as a species of the genus *Cobitis* Linnaeus, 1758 occurring in the lowest Danube basin including the Danube River itself. Junior synonyms are *C. albicoloris* Chichkoff, 1932 and *C. taenia tesselatus* Pietschman, 1937 (Kottelat 2012). For some time, most authors included *S. bulgarica* in the polytypic Balkan golden loach complex, with *Sabanejewia balcanica*, as a subspecies (Nalbant 1957; Bănărescu 1964; Bănărescu et al. 1972). Vasil'eva and Vasil'ev (1988) analysed the geographical variations of characters among a number of local forms of golden loach throughout its range they showed that *bulgarica* is quite distinct from all other populations and deserves species rank (Nalbant 1994; Kottelat 1997, 2012). The main differences include the overall colour of body, character of Gambetta´s zone, pigmentation and character of spots on the end of the caudal peduncle and basis of caudal fin as well as the habitat, as it is the only deep-water riverine form of *Sabanejewia*. Recently, Križek et al. (2020) based on several lines of evidence and analysis of individuals from type localities of both *S. balcanica* and *S. bulgarica sensu* Bănărescu et al. (1972) have claimed that Danubian golden loaches are genetically closer to *S. bulgarica* and *S. balcanica* itself is to be restricted to the drainage of the Vardar River.

Chromosomes of the loaches of the genus *Sabanejewia* remain poorly studied. Ráb et al. (1991) summarized three previous studies dealing with *S. larvata* (De Filippi, 1859), *S. caspia* (Eichwald, 1838) and *S. kubanica* Vasil'eva et Vasil'ev, 1988. Vasil'ev and Vasil'eva (1994, 2019) reported karyotypes of *S. caspia* and *S. aurata* (De Filippi, 1863) from other locations (Kizylach Bay of the Caspian Sea, Kura River in Tbilisi and southern basin of the Bug River of the Black Sea, respectively). The karyotype of *S. baltica* Witkowski, 1994 from the Bug River in Poland was reported by Boroń (2000). All these studies document invariable 2n = 50 and remarkably small karyotype variability among the loaches of this genus. To contribute to the cytotaxonomy of *Sabanejewia*, the present report describes the karyotype and C-banding pattern of *S. bulgarica* from Danube Delta, Romania. Additionally, the comparison of karyotypes of all karyologically studied species of *Sabanejewia* was conducted.

Materials and methods

The five *Sabanejewia bulgarica* specimens examined (all females) were collected in the Saint George Branch, Danube Delta (July 1997). The loach individuals arrived at the laboratory (Danube Delta Institute, Tulcea) in very bad condition and died subsequently. One female displayed the colour pattern and high body as described in original description by Drensky (1928), while four other individuals displayed other colour patterns (Fig. 1).

Standard procedure for chromosome preparation followed Ráb and Roth (1988) and C-banding technique followed Haaf and Schmid (1984). Chromosomes were clas-
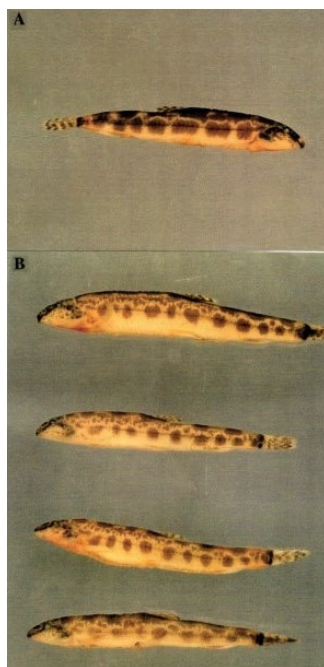
Figure 1. Analysed females of *Sabanejewia bulgarica* from the Saint George Branch, Danube Delta **A** individual with colour pattern and high body as described in original description by Drensky (1928) **B** individuals of *S. bulgarica* with other colour patterns.

sified according to the system of Levan et al. (1964). The analysed individuals are deposited as vouchers in fish collection of Laboratory of Fish Genetics, IAPG, AS CR, Liběchov (Accs. code. SB 7/97).

Results and discussion

The diploid chromosome number was determined as 2n = 50 in all five individuals. The detailed karyotype analysis was carried out on a single specimen with morphological characters corresponding to the original description (Fig. 1, top individual). The karyotype comprised 2 pairs of metacentric (m), 6 pairs of submetacentric (sm) and 17 pairs of subtelocentric (st) to acrocentric (a) chromosomes (Fig. 2). C-banding procedure reveals conspicuous pericentromeric blocks of heterochromatin in most of the chromosome pairs with very prominent blocks in the largest m chromosome and the first sm element.

Our results enable us to compare the karyotype of *Sabanejewia bulgarica* with karyotypes of other species of this genus analysed so far (Table 1). It is evident that the karyotype of *S. bulgarica* shows some apparent similarities with those of congeneric species: i) the same diploid chromosome number 2n = 50, ii) the low number of m chromosomes (two pairs), iii) the numbers of sm chromosomes (six pairs) and st to a chromosomes (17 pairs), iv) likely C-banding pattern correspond to those
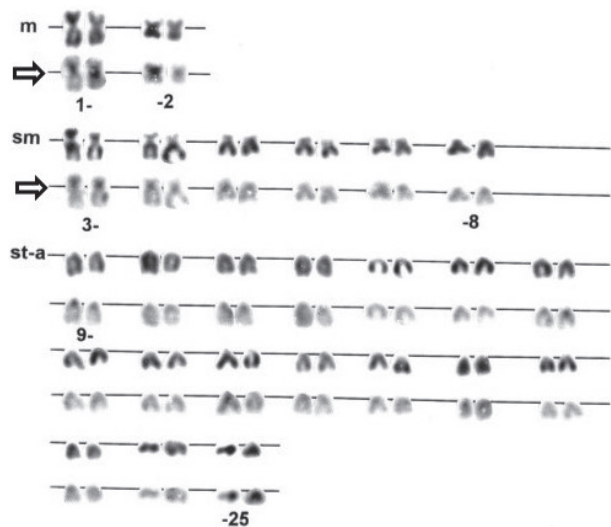
Figure 2. Karyotype of female *Sabanejewia bulgarica* arranged from conventionally Giemsa-stained (the first row) and sequentially C-banded (the second row, large blocks denoted by arrows) chromosomes. m – metacentric, sm – submetacentric, st – subtelocentric and a – acrocentric chromosomes.

Table 1. Diploid chromosome numbers (2n) and karyotype structure of karyologically studied species of genus *Sabanejewia*. Types of chromosomes: m –metacentric, sm – submetacentric, st – subtelocentric, a – acrocentric.

Species	2n	Karyotype characteristics				References
		m	sm	st	a	
S. larvata	50	2	3	11	9	Lodi and Marchioni 1980
S. caspia	50	2	3	11	9	Vasil'ev 1985; Vasil'ev and Vasil'eva 1994
S. kubanica	50	3	7	15		Vasil'eva and Vasil'ev 1988
S. aurata aurata	50	3	6	16		Vasil'ev and Vasil'eva 1994
S. (aurata) balcanica	50	2	6	17		Ráb et al. 1991
S. (aurata) balcanica	50	2	6	17		Vasil'ev and Vasil'eva 1994
S. baltica	50	2	8	15		Boroń 2000
S. bulgarica	50	2	6	17		This study

of *S. balcanica* (Ráb et al. 1991). The comparison also demonstrates some differences in karyotype structure, which, however, need explanation. Lodi and Marchioni (1980) and Vasil'ev (1985) distinguished the categories of st and a chromosome while Vasil'eva and Vasil'ev (1988), Ráb et al. (1991) and Vasil'ev and Vasil'eva (1994, 2019) did not recognize these morphological types of chromosomes as different ones but combined them together. This is especially due to the presence of small acrocentric-like chromosomes with their centromere position ranging gradually from subterminal to nearly terminal which makes difficult the precise assignment of these chromosomes to particular morphological types. The possible interspecific differences in proportion of such st and a chromosomes are in addition masked by the effect of chromosome arms contraction during mitosis due to the effect of timing and dose

of colchicine treatment. Thus, it is difficult to interpret the karyotype descriptions of different authors.

In conclusion, the karyotype of *S. bulgarica* from the Danube Delta is very similar to karyotypes of other *Sabanejewia* species in respect of the chromosome number and morphological types of chromosomes, and is nearly identical with karyotype of *S. balcanica* from the Northern Carpathian range in Slovakia. Further detailed cyto-taxonomic as well as other genetic surveys of populations of *Sabanejewia* throughout its distribution will answer problems of systematics of Danubian golden loaches.

Acknowledgements

References

Bănărescu P (1964) Fauna Republicii Populare Romine. Pisces–Osteichthyes. Editura Academiei Republicii Populare Române, Bucureşti, 962 pp.

Bănărescu P, Nalbant T, Chelmu S (1972) Revision and geographical variation of *Sabanejewia aurata* in Romania and the origin of *S. bulgarica* and *S. romanica* (Pisces, Cobitidae). Annotationes Zoologicae et Botanicae 75: 1–49.

Boroń A (2000) Cytogenetic characterization of the loaches of the genera *Sabanejewia*, *Misgurnus* and *Barbatula* (Pisces, Cobitidae). Folia Zoologica 49(Suppl. 1): 37–44.

Drensky P (1928) Die Fische der Familie Cobitidae in Bulgarien (nach der Sammlung der Konigl. Natur-Historischen Museums in Sofia). Bulletin de l'Institut Royal des Sciences Naturelles Sofia 1: 156–181. [in Bulgarian, German summary]

Haaf T, Schmid M (1984) An early stage of ZW/ZZ sex chromosome differentiation in *Poecilia sphenops* var. *melanistica* (Poeciliidae, Cyprinodontiformes). Chromosoma 89: 37–41. https://doi.org/10.1007/BF00302348

Kottelat M (1997) European freshwater fishes. Biologia, Bratislava 52/Supplement 5: 1–271.

Kottelat M (2012) Conspectus Cobitidum: an inventory of the loaches of the world (Teleostei: Cypriniformes: Cobiotoidei). Raffles Bulletin of Zoology Supplement 26: 1–199.

Kottelat M, Freyhof J (2007) Handbook of European Freshwater Fishes. Cornol, Publications Kottelat, CH, 646 pp.

Križek P, Mendel J, Fedorčák J, Koščo J (2020) In the foothill zone – *Sabanejewia balcanica* (Karaman 1922), in the lowland zone – *Sabanejewia bulgarica* (Drensky, 1928): Myth or reality? Ecology and Evolution 10(14): 7929–7947. https://doi.org/10.1002/ece3.6529

Levan A, Fredga K, Sandberg AA (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201–220. https://doi.org/10.1111/j.1601-5223.1964.tb01953.x

Lodi E, Marchioni V (1980) Chromosome complement of the masked loach *Sabanejewia larvata* (De Fil.) (Pisces, Osteichthyes). Caryologia 33: 435–440. https://doi.org/10.1080/00087114.1980.10796857

Nalbant TT (1957) *Cobitis aurata vallachica*, eine neue Unterart des Balkan-Steinpeitzgers (Pisces, Cobitidae). Senckenbergiana biologica 38: 209–232.

Nalbant TT (1994) Studies on loaches (Pisces: Ostariophysi: Cobitidae). I. An evaluation of the valid genera of Cobitinae. Travaux du Museum National d'Histoire Naturelle „Grigore Antipa" 34: 375–380.

Perdices A, Doadrio I (2001) The molecular systematics and biogeography of the European cobitids based on mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 19: 468–478. https://doi.org/10.1006/mpev.2000.0900

Pietschman V (1937) Eine neue Lokalform von *Cobitis taenia* L. (Cyprinoidea), *Cobitis taenia tesselatus*. Anzeiger der Akademie der Wissenschaften in Wien 74: 29–30.

Ráb P, Roth P (1988) Cold-blooded vertebrates. In: Balíček P, Forejt J, Rubeš J (Eds) Methods of Chromosome Analysis, Czechoslovak Biological Society Publishers, Brno, 115–124. [in Czech]

Ráb P, Roth P, Vasil'eva ED (1991) Chromosome banding study of the golden loach *Sabanejewia aurata balcanica* from Slovakia (Cobitidae). Japanese Journal of Ichthyology 38: 141–146. https://doi.org/10.1007/BF02905539

Šlechtová V, Bohlen J, Perdices A (2008) Molecular phylogeny of the freshwater fish family Cobitidae (Cypriniformes: Teleostei): Delimitation of genera, mitochondrial introgression and evolution of sexual dimorphism. Molecular Phylogenetics and Evolution 47: 812–831. https://doi.org/10.1016/j.ympev.2007.12.018

Vasil'ev VP (1985) Evolutionary Karyology of Fishes. Nauka Publishers, Moscow, 300 pp. [in Russian]

Vasil'ev VP, Vasil'eva ED (1994) The karyological diversity in spined loaches from genera *Cobitis* and *Sabanejewia*. Abstract VIII Congress of Societatis Europaea Ichthyologorum "Fishes and their Environment", Oviedo, Spain, 26 Sept.–2 Oct. 1994, 67 pp.

Vasil'eva ED, Vasil'ev VP (1988) Studies in intraspecific structure of *Sabanejewia aurata* (Cobitidae) with the description of new subspecies *S. aurata kubanica* subsp. nov. -Voprosy Ichtiologii 28: 192–212. [in Russian]

Vasil'eva ED, Vasil'ev VP (2019) Caspian spined loach *Sabanejewia caspia*: well-known but practically unexplored species of the Cobitidae family: pecularities of morphology, karyotype, distribution, and postulated phylogenetic links. Journal of Ichthyology 59(2): 144–159. https://doi.org/10.1134/S0032945219020206

Vladykov V (1925) Über eine neue *Cobitis*-art aus der Tschechoslowakei: *Cobitis montana* sp. n. Zoologische Jahrbucher, Abteilung für Systematik, Geographie und Biologie der Tiere 50: 320–338.

ORCID

Eva Hnátková https://orcid.org/0000-0002-0892-9557

RESEARCH ARTICLE

Complete chloroplast genome sequence of *Rhododendron mariesii* and comparative genomics of related species in the family Ericaeae

Zhiliang Li[1], Zhiwei Huang[1], Xuchun Wan[1], Jiaojun Yu[1], Hongjin Dong[1], Jialiang Zhang[1], Chunyu Zhang[1,2], Shuzhen Wang[1]

1 *College of Biology and Agricultural Resources, Huanggang Normal University, Huanggang, 438000, Hubei Province, China* **2** *College of Plant Science & Technology, Huazhong Agricultural University, Wuhan, 430070, Hubei Province, China*

Corresponding authors: Shuzhen Wang (wangshzhen710@whu.edu.cn); Chunyu Zhang (zhchy@mail.hzau.edu.cn)

Abstract

Rhododendron mariesii Hemsley et Wilson, 1907, a typical member of the family Ericaeae, possesses valuable medicinal and horticultural properties. In this research, the complete chloroplast (cp) genome of *R. mariesii* was sequenced and assembled, which proved to be a typical quadripartite structure with the length of 203,480 bp. In particular, the lengths of the large single copy region (LSC), small single copy region (SSC), and inverted repeat regions (IR) were 113,715 bp, 7,953 bp, and 40,918 bp, respectively. Among the 151 unique genes, 98 were protein-coding genes, 8 were tRNA genes, and 45 were rRNA genes. The structural characteristics of the *R. mariesii* cp genome was similar to other angiosperms. Leucine was the most representative amino acid, while cysteine was the lowest representative. Totally, 30 codons showed obvious codon usage bias, and most were A/U-ending codons. Six highly variable regions were observed, such as *trnK-pafI* and *atpE-rpoB*, which could serve as potential markers for future barcoding and phylogenetic research of *R. mariesii* species. Coding regions were more conserved than non-coding regions. Expansion and contraction in the IR region might be the main length variation in *R. mariesii* and related Ericaeae species. Maximum-likelihood (ML) phylogenetic analysis revealed that *R. mariesii* was relatively closed to the *R. simsii* Planchon, 1853 and *R. pulchrum* Sweet,1831. This research will supply rich genetic resource for *R. mariesii* and related species of the Ericaeae.

Keywords

chloroplast genome, comparative genomics, conservation genetics, phylogeny, *Rhododendron mariesii*

Introduction

Rhododendron mariesii Hemsley et Wilson, 1907, a typical member of the family Ericaeae, is mainly distributed in central China (Wang et al. 2018). Well known for leaf shape and bright-colored flowers, *R. mariesii* possesses valuable medicinal and horticultural properties (Wang et al. 2018). The deciduous species *R. mariesii* attracted great interest of *Rhododendron* breeders and geneticists. Furthermore, *R. mariesii* is very important in the woodland flora of the Dabie Mountains, and plays critical roles in ecological balance (Wang et al. 2018). Recently, *Rhododendron*-based ecological tourism, habitat fragmentation, and human activities have exerted significant effects towards natural growth of the wild *Rhododendron* population (Wang et al. 2019). Therefore, the research on population genetics and ecological conservation of wild *R. mariesii* is vital and necessary. However, limited genome information is available for *R. mariesii*, which has largely hindered corresponding genetic and molecular research.

In higher plants, the majority of plastomes are circular and quadripartite architecture consisting of two inverted repeat regions (IRa and IRb), a large single-copy region (LSC), and a small single-copy region (SSC) (Daniell et al. 2016; Hu et al. 2020; Abdullah et al. 2021). As maternally inherited organelle, the angiosperm plastome has a relatively conserved gene content and stable structure, which offers genetic markers sufficient for genome-wide evolutionary investigation at various taxonomic levels (Asaf et al. 2016; Zhang et al. 2017; Givnish et al. 2018). In plants, the size of cp genomes varied from 107 to 280 kb, containing approximately 130 genes related to photo synthesis and carbon fixation (Daniell et al. 2016; Rossini et al. 2021). The substitution rate of cp genome is lower than that of the nuclear genome, and 115–165 kb in cp genome is highly evolutionarily conserved (Smith 2015). However, specific genes exhibit accelerated evolution rates, such as *ycf1*, *matK*, and *rbcL*, which often serve as DNA barcoding (Dong et al. 2015; Wambugu et al. 2015; Zhang et al. 2017).

Next generation sequencing (NGS) has greatly increased the availability of genome data for non-species model, which facilitates the comparative cp genomics and phylogenetic studies at interspecific level (Santos and Almeida 2019; Pervez et al. 2022). In this research, the cpDNA of *R. mariesii* was assembled and annotated, SSR loci were characterized, comparative genomics and phylogenetic studies were also performed, hoping to benefit the studies of population evolution and conservation genetics of *R. mariesii* and related species.

Material and methods

Materials sampling and DNA extraction

Young and disease-free leaves of wild *R. mariesii* were sampled from the Dabie Mountains (central China, 29°16.13'N, 115°27.07'E, 1,005 m), dried in silica, and stored at -20 °C until further usage. In particular, sample collection was authorized by the

Biodiversity Conservation of Huanggang Normal University. The specimens were identified by Hongjin Dong (Huanggang Normal University), who possesses a doctoral degree in botany. All materials were well conserved in the Huanggang Normal University Herbarium (Hubei province, China). Total genomic DNA was extracted and purified from fresh leaves according to Wang et al. (2019). Subsequently, the quality of total genome DNA was verified in 1% agarose gel stained by GelRed and quantified by spectophotometer (NanoDrop 1000, Thermofisher Scientific, USA).

Genome sequencing, assembly, and annotation

Nextera DNA library preparation kit was used to construct the paired-end Illumina libraries. These libraries were sequenced on Illumina NovaSeq6000 Sequencing System (Illumina, Hayward, CA) in a paired-end run (500 cycles, 1 × 250 pb). After trimming adapter sequences and removing low-quality sequences, raw data was filtered by SOAPnuke software (Chen et al. 2018). Then, the high-quality reads were *de novo* assembled by GetOrganelle pipeline (Jin et al. 2020). BOWTIE2 were used to validate the assembled sequence error of *R. mariesii* cp genome through mapping raw sequencing reads to the assembled plastome (Hanussek et al. 2021). Online program Organelle Genome DRAW (OGDRAW) was used to draw the physical map of *R. mariesii*cp genome (Greiner et al. 2019). Furthermore, gene annotation and analysis were carried out with DOGMA and C$_p$GAVAS softwares, respectively (Liu et al. 2012). The final annotations were also manually verified by Geneious (ver.8.0.2) (Yu et al. 2022). The cp genome data had been submitted to the National Center for Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov/).

Codon usage and nucleotide diversity analysis

Codon usage frequency was analyzed by CodonW software (https://sourceforge.net/projects/codonw/). Particularly, all protein coding genes were used for analysis. Relative synonymous codon usage (RSCU) analysis was carried out to measure codon usage bias (Rossini et al. 2021). The RSCU referred to the ratio of observed frequency of codons to frequency expected in regarding to the equal usage of synonymous codons for a certain amino acid (Rossini et al. 2021). In particular, RSCU value more than 1 means a preferred codon, otherwise the value less than and equal to 1 are considered as no codon usage bias (Morton 2022).

In total, eleven full chloroplast genomes of genus *Rhododendron* were downloaded from NCBI database: *R. molle* Siebold et Zuccarini, 1846; *R. griersonianum* Balfour filius et Forrest, 1919; *R. pulchrum* (Sweet) George Don, 1834; *R. henanense* Fang, 1983; *R. micranthum* Maximowicz, 1870; *R. delavayi* Franchet, 1886; *R. concinnum* Hemsley, 1890; *R. simsii* Planchon, 1876; *R. platypodum* Diels, 1990; *R. datiandingense* Feng, 1996; and *R. kawakamii* Hayata, 1911. Unique genes of these ten downloaded and the newly assemble *R. mariesii* cp genomes were extracted with PHYLOSUITE v1.2.2 and aligned by Windows version of MAFFT software, then nucleotide diversity (Pi) was calculated for each unique gene with DNASP ver 6.12.03 (Rossini et al. 2021).

Simple sequence repeats (SSR) analysis

MISA software (MicroSAtellite identification tool v2.1, http://pgrc.ipk-gatersleben.de/misa) was used to identify SSR motifs. Minimum number of tandem repeat units were set as follows: five repeat units for tri-, tetra-, penta-, and hexanucleotide SSRs; six repeat units for di-nucleotide SSRs; 10 repeat units for mono-nucleotide SSRs. The maximal number of bases interrupting two SSRs in a compound microsatellite was 100 bp.

Phylogenetic analysis

Through searing NCBI database, 21 cp genomes of Ericaceae species were found and downloaded: 12 species of *Rhododendron*; two species of *Vaccinium* Linnaeus, 1753; *Arbutus unedo* Sims, 1822; *Hemitomes congestum* Asa Gray, 1858; *Allotropa virgata* Torrey et Gray, 1868, *Monotropa hypopitys* Linnaeus, 1753; *Pityopus californicus* (Eastwood) H.F.Copeland, 1935; and 2 species of *Gaultheria* Kalm, 1753. Together with the newly assembled *R. mariesii* cp genome, these 22 cp genomes were used to construct phylogeny tree. These cp genomes were initially aligned with MAFFT for phylogenetic analysis (Yu et al. 2020). RAxML (version 8.2.8 for Windows) was used to run maximum likelihood (ML) analysis with a bootstrap value of 1000 (Alexandros 2014). FIGTREE v1.4 was used to visualize and adjust the ML trees (Yu et al. 2022). In particular, cp genome of *Pyrola rotundifolia* Benth. (1840) played the roles of an out-group.

Comparative analysis of genome structure

The structural characteristics of cp genomes, containing newly assembled *R. mariesii* and 10 cp genomes of the genus *Rhododendron* (*R. delavayi*, *R. henanense*, *R. micranthum*, *R. concinnum*, *R. griersonianum*, *R. simsii*, *R. kawakamii*, *R. molle*, *R. platypodum*, and *R. datiandingense*) were compared and analyzed with mVISTA online tool (using Shuffle-LAGAN alignment program). In particular, the annotated cp genome of *R. mariesii* served as a reference against the other cp genome. Genome alignments, including rearrangements or inversions, was detected with MAUVE (Darling et al. 2004). For investigating whether expansion or contraction occurred in *R. mariesii* cp genome, the IR/LSC and IR/SSC junction regions were compared with IRscope software (Amiryousefi et al. 2018).

Results

General features of *R. mariesii* chloroplast genome

In total, 19,498,900 reads were obtained from NovaSeq paired-end run. After stringent quality assessment and filtering, 19,309,162 clean reads (2.891 Gb) with an average of 149 bp read length were obtained. The percentage of clean reads was 99.03%, and the clean bases were 2,891,089,781 bp. In particular, GC content was 39.52%.

In addition, Q20 (a base with quality value greater than 20) and Q30 (a base with quality value greater than 20) values were 97.28% and 92.34%, respectively. The size of *R. mariesii* cp genome is 203,480 bp. Moreover, typical quadripartite structure was observed, as a large single-copy (LSC) region (113,715 bp) and a small single-copy (SSC) region (7,953 bp) were separated by a pair identical inverted repeat regions (IRs) (40,918 bp) (Fig. 1).

In total, 151 genes were successfully annotated, including 98 protein-coding genes, 45 tRNA genes, and 8 rRNA genes. The lengths of CDS, rRNA, tRNA, intergenic regions, and introns were 65,889 bp (32.38%), 8,998 bp (4.42%), 3,449 bp (1.7%),
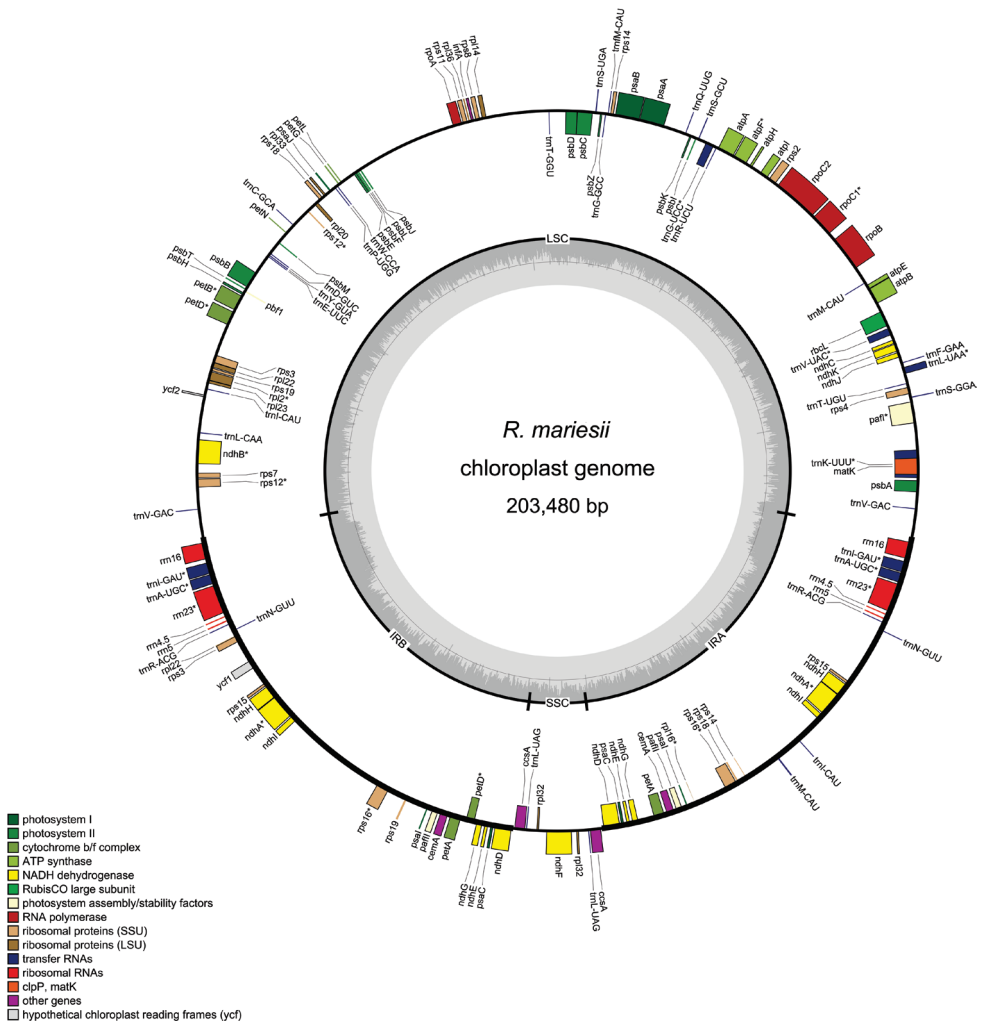
Figure 1. The chloroplast genome map of *R. mariesii*. Thick lines represented LSC, SSC, and IR regions, respectively. Genes shown inside circle were transcribed counterclockwise, and the outside outer circle were transcribed clockwise. Different gene groups were represented by different colors.

45,409 bp (22.32%), and 80,033 bp (39.33%), respectively. The GC content of CDS, rRNA, tRNA, intergenic regions, and intron were 37.67%, 54.87%, 51.49%, 32.06%, and 33.75%, respectively. A set of 55 photosynthesis-related genes were found, containing six subunits of ATP synthase (*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, and *atpI*), seven subunits of photosystem I, 17 subunits of photosystem II, 17 subunits of NADH-dehydrogenase, seven subunits of cytochrome b/f complex, and one subunit of rubisco (*rbcL*) (Table 1). Considering the self replication, 11 genes were large subunits of ribosome, four genes were DNA-dependent RNA polymerase, one gene was translational initiation factor, eight genes were rRNA genes, 45 genes were tRNA genes, and 18 genes were small subunit of ribosome (Table 1). The other genes were related to acetyl-CoA carboxylase, c-type cytochrome synthesis gene, envelope membrane protein, and maturase (Table 1). In addition, there were three conserved open reading frames, including one *ycf3* and two *ycf4*. Totally, 16 genes contained introns, containing *trnK-UUU*, *ycf3*, *trnL-UAA*, *trnV-UAC*, *ropB*, *atpF*, *trnS-CGA*, *accD*, *rpl16*, *ndhB*, *trnE-UUC*, *trnA-UGC*, *ndhA*, *trnA-UGC*, and *trnE-UUC* (Table 2). Besides *ycf3* and *accD* genes (three exons and two introns), the other 14 genes all had two exons and one intron.

Table 1. Gene content of *R. mariesii* chloroplast genome. The duplicated genes were included into brackets.

Category of genes	Group of genes	Name of genes
Genes for photosynthesis	Subunits of ATP synthase	*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *atpI*
	Subunits of photosystem I	*psaA*, *psaB*, *psaC* (2×), *psaI* (2×), *psaJ*
	Subunits of photosystem II	*psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ* (3×), *psbK*, *psbL*, *psbM*, *psbN*, *psbT*, *psbZ*
	Subunits of NADH-dehydrogenase	*ndhA* (2×), *ndhB*, *ndhC*, *ndhD* (2×), *ndhE* (2×), *ndhF*, *ndhG* (2×), *ndhH* (2×), *ndhI* (2×), *ndhJ*, *ndhK*
	Subunits of cytochrome b/f complex	*petA* (2×), *petB*, *petD*, *petG*, *petL*, *petN*
	Subunit of rubisco	*rbcL*
Self replication	Large subunit of ribosome	*rpl2*, *rpl14*, *rpl16*, *rpl20*, *rpl22* (3×), *rpl32*(2×), *rpl33*, *rpl36*
	DNA dependent RNA polymerase	*rpoA*, *rpoB*, *rpoC1*, *rpoC2*
	Translational initiation factor	*infA*
	Ribosomal RNA genes	*rrn5S* (2×), *rrn16S* (4×), *rrn23S* (2×)
	Transfer RNA genes	*trnK-UUU*, *trnH-GUC*, *trnS-GGA*, *trnT-UGU* , *trnT-GGU*, *trnL-UAA*, *trnL-CAA*, *trnL-UAG* (2×), *trnM-CAU* (12×), *trnF-GAA*, *trnV-UAC*, *trnV-GAC* (2×), *trnR-UCU*, *trnR-ACG* (2×), *trnS-CGA*, *trnS-GCU*, *trnS-UGA*, *trnQ-UUG*, *trnW-CCA*, *trnP-UGG*, *trnC-GCA*, *trnD-GUC*, *trnY-GUA*, *trnE-UUC* (2×), *trnA-UGC* (2×), *trnN-GUU*, *trnA-UGC*, *trnE-UUC* (2×)
	Small subunit of ribosome	*rps2*, *rps3* (3×), *rps4*, *rps7*, *rps8*, *rps11*, *rps14*, *rps15* (3×), *rps16*, *rps18* (2×), *rps19* (3×)
Other genes	Subunit of acetyl-CoA-carboxylase	*accD*
	c-type cytochrom synthesis gene	*ccsA* (2×)
	Envelop membrane protein	*cemA* (2×)
	Maturase	*matK*
Unkown function	Conserved open reading frames	*ycf3*, *ycf4* (2×)

Table 2. The characteristics list of genes possessing introns.

Gene	Strand	Start	End	ExonI	IntronI	ExonII	IntronII	ExonIII
trnK-UUU	-	1,834	4,404	37	2499	35		
ycf3	-	6,794	8,753	124	711	232	742	151
trnL-UAA	+	11,313	11,909	35	512	50		
trnV-UAC	-	15,031	15,692	39	588	35		
rpoB	+	21,836	25,719	3,169	677	38		
atpF	+	35,554	36,820	161	700	406		
trnS-CGA	-	38,853	39,609	31	666	60		
accD	+	55,356	56,894	571	159	150	54	605
rpl16	-	59,253	167,784	9	108,121	402		
ndhB	-	101,387	103,550	721	685	758		
trnE-UUC	+	112,521	113,535	32	943	40		
trnA-UGC	+	113,600	114,490	37	818	36		
ndhA	+	126,079	128,272	563	1090	541		
ndhA	-	181,938	184,131	563	1090	541		
trnA-UGC	-	195,720	196,610	37	818	36		
trnE-UUC	-	196,675	197,689	32	943	40		

Codon usage analysis and nucleotide diversity analysis

In the *R. mariesii* chloroplast genome, the protein-coding regions presented 40,013 codons (Table 3). Particularly, leucine (Leu) was the main amino acid (10.477%), followed by isoleucine (Ile, 8.972%) and glycine (Gly, 7.148%) (Fig. 2). In particular, cysteine (Cys) and tryptophan (Trp) were the lowest representative amino acids, accounting for 1.180% and 1.869%, respectively. According to RSCU values, a total of 30 codons showed obvious codon usage bias, as RSCU value were more than 1 (Table 3). Except Leu codon (UUG), all the other 29 codons were A/U-ending. For the 34 codons with RSCU values less than 1, 31 were C/G-ending, while 3 were A/U-ending.

Nucleotide diversity analysis showed that sequence level of divergence existed between different *Rhododendron* cp genomes. Pi values for each gene region varied from 0 to 0.06896. High level of genetic variation mainly existed in SSC region (Pi =
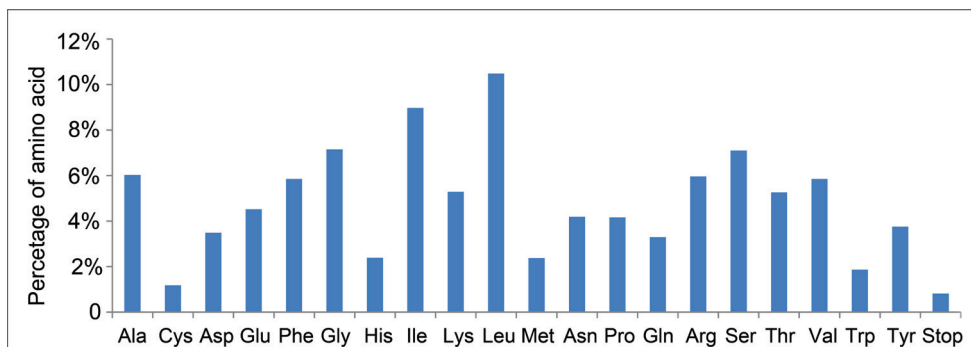
Figure 2. Occurrence percentage of amino acids in *R. mariesii* chloroplast genome.

Table 3. The relative synonymous codon usage in *R. mariesii* cp genome.

Amino acid	Codon	No	RSCU	The codon frequency per amino acid(%)	Amino acid	Codon	No	RSCU	The codon frequency per amino acid(%)
Ala	GCA	703	1.17	29.16	Pro	CCA	509	1.22	30.55
	GCC	345	0.57	14.31		CCC	294	0.71	17.65
	GCG	275	0.46	11.41		CCG	202	0.48	12.12
	GCU	1088	1.81	45.13		CCU	661	1.59	39.67
Cys	UGC	118	0.5	24.99	Gln	CAA	1053	1.59	79.65
	UGU	354	1.5	74.97		CAG	269	0.41	20.35
Asp	GAC	273	0.39	19.57	Arg	AGA	653	1.64	27.37
	GAU	1122	1.61	80.44		AGG	179	0.45	7.5
Glu	GAA	1395	1.54	77.2		CGA	626	1.57	26.24
	GAG	412	0.46	22.8		CGC	147	0.37	6.16
Phe	UUC	753	0.64	32.19		CGG	158	0.4	6.62
	UUU	1586	1.36	67.8		CGU	623	1.57	26.11
Gly	GGA	1079	1.51	37.73	Ser	AGC	188	0.4	6.61
	GGC	327	0.46	11.43		AGU	580	1.22	20.41
	GGG	465	0.65	16.26		UCA	507	1.07	17.84
	GGU	989	1.38	34.58		UCC	415	0.88	14.6
His	CAC	223	0.47	23.28		UCG	240	0.51	8.44
	CAU	735	1.53	76.73		UCU	912	1.93	32.09
Ile	AUA	1125	0.94	31.34	Thr	ACA	640	1.22	30.39
	AUC	674	0.56	18.78		ACC	411	0.78	19.52
	AUU	1791	1.5	49.89		ACG	213	0.4	10.11
Lys	AAA	1631	1.54	77.11		ACU	842	1.6	39.98
	AAG	484	0.46	22.88	Val	GUA	845	1.44	36.11
Leu	CUA	518	0.74	12.36		GUC	302	0.52	12.91
	CUC	251	0.36	5.99		GUG	319	0.55	13.63
	CUG	248	0.35	5.92		GUU	874	1.49	37.35
	CUU	889	1.27	21.21	Trp	UGG	748	1	100.02
	UUA	1475	2.11	35.18	Tyr	UAC	306	0.41	20.32
	UUG	811	1.16	19.35		UAU	1200	1.59	79.68
Met	AUG	954	1	100.01	Stop*	UAA	133	1.22	40.69
Asn	AAC	382	0.46	22.78		UAG	88	0.81	26.92
	AAU	1295	1.54	77.22		UGA	106	0.97	32.42

0.01723), followed by LSC (Pi = 0.00697) and IR (Pi = 0.001224) regions (Fig. 3). In total, six gene regions showed high levels of nucleotide diversity (Pi > 0.02), containing *trnI-GAU* (Pi = 0.06896), *trnG-UCC* (Pi = 0.06721), *rps3* (Pi = 0.04509), *rps12* (Pi = 0.03947), *trnV-UAC* (Pi = 0.03622), and *trnK-UUU* (Pi = 0.02554).

SSR analysis of *R. mariesii* plastome

A set of 70 SSRs were identified from *R. mariesii* cp genome, and 5 SSRs were present in compound formation. Particularly, 65 SSRs (92.86%) were mononucleotide motifs, 2 were dinucleotide motifs (2.86%), 2 were trinucleotide motifs (2.86%), and 1 were hexanucleotide repeats (1.43%) (Table 4). Dominant mononucleotide repeats were
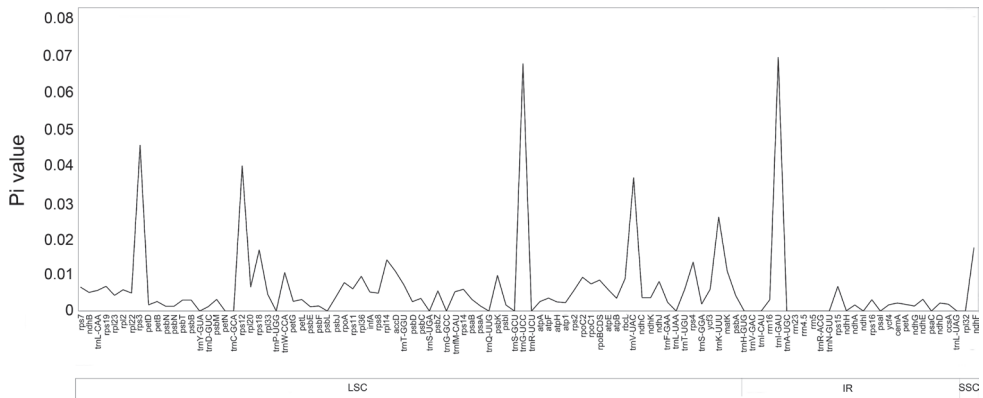
Figure 3. The nucleotide diversity (Pi) of 11 *Rhododendron* species chloroplast genomes. X-axis presented the position of aligned chloroplast genomes, and Y-axis referred to Pi value. Below the X-axis, large single-copy (LSC), small single-copy (SSC), as well as inverted repeat (IR) regions were displayed with arrow bars.

Table 4. The frequency of each type of microsatellite in *R. mariesii* cp genome.

Repeats	5	6	7	8	9	10	11	12	13	14	15	total	Percentage
A/T	-	-	-	-	-	32	14	9	7	2		64	91.429%
C/G	-	-	-	-	-						1	1	1.429%
AT/AT	-		2									2	2.857%
AAG/CTT	1											1	1.429%
AAT/ATT							1					1	1.429%
AAGGGT/ACCCTT	1											1	1.429%

A/T (91.429%), while C/G repeats accounted for 1.429%. In related to dinucleotide motifs, only AT/AT type was found (2.857%). For trinucleotide motifs, only one (AAG/CTT)5 and one (AAT/ATT) 11 motif were found. For hexanucleotide repeats, one (AAGGGT/ACCCTT)5 was found, accounting for 1.429%.

Mononucleotide A/T repeats with repeat numbers of 10–14 were the most abundant. Meanwhile, (C/G)n microsatellites were all repeated 15 times. In relation to dinucleotide repeats, the identified SSRs all have 7 repeat motifs. Regarding to trinucleotide motifs, AAG/CTT and AAT/ATT microsatellites repeated 5 and 11 times, respectively. The hexanucleotide motif AAGGGT/ACCCTT repeated 5 times. Totally, 34 SSRs were present in the intergenic spacer region, accounting for 41.43%. Moreover, 28 SSRs were present in *rpl16* gene. All the remaining 13 microsatellites were found in *ccsA*, *cemA*, *ndhA*, *rpoA*, *rpoC2*, *rps7*, *rps8*, and *trnL-UAA* genes.

Phylogenetic analysis

For clarifying the phylogenetic location of *R. mariesii* among the Ericaeae, complete plastomes of *R. mariesii* and other 21 species in the Ericaeae with fully sequenced

chloroplast genomes were used in reconstructing phylogenetic relationships. The phylogenetic tree revealed that *R. mariesii* had a close genetic relationship with *R. simsii* and *R. pulchrum* (Fig. 4). In particular, all these 22 taxa belonging to Ericaeae were grouped into one clade and clustered into two subclades. Topological structure was almost consistent with the previously published phylogeny (Liu et al. 2021). *A. unedo* (NCBI: JQ067650), *H. congestum* (NCBI: NC_035581), *A. virgate* (NCBI: NC_035580), *M. hypopitys* (NCBI: NC_029704), *P. californicus* (NCBI: NC_035584), *Vaccinium oldhamii* Miquel, 1866 (NCBI: MK049537), *Vaccinium macrocarpon* Aiton, 1789 (NCBI: JQ757046), *Gaultheria griffithiana* Wight, 1847 (NCBI: MW528025), *Gaultheria fragrantissima* Wallich, 1820 (NCBI: MW563322), and *P. rotundifolia* (NCBI: KU833271), were relatively distant related with *R. mariesii*.

Comparative plastome sequence divergence and hotspot regions

Structural characteristics of 11 *Rhododendron* cp genomes were investigated with mVISTA software, containing the newly assembled *R. mariesii* cp genome and 10 download cp genome of the *Rhododendron* genus. In particular, the annotated *R. mariesii* cp genome served as a reference. Relatively high similarity was detected among
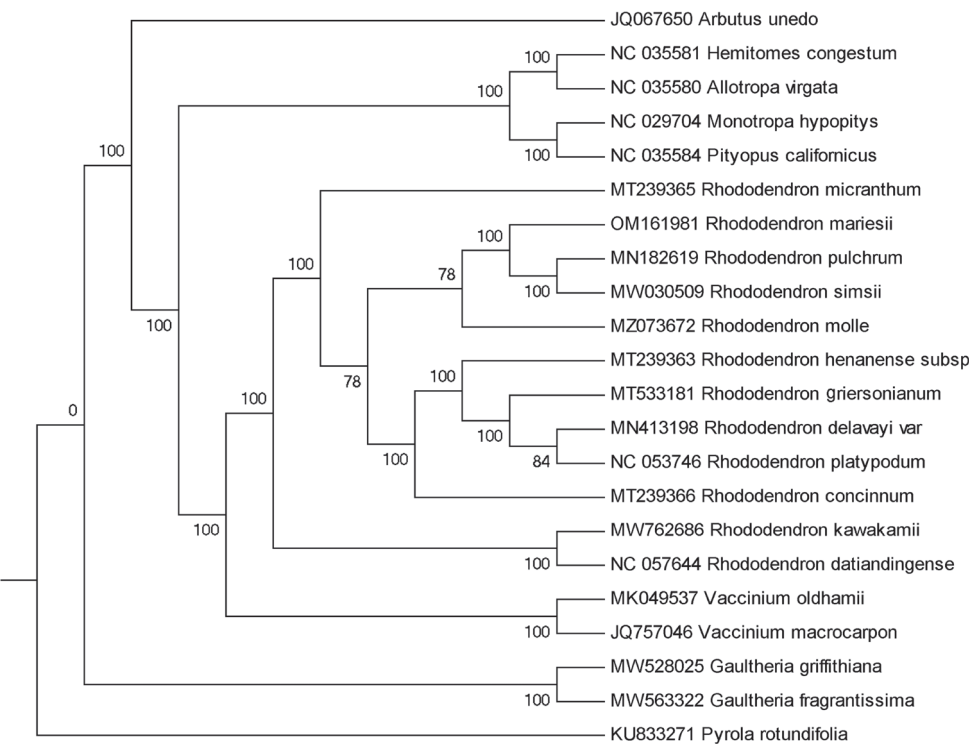
Figure 4. The maximum-likelihood phylogenetic tree for *R. mariesii*. Numbers on each node referred to bootstrap support value.

these 11 *Rhododendron* species. Coding regions were more conserved than non-coding regions (CNS in Fig. 5). The LSC and SSC regions were relatively more stable than IR regions. Among these coding regions, *rpoB*, *rpoC2*, *rps8*, *petD*, *rpl23*, *rpl22*, and *ndhF* were relatively divergent because of intron regions. In *R. mariesii* plastid genome, highly variable regions mostly existed in the intergenic spacer, such as *trnK-pafI*, *atpE-rpoB*, *trnT-rpl14*, *rpoA-psbJ*, *rpl20-trnE*, *ndhI-rps19*, and *rpl16-ndhI*. Compared with intergenic spacer, protein coding regions were highly conserved, such as *rps4*, *ndhJ*, *ndhK*, *rpcL*, rps2, atpI, *psaA*, *psaB*, *psbB*, *cemA*, and *petA*. No rearrangements and inversions occurred in these 11 cp genomes of *Rhododendron* species.

Particularly, lengths of the IR regions of 6 cp genomes ranged from 14,194 bp (*R. mariesii* cp genome) to 47,467 (*R. griersonianum* cp genome) (Fig. 6). Expansion and contraction existed in these cp genomes. In *R. griersonianum* cp genome, JLB line (line between LSC and IRb) was located between genes *ycf15* and *trnR*, while *ycf15* was located in the LSC region with 166 bp extending to IRb region. In *R. mariesii* cp genome, JLB line was located between *trnV* and *rrn16* (476 bp extending to LSC region). In *R. micranthum* and *R. henanense* cp genomes, the JLB lines were located between
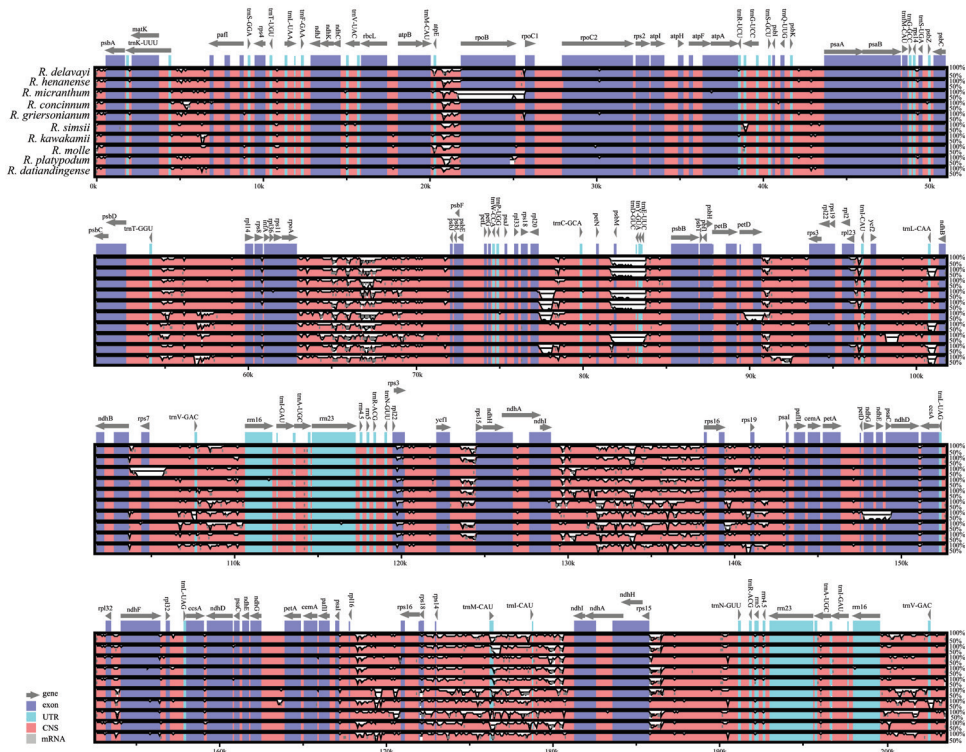
Figure 5. Comparison of cp genomes with *R. mariesii* annotation serving as the reference. Vertical scale indicated the percentage of identity (50–100%), and horizontal axis was coordinates within cp genome. The genome regions were color-coded as exons, introns, and conserved non-coding sequences, respectively.
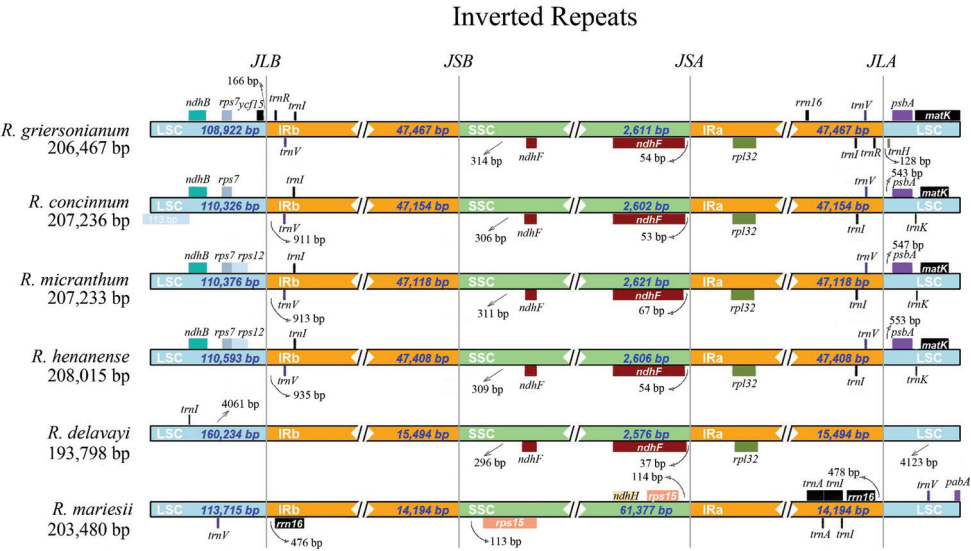
Figure 6. The comparison of LSC, SSC, and IR regional boundaries of cp genome between *R. mariesii* and related taxa. JLB, JSB, JSA, and JLA respected "junction line between LSC and IRb", "junction line between IRb and SSC", "junction line between SSC and IRa" , as well as "junction line between IRa and LSC", respectively.

rps12 and *trnV*, while *trnV* was located in IRb region with 913 bp and 935 bp extending to LSC region, respectively. However, JLB line was located between *rps7* and *trnI*, and *trnI* was located in IRb region with 911 bp extending to LSC region in *R. concinnum*. Except for *R. mariesii*, *ndhF* was located in SSC region with 296 bp–314 bp extending to IRb region in the five other cp genomes (Fig. 6). Meanwhile, *rps15* was located in SSC region of *R. mariesii* cp genome. The JSA line (line between SSC and IRa) was located between *ndhF* and *rpl32*, and *ndhF* was distributed in SSC region with 54 bp, 53 bp, 67 bp, 54 bp, and 37 bp to IRa region in *R. griersonianum*, *R. concinnum*, *R. micranthum*, *R. henanense*, and *R. delavayi* cp genomes, respectively. Besides *ndhH*, *rps15* was also located in SSC region with 114 bp extending to IRa region in *R. mariesii* cp genome. Furthermore, JLA line (line between IRa and LSC) was located between *trnV* and *psbA* in *R. concinnum*, *R. micranthum*, and *R. henanense*. However, JLA line was located between *trnR* and *trnH* in *R. griersonianum* cp genome. In *R. mariesii* cp genome, *rrn16* and *trnV* were located besides the JLA line.

Discussion

The chloroplast genome is the main organelle for plant transforming light energy into chemical energy (Zhang et al. 2018). Plastome genome is useful for comparative genomic research and phylogenomic analyses due to polymorphic regions generated through genomic expansion, inversion, contraction, and gene rearrangement (Sanitá

Lima et al. 2016; Kahraman et al. 2019; Li et al. 2019; Li et al. 2020; Liu et al. 2020; Wang et al. 2020). The single circular cp genome structure of *R. mariesii* was the same as other species belonging to the Ericaeae with a typical quadripartite structure and similar GC content unevenly distributed across the cp genome (Liu et al. 2021; Xu et al. 2022). The GC content of *R. mariesii* cp genome (39.52%) was slightly larger than that of *Myracrodruon urundeuva* Allemão, 1862 (37.8%) (Rossini et al. 2021). Relative to both LSC (35.85%) and SSC (36.49%) regions, the GC content in IR region (30.48%) is lower. However, the GC content in IR region was larger than LSC and SSC regions in cp genome *Xanthium spinosum* Linnaeus, 1753 (Raman et al. 2020). The length of total genome size and each region were similar to other plant cp genomes, such as *R. molle*, *Rubus* species (Rosaceae), and rubber dandelion (*Taraxacum kok-saghyz* Rodin) (Zhang et al. 2017; Xu et al. 2022; Yu et al. 2022).

The size of *R. mariesii* cp genome (203,480 bp) was larger than that of *R. pulchrum* (146,941 bp), *R. simsii* (152,214 bp), *R. molle* (197,877 bp), *R. delavayi* (193,798 bp), and *R. platypodum* (201,047 bp), but smaller than that of *R. kawakamii* (230,777 bp), *R. micranthum* (207,233 bp), *R. henanense* (208,015 bp), *R. griersonianum* (206,467 bp), *R. concinnum* (207,236 bp), and *R. datiandingense* (207,311 bp). Totally, 151 genes existed in *R. mariesii* cp genome, which were more than that of *R. molle* (149 genes) and *R. pulchrum* (73 genes) (Shen et al. 2020; Xu et al. 2022). In particular, protein-coding genes accounted for 64.901% in *R. mariesii* cp genome, which were lower than that of *R. molle* cp genome (65.101%) but higher than that of *R. pulchrum* cp genome (57.534%) (Shen et al. 2020; Xu et al. 2022). Moreover, 45, 44, and 29 tRNA genes were found in cp genomes of *R. mariesii*, *R. molle*, and *R. pulchrum*, respectively. In both *R. mariesii* and *R. molle* cp genomes, eight rRNA genes were annotated, but only two were found in *R. pulchrum* cp genome (Shen et al. 2020; Xu et al. 2022).

Besides genes involved in photosynthesis transforming light energy into chemical energy, other genes also existed in *R. mariesii* cp genome. For example, *accD* gene, encoding plastid beta carboxyl transferase subunit of acetyl-CoA carboxylase (ACCase) important for plant growth (leaf growth, leaf longevity, fatty acid biosynthesis, and embryo development), has been reported to be involved in the adaptation to specific ecological niches during radiation of dicotyledonous plants (Hu et al. 2015). In *R. mariesii* cp genome, one copy of *accD* gene was also found. Codons coding Leu (10.477%), Ile (8.972%), and Gly (7.148%) were dominant, while Cys (1.180%) and Trp (1.869%) were the least, which were the same as that of *M. urundeuva* cp genome (Rossini et al. 2021). Codon bias, an efficient mechanism of translation influenced by natural selection and mutation pressure, takes place if synonymous codons are used at different frequencies (Zhang et al. 2022). A total of 30 codons showed codon usage bias, and most were A/U-ending codons, which were the same as that observed in *M. urundeuva* and *Solanum* (Zhang et al. 2018a; Rossini et al. 2021). In total, six gene regions showed high levels of nucleotide diversity (Pi > 0.02), containing *trnI-GAU*, *trnG-UCC*, *rps3*, *rps12*, *trnV-UAC*, and *trnK-UUU*, serving as the first candidate for developing molecular markers to identify *Rhododendron* species.

A total of 70 SSRs were identified from *R. mariesii* cp genome, more than that of *M. urundeuva* (36 SSRs), *Spondias bahiensis* P. Carvalho, 2015 (53 SSRs) and *Mangifera indica* Wallich, 1847 (57 SSRs), but fewer than that of *Syringa pinnatifolias* Hemsley, 1906 (253 SSRs) (Jo et al. 2017; Santos and Almeida 2019). Variation in the number and type of microsatellites might play important roles in plastome organization. The main motifs were A/T repeats (91.429%), which was the same with that of *M. urundeuva*, *S. bahiensis*, and *M. indica* (Jo et al. 2017; Santos and Almeida 2019; Rossini et al. 2021). However, no correlation was found between large repeat regions and rearrangement endpoints, which was similar with Liu et al. (2013). Very limited tandem (G/C)n-containing microsatellites were observed, which might be due to the low content of G and C bases in chloroplast genome. Molecular markers developed for the intergenic regions could be used for phylogenetic, phylogeographic, and barcoding studies of *Rhododendron* species.

Non-coding regions often mutate relatively faster than coding regions (Yu et al. 2022). In *R. mariesii* cp genome, coding regions were more conserved than the non-coding regions. Relatively high similarity was detected among these *Rhododendron* cp genomes, but expansion and contraction also existed in IR regions, which might be the dominant reason for variation in cp genome size. Obvious differences were found in cp IR boundary regions, containing gene contents and locations. However, IR regions were least divergent, which were mainly due to the presence of four highly conserved rRNA sequences in *X. spinosum* (Raman et al. 2020). Furthermore, LSC and SSC regions were relatively more stable than IR regions in *R. mariesii* cp genome. These genetic variations may significantly facilitate *R. mariesii* adapting to the changes of survival conditions. According to neutral theory, nucleotide substitution in non-coding regions (intergenic spacer, intron region, and pseudogenes) are neutral or near-neutral, which could not be affected by natural selection (Akashi et al. 2012). Therefore, evolutionary history of *R. mariesii* could be well calculated from the rate of molecular evolution in non-coding region.

This research aimed to expand the molecular genetic resources available for *R. mariesii* through high-throughput sequencing and cp genome assembly. The *R. mariesii* cp genome sequence could be used in distinguishing and resolving phylogenetic relationships within Ericaeae lineage. Moreover, this research will be vital for further genetic analysis on *R. mariesii* and other species in the Ericaeae family.

Conflicts of interests

The authors declare that they have no competing interests.

Data availability statement

The cp genome of *R. mariesii* was submitted to GenBank database under the accession number of OM161981.

Acknowledgements

References

Abdullah, Henriquez CL, Mehmood F, Hayat A, Sammad A, Waseem S, Waheed MT, Matthews PJ, Croat TB, Poczai P, Ahmed I (2021) Chloroplast genome evolution in the *Dracunculus* clade (Aroideae, Araceae). Genomics 113(1): 183–192. https://doi.org/10.1016/j.ygeno.2020.12.016

Akashi H, Osada N, Ohta T (2012) Weak selection and protein evolution. Genetics 192: 15–31. https://doi.org/10.1534/genetics.112.140178

Alexandros S (2014) Raxml version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics (9): 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Amiryousefi A, Hyvönen J, Poczai P (2018) IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics 34: 3030–3031. https://doi.org/10.1093/bioinformatics/bty220

Asaf S, Khan AL, Khan AR, Waqas M, Kang SM, Khan MA, Lee SM, Lee IJ (2016) Complete chloroplast genome of *Nicotiana otophora* and its comparison with related species. Frontiers in Plant Science 7: 843. https://doi.org/10.3389/fpls.2016.00843

Daniell H, Lin CS, Yu M, Chang WJ (2016) Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biology 17: 134. https://doi.org/10.1186/s13059-016-1004-2

Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Research 14: 1394–1403. https://doi.org/10.1101/gr.2289704

Dong W, Xu C, Cheng T, Lin K, Zhou S (2013) Sequencing angiosperm plastidgenomes made easy: a complete set of universal primers and a case study on thephylogeny of Saxifragales. Genome Biology and Evolution 5: 989–997. https://doi.org/10.1093/gbe/evt063

Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, Cheng T, Guo J, Zhou S (2015) ycf1, the most promising plastid DNA barcode of land plants. Scientific Reports 5: 8348. https://doi.org/10.1038/srep08348

Greiner S, Lehwark P, Bock R (2019) Organellar genome DRAW(OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Research 47(W1): W59–W64. https://doi.org/10.1093/nar/gkz238

Givnish TJ, Zuluaga A, Spalink D, Soto Gomez M, Lam VKY, Saarela JM, Sass C, Iles WJD, de Sousa DJL, Leebens-Mack J, Chris Pires J, Zomlefer WB, Gandolfo MA, Davis JI, Stevenson DW, de Pamphilis C, Specht CD, Graham SW, Barrett CF, Ané C (2018) Monocot

plastid phylogenomics, timeline, net rates of species diversification, the power of multi-gene analyses, and a functional model for the origin of monocots. American Journal of Botany 105: 1888–1910. https://doi.org/10.1002/ajb2.1178

Hanussek M, Bartusch F, Krüger J (2021) Performance and scaling behavior of bioinformatic applications in virtualization environments to create awareness for the efficient use of compute resources. PLoS Computational Biology 17(7): e1009244. https://doi.org/10.1371/journal.pcbi.1009244

Hu G, Cheng L, Huang W, Cao Q, Zhou L, Jia W, Lan Y (2020) Chloroplast genomes of seven species of Coryloideae (Betulaceae): structures and comparative analysis. Genome 63(7): 1–12. https://doi.org/10.1139/gen-2019-0153

Hu S, Sablok G, Wang B, Qu D, Barbaro E, Viola R, Li M, Varotto C (2015) Plastome organization and evolution of chloroplast genes in *Cardamine* species adapted to contrasting habitats. BMC Genomics 16: 306. https://doi.org/10.1186/s12864-015-1498-0

Jin JJ, Yu WB, Yang JB, Song Y, de Pamphilis CW, Yi TS, Li DZ (2020) GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology 21(1): 241. https://doi.org/10.1186/s13059-020-02154-5

Jo S, Kim HW, Kim YK, Sohn JY, Cheon SH, Kim KJ (2017) The complete plastome sequences of *Mangifera indica* L. (Anacardiaceae). Mitochondrial DNA Part B-Resources 2(2): 698–700. https://doi.org/10.1080/23802359.2017.1390407

Kahraman K, Lucas SJ (2019) Comparison of different annotation tools for characterization of the complete chloroplast genome of *Corylus avellana* cv Tombul. BMC Genomics 20(1): 874. https://doi.org/10.1186/s12864-019-6253-5

Li X, Zuo Y, Zhu X, Liao S, Ma J (2019) Complete chloroplast genomes and comparative analysis of sequences evolution among seven *Aristolochia* (Aristolochiaceae) medicinal species. International Journal of Molecular Sciences 20(5): 1045. https://doi.org/10.3390/ijms20051045

Li CJ, Wang RN, Li DZ (2020) Comparative analysis of plastid genomes within the Campanulaceae and phylogenetic implications. PLoS ONE 15(5): e0233167. https://doi.org/10.1371/journal.pone.0233167

Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X (2012) CpGAVAS, an integrated web server for the annotation, visualization, analysis, and genbank submission of completely sequenced chloroplast genome sequences. BMC Genomics 13(1): 715. https://doi.org/10.1186/1471-2164-13-715

Liu Q, Li X, Li M, Xu W, Schwarzacher T, Heslop-Harrison JS (2020) Comparative chloroplast genome analyses of *Avena*: insights into evolutionary dynamics and phylogeny. BMC Plant Biology 20: 406. https://doi.org/10.1186/s12870-020-02621-y

Liu Y, Li Q, Wang L, Wu L, Huang Y, Zhang J, Song Y, Liao J (2021) The complete chloroplast genome of *Rhododendron molle* and its phylogenetic positionwithin Ericaceae. Mitochondrial DNA Part B-Resources 6(9): 2587–2588. https://doi.org/10.1080/23802359.2021.1959458

Liu Y, Huo N, Dong L, Wang Y, Zhang S, Young HA, Feng X, Gu YQ (2013) Complete chloroplast genome sequences of Mongolia medicine *Artemisia frigida* and phylogenetic relationships with other plants. PLoS ONE 8(2): e57533. https://doi.org/10.1371/journal.pone.0057533

Morton BR (2022) Context-dependent mutation dynamics, not selection, explains the codon usage bias of most angiosperm chloroplast genes. Journal of Molecular Evolution 90: 17–29. https://doi.org/10.1007/s00239-021-10038-w

Pervez MT, Hasnain MJU, Abbas SH, Moustafa MF, Aslam N, Shah SSM (2022) A comprehensive review of performance of next-generation sequencing platforms. Biomed Research International 2022: 3457806. https://doi.org/10.1155/2022/3457806

Raman G, Park KT, Kim JH, Park S (2020) Characteristics of the completed chloroplast genome sequence ofxanthium spinosum: comparative analyses, identification of mutational hotspots and phylogenetic implications. BMC Genomics 21(1): 855. https://doi.org/10.1186/s12864-020-07219-0

Rossini BC, de Moraes MLT, Marino CL (2021) Complete chloroplast genome of *Myracrodruon urundeuva* and its phylogenetics relationships in Anacardiaceae family. Physiology and Molecular Biology of Plants 27(4): 801–814. https://doi.org/10.1007/s12298-021-00989-1

Santos V, Almeida C (2019) The complete chloroplast genome sequences of three *Spondias* species reveal close relationship among the species. Genetics and Molecular Biology 42(1): 132–138. https://doi.org/10.1590/1678-4685-gmb-2017-0265

Sanitá Lima M, Woods LC, Cartwright MW, Smith DR (2016) The (in)complete organelle genome: exploring the use and nonuse of available technologies for characterizing mitochondrial and plastid chromosomes. Molecular Ecology Resources 16: 1279–1286. https://doi.org/10.1111/1755-0998.12585

Shen J, Li X, Zhu X, Huang X, Jin S (2020) The complete plastid genome of *Rhododendron pulchrum* and comparative genetic analysis of Ericaceae species. Forests 11(2): 158. https://doi.org/10.3390/f11020158

Smith DR (2015) Mutation rates in plastid genomes: they are lower than you might think. Genome Biology and Evolution 7: 1227–1234. https://doi.org/10.1093/gbe/evv069

Wambugu PW, Brozynska M, Furtado A, Waters DL, Henry RJ (2015) Relationships of wild and domesticated rices (*Oryza* AA genome species) based upon whole chloroplast genome sequences. Scientific Reports 5: 13957. https://doi.org/10.1038/srep13957

Wang S, Li Z, Guo X, Fang Y, Xiang J, Jin W (2018) Comparative analysis of microsatellite, SNP, and Indel markers in four *Rhododendron* species based on RNA-seq. Breeding Science 68(5): 536–544. https://doi.org/10.1270/jsbbs.18092

Wang S, Jin Z, Luo Y, Li Z, Fang Y, Xiang J, Jin W (2019) Genetic diversity and population structure of *Rhododendron simsii* (Ericaceae) as reveled by microsatellite markers. Nordic Journal of Botany 37(4): 1–10. https://doi.org/10.1111/njb.02251

Wang X, Rhein HS, Jenkins J, Schmutz J, Grimwood J, Grauke LJ, Randall JJ (2020) Chloroplast genome sequences of *Carya illinoinensis* from two distinct geographic populations. Tree Genetics and Genomes 16(4): 48. https://doi.org/10.1007/s11295-020-01436-0

Yu J, Dong H, Xiang J, Xiao Y, Fang Y (2020) Complete chloroplast genome sequence and phylogenetic analysis of *Magnolia pilocarpa*, a highly ornamental species endemic in central China. Mitochondrial DNA Part B-Resources 5(1): 720–722. https://doi.org/10.1080/23802359.2020.1714511

Yu J, Fu J, Fang Y, Xiang J, Dong H (2022) Complete chloroplast genomes of *Rubus* species (Rosaceae) and comparative analysis within the genus. BMC Genomics 23: 32. https://doi.org/10.1186/s12864-021-08225-6

Zhang Y, Iaffaldano BJ, Zhuang X, Cardina J, Cornish K (2017) Chloroplast genome resources and molecular markers differentiate rubber dandelion species from weedy relatives. BMC Plant Biology 17: 34. https://doi.org/10.1186/s12870-016-0967-1

Zhang R, Zhang L, Wang W, Zhang Z, Du H, Qu Z, Li XQ, Xiang H (2018) Differences in codon usage bias between photosynthesis-related genes and genetic system-related genes of chloroplast genomes in cultivated and wild *Solanum* species. International Journal of Molecular Sciences19: 3142. https://doi.org/10.3390/ijms19103142

Zhang J, Huang H, Qu C, Meng X, Meng F, Yao X, Wu J, Guo X, Han B, Xing S (2021) Comprehensive analysis of chloroplast genome of *Albizia julibrissin* durazz. (Leguminosae sp.). Planta 255(1): 26. https://doi.org/10.1007/s00425-021-03812-z

Supplementary material 1

Taxonomic and accession information on cp genomes downloaded from NCBI database
Authors: Zhiliang Li, Zhiwei Huang, Xuchun Wan, Jiaojun Yu, Hongjin Dong, Jialiang Zhang, Chunyu Zhang, Shuzhen Wang
Data type: wps

Link: https://doi.org/10.3897/compcytogen.17.101427.suppl1

RESEARCH ARTICLE

A comparative cytogenetic study of *Hypsibarbus malcolmi* and *H. wetmorei* (Cyprinidae, Poropuntiini)

Sudarat Khensuwan[1], Weerayuth Supiwong[2], Chatmongkon Suwannapoom[3], Phichaya Buasriyot[4], Sitthisak Jantarat[5], Weera Thongnetr[6], Nawarat Muanglen[7], Puntivar Kaewmad[8], Pasakorn Saenjundaeng[2], Kriengkrai Seetapan[9], Thomas Liehr[10], Alongklod Tanomtong[1]

1 *Department of Biology, Faculty of Science, Khon Kaen University, Muang, Khon Kaen 40002, Thailand* **2** *Faculty of Interdisciplinary Studies, Khon Kaen University, Nong Khai Campus, Muang, Nong Khai 43000, Thailand* **3** *Department of Fishery, School of Agriculture and Natural Resources, University of Phayao, Muang, Phayao 56000, Thailand* **4** *Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi, Mueang Nonthaburi, Nonthaburi 11000, Thailand* **5** *Department of Science, Faculty of Science and Technology, Prince of Songkla University, Pattani 94000, Thailand* **6** *Division of Biology, Department of Science, Faculty of Science and Technology, Rajamangala University of Technology Krungthep, Bangkok 10120, Thailand* **7** *Department of Fisheries, Faculty of Agricultural Technology, Sakon Nakhon Rajabhat University, Sakon Nakhon 47000, Thailand* **8** *Faculty of Science and Technology, Mahasarakham Rajabhat University, Muang, Maha Sarakham 44000, Thailand* **9** *School of Agriculture and Natural Resources, University of Phayao, Tumbol Maeka, Muang District, Phayao Province, 56000 Thailand* **10** *Institute of Human Genetics, University Hospital Jena, Jena 07747, Germany*

Corresponding author: Weerayuth Supiwong (weersu@kku.ac.th)

Citation: Khensuwan S, Supiwong W, Suwannapoom C, Buasriyot P, Jantarat S, Thongnetr W, Muanglen N, Kaewmad P, Saenjundaeng P, Seetapan K, Liehr T, Tanomtong A (2023) A comparative cytogenetic study of *Hypsibarbus malcolmi* and *H. wetmorei* (Cyprinidae, Poropuntiini). Comparative Cytogenetics 17: 181–194. https://doi.org/10.3897/compcytogen.17.107703

Abstract

Cyprininae are a highly diversified but demonstrably monophyletic lineage of cypriniform fishes. Here, the karyotype and chromosomal characteristics of *Hypsibarbus malcolmi* (Smith, 1945) and *H. wetmorei* (Smith, 1931) were examined using conventional, nucleolus organizing regions (NORs) and molecular cytogenetic protocols. The diploid chromosome number (2n) of *H. malcolmi* was 50, the fundamental number (FN) was equal to 62, and the karyotype displayed 8m + 4sm + 38a with NORs located at the centromeric and telomeric positions of the short arms of chromosome pairs 1 and 2, respectively. 2n of *H. wetmorei* was 50, FN 78, karyotype 14m + 14sm + 22a with the NORs at the telomeric position of

the short arm of chromosome pair 2. 2n and FN in males and females were identical. Fluorescence *in situ* hybridization using different microsatellite motifs as probes also showed substantial genomic divergence between both studied species. In *H. wetmorei*, $(CAG)_n$ and $(CAC)_n$ microsatellites accumulated in the telomeric regions of all chromosomes, while in *H. malcolmi*, they had scattered signals on all chromosomes. Besides, the $(GAA)_n$ microsatellites were distributed along all chromosomes of *H. malcolmi*, but there was a strong hybridization pattern in the centromeric region of a single pair in *H. wetmorei*. These cytogenomic difference across the genomes of these *Hypsibarbus* Rainboth, 1996 species are markers for specific evolutionary differentiation within these two species.

Keywords
Fish cytogenetics, Cyprinidae, microsatellites , chromosomes

Introduction

The Cyprininae are the largest subfamily of the family Cyprinidae, which are the most diverse group of freshwater fish worldwide. This subfamily currently includes 33 genera, with 228 species being widely distributed in the freshwater systems of Eurasia (Fricke et al. 2023). *H. malcolmi* (Smith, 1945) and *H. wetmorei* (Smith, 1931), two yet understudied examples of Cyprininae, are widely distributed in Thailand's rivers Mekong, Songkhram, Chao Phraya and Sirindhorn peat swamp forest. The two species mentioned have been shown to be the most similar to each other in external morphology and coloration (Fig. 1) and may be considered a species complex (Rainboth 1996). In addition, these two species are placed in the tribe of Poropuntiini on the phylogenetic reconstruction proposed by Yang et al. (2015). The classification of these fishes has been extremely difficult (Nelson 1994; Rainboth 1996; Ruber et al. 2007; Fang et al. 2009; Yang et al. 2010).

The diploid chromosome number of *H. malcolmi* and *H. wetmorei* has been reported as 2n = 50, but the karyotype and NF of *H. malcolmi* seem to be different (Magtoon and Arai 1989; Piyapong 1999; Donsakul and Magtoon 2002; Donsakul et al. 2007; Chantapan 2015; Khensuwan et al. 2023). Cytogenetics has become an important tool for fish classification, including cyprinids (Yang et al. 2015). Hereby, an important characteristic is the localization of nucleolus organizer region(s) (NOR(s)) as an inter- and intraspecies-specific marker for cytotaxonomic studies; NORs have been used for studying phylogenetic relationships between Cyprinids (Amemiya and Gold 1988; Galetti Jr 1998; Almeida-Toledo et al. 2000).

Classical and molecular cytogenetics play a crucial role in elucidating evolutionary patterns in cyprinid fish, especially in cases when species exhibit conserved diploid numbers. The abundance and chromosomal location of specific repetitive DNAs (microsatellites) change significantly between genomes of closely related species, and these variations are generally species-specific (Pereira et al. 2013). For example, the dinucleotides (CA)15 and (GA)15 accumulated exclusively in telomeric and subtelomeric chromosomal regions, corroborating findings from other fish groups studied to date (Terencio et al. 2013; Xu et al. 2013; Yano et al. 2014; Oliveira et al. 2015, 2018; Pucci et al. 2016). Otherwise,

the genome of the wolf fish *Hoplias malabaricus* (Bloch, 1794), with 12 different microsatellite repeats ((A)30, (C)30, (CA)15, (GA)15, (GC)15, (CAC)10, (CAA)10, (CAG)10, (CAT)10, (GAG)10, (TAA)10 and (CGG)10) showed strong hybridization signals at subtelomeric and heterochromatic regions of several autosomes, with a varied amount of signal on the sex chromosomes (Cioffi et al. 2011). So, in our study using trinucleotides (CAG)10, (GAA)10 and (CAC)10 observed patterns in the dynamics of the *Hypsibarbus* Rainboth, 1996 genome. Such microsatellites are predominantly located in the heterochromatic regions (telomeres, centromeres and sex chromosomes) of fish chromosomes, where a significant fraction of repetitive DNA is localized (Yüksel and Gaffaroğlu 2008; Knytl et al. 2013; Saenjundaeng et al. 2018, 2020; Phimphan et al. 2020; Saenjundaeng et al. 2020; Wang et al. 2022; Kentachalee et al. 2023). Single short repeats (SSRs) are short motifs that are repeated across the genome and consist of one to six nucleotides (Cioffi and Bertollo 2012; López-Flores and Garrido Ramos 2012). By supporting the correct pairing of the DNA double strand and preventing replication errors such as the creation of loops or other structures, they contribute to the stability of DNA molecules (Schueler et al. 2001). Furthermore, repeated DNAs play an essential role in speciation, sex differentiation, and biodiversity (Vicari et al. 2005; Cioffi et al. 2009; Sember et al. 2018).

The present study includes in-depth cytogenetic analyses of *H. malcolmi* and *H. wetmorei* (not a hybrid), comprising conventional Giemsa- and Ag-NOR staining and fluorescence in situ hybridization (FISH) approaches with chromosomal mapping of several repetitive DNA classes (microsatellites).

Material and methods

Animals

Individuals of *H. malcolmi* (12♂ and 6♀) and *H. wetmorei* (8♂ and 8♀) were collected in the Mekong River basin (Thailand) (Fig. 1). Fish were transferred to the laboratory and identified according to the morphological criteria of Rainboth et al. (2012). Experiments were performed in accordance with ethical protocols, with anesthesia using clove oil (Eugenol 3%) prior to the euthanasia, as approved by the Ethics Committee of Khon Kaen University (Record No. IACUC-KKU-105/63). The specimens were deposited in the fish collection of the Cytogenetic Laboratory, Department of Biology, Faculty of Science, Khon Kaen University (Thailand). DOI: dx.doi.org/10.17504/protocols.io.36wgq3r8klk5/v1.

Chromosome preparation and NOR staining

Mitotic chromosomes were obtained from the anterior kidney following the drop onto microscopic slides and the air-dry method to visualize the chromosomes (Bertollo 2015). Conventional staining was performed using 5% Giemsa for 8 min (Khensuwan

et al. 2023). In addition, the distribution of NORs was visualized according to the standard protocol using silver (Ag) staining (Howell and Black 1980). The slides were then sealed with cover slips and incubated at 60 °C for 5 minutes. After that, they were soaked in distilled water until the cover slips were separated. The glass slides were stained with 5% Giemsa for 1 minute.

Probe preparation and FISH experiments

FISH experiments were performed under high stringency conditions (Yano et al. 2017) to classify microsatellite sequences, specifically (CAG)10, (GAA)10, and (CAC)10. These sequences were directly labeled by Cy3 at the 5'end during synthesis (Sigma, St. Louis, MO, USA) as described by Kubat et al. (2008). FISH was performed under stringent conditions and hybridization occurred overnight in a moist chamber at 37 °C (Sassi et al. 2023). Chromosomes were counterstained with 4',6-Diamidino-2-phe-nylindole dihydrochloride (DAPI, 1.2 µg/ml) mounted in antifade solution (Vector, Burlingame, CA, USA,).

Figure 1. Collection sites of *Hypsibarbus malcolmi* (1) and *H. wetmorei* (2) in the Mekong River, North-East Thailand (18°17'48.2"N, 104°00'16.9"E and 18°17'59.9"N, 104°00'09.5"E). Scale bar for fish: 1 cm.

Image processing

At least 20 metaphase spreads per individual were analyzed to confirm the diploid number, karyotype structure, NORs and FISH data. Images were captured using an Axioplan II microscope (Carl Zeiss Jena GmbH, Germany) with CoolSNAP and processed using Image Pro Plus 4.1 software (Media Cybernetics, Silver Spring, MD, USA). Chromosomes were classified according to centromere position as metacentric (m), submetacentric (sm) and acrocentric (a) (Levan et al. 1964). For the chromosomal arm number (FN; fundamental number) m+sm were scored as bi-armed while a as mono-armed.

Results

Chromosome number, karyotype and fundamental number

Cytogenetic analysis of *H. malcolmi* revealed 2n = 50 and FN = 62 in both sexes with a karyotype composed of 8 metacentric, 4 submetacentric and 38 acrocentric chromosomes (Fig. 2A). On the other hand, although *H. wetmorei* also showed 2n = 50, its FN was equal to 78, given its karyotype being composed of 14 metacentric, 14 submetacentric and 22 acrocentric chromosomes.
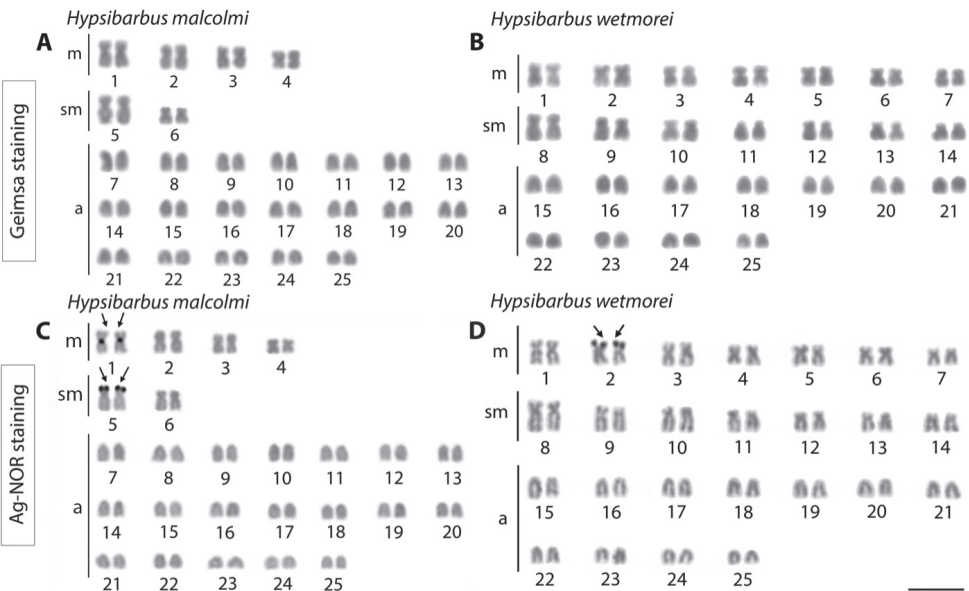
Figure 2. Karyotypes after conventional Giemsa (**A, B**) and NOR staining (arrows) (**C, D**) of *Hypsibarbus malcolmi*, 2n = 50 (**A, C**) and *H. wetmorei*, 2n = 50 (**B, D**). Scale bar: 5 μm.

Table 1. Available cytogenetic data for *Hypsibarbus* species.

Species	2n	FN	Karyotype	Locality	NORs site	References
Hypsibarbus lagleri Rainboth, 1996	50	74	4m + 20sm + 26a	Thailand	-	Donsakul et al. 2002
Hypsibarbus malcolmi (Smith, 1945)	50	64	10m + 4sm + 36a	Thailand	-	Donsakul et al. 2007
	50	62	8m + 4sm + 38a	Thailand	1, 5	Khensuwan et al. 2023
	50	62	8m + 4sm + 38a	Thailand	1, 5	Present study
Hypsibarbus vernayi (Norman, 1925)	50	58	6m + 2sm + 4st + 38a	Thailand	-	Donsakul et al. 2002
Hypsibarbus wetmorei (Smith, 1931)	50	70	12m + 8sm + 6st + 24a	Thailand	-	Magtoon and Arai 1989
	50	74	12m + 12sm + 4st + 22a	Thailand	2	Piyapong 1999
	50	74	12m + 12sm + 2st + 24a	Thailand	-	Donsakul et al. 2002
	50	82	10m + 14sm + 8st + 18a	Thailand	6	Chantapan 2015
	50	78	14m + 14sm + 22a	Thailand	2	Khensuwan et al. 2023
	50	78	14m + 14sm + 22a	Thailand	2	Present study

NOR- staining and FISH results

While *H. malcolmi* had two pairs of NOR-bearing chromosomes, *H. wetmorei* had only one such pair. In the first, Ag-NOR regions were located at the centromeric and telomeric positions of the short arms on metacentric pairs 1 and 5 (Fig. 2C), while in *H. wetmorei* they were restricted to the telomeres on the short arms of pair 2 (Fig. 2D).

In *H. wetmorei* the (CAG)n and (CAC)n microsatellites accumulated in the telomeric regions of all chromosomes, while *H. malcolmi* had scattered signals along all 50

Figure 3. Hybridization patterns with microsatellite probes $(CAG)_{10}$, $(GAA)_{10}$ and $(CAC)_{10}$ (red signals) on metaphase plates of *Hypsibarbus malcolmi* and *H. wetmorei*. Chromosomes were counterstained with DAPI (blue). Scale bar: 5 μm.

chromosomes. (GAA)n presented strong signals in the centromeric regions of a single chromosomal pair in *H. malcolmi*, but a scattered distribution among all chromosomes in *H. wetmorei* (Fig. 3).

Discussion

Our study has characterized populations of *H. malcolmi* and *H. wetmorei* by classical and molecular cytogenetics. For both species, diploid number and other features described in the scientific literature were confirmed (Magtoon and Arai 1989; Piyapong 1999; Donsakul and Magtoon 2002; Donsakul et al. 2007; Chantapan 2015; Khensuwan et al. 2023). 2n = 50 was reported for both species, as in the whole Cyprinidae lineage, which has been cytogenetically investigated so far and all of them exhibit a remarkable 2n conservation of 50 chromosomes. But with distinct karyotype organization in different species and populations. However, such a preserved 2n is clearly linked to substantial intrachromosomal changes, as also demonstrated by the discrepant NOR- and microsatellite patterns obtained in this study, emphasizing the importance of structural rearrangements in the evolution of this family, such as chromatin duplications/deletions, pericentric inversions, transpositions, and translocations (Pereira et al. 2011; Saenjundaeng et al. 2020; Khensuwan et al. 2023). Such rearrangements can also be observed as distinct patterns of NOR and microsatellite distribution among populations.

The position of NOR was consistent with the previous report for both species, with two pairs in *H. malcolmi* and one in *H. wetmorei*. The occurrence of multiple NORs in fish was considered to be apomorphic, whereas a single pair of NORs is considered to be plesiomorphic (Gold and Amemiya 1986). In species with multiple NORs, interindividual variation is common suggesting that transposition of rDNA genes from one chromosome to another may occur (reviewed in Phillips and Rab 2001). For example, *Salvelinus namaycush* (Walbaum, 1792) (Phillips and Ihssen 1989), *S. alpinus* (Linnaeus, 1758) (Phillips et al. 1988) and *Salmo trutta* Linnaeus, 1758 (Castro et al. 1994) found differences in the number of NORs. Previous investigations in Cyprinidae have shown that almost all NOR sites correspond to active 18S rDNA loci (Khensuwan et al. 2023). The 18S rDNA is clustered with the 5.8S and the 25S rDNAs in plants, although only the first composes the small subunit of ribosomes (Goffová and Fajkus 2021). In fish genomes, 18S rDNA is usually located at the terminal position on chromosomes (Sochorová et al. 2018). This was also observed in both studied *Hypsibarbus* species, in addition to a centromeric site at the first chromosome pair in *H. malcolmi*. Although it is known that the terminal position of this rDNA facilitates the arrangement of the NOR in the interphase nucleus, centromeric NORs were found in the karyotypes of several species (e.g. Barth et al. 2013; Sassi et al. 2021), including species that only harbor a single rDNA locus (e.g. Sing and Barman 2013). Indeed, the number of NORs presented in the genome varies by species, and the rDNA content of NORs can differ between individuals of the same species and even between cells within an individual (Stults et al. 2008, 2009). Because

ribosomal gene arrays are extremely repetitive, they are prone to homologous recombination (HR), creating unstable areas that could favor chromosomal rearrangements (Kobayashi 2008). The pattern observed in *H. malcolmi* could be a hint at a paracentric inversion of the short arm of chromosome 1. Normally, the NORs/18S rDNA is commonly found in a terminal location inside chromosomes (Sochorová et al. 2018), except for *H. malcolmi* located in the centromeric region. It is also remarkable that a large variety of karyotype re-organization occurs among populations.

The instability of repetitive regions of the genome can also be observed by microsatellites. These small repetitive motifs have been shown to stall and reverse replication forks, and to be hotspots of chromosomal double strand breaks in model organisms (Pelletier et al. 2003; Kerrest et al. 2009; reviewed in Gadgil et al. 2017). In fish, they are also accumulated in sex chromosomes (Schemberger et al. 2019). Closely related species can have very distinct patterns of microsatellite accumulation, as observed in the two species of *Hypsibarbus* here studied. Such discrepancy is more notable when comparing the (CAG)n, (GAA)n and (CAC)n motifs that are dispersed in the *H. malcolmi* genome but accumulate in the telomeres of *H. wetmorei*. Microsatellite motifs had a preferential accumulation in heterochromatic regions (reviewed in Cioffi and Bertollo,2012). However, the majority of the (CAG)n, (GAA)n and (CAC)n microsatellite sequences in *H. malcolmi* showed a scattered pattern on chromosomes, without a specific relationship with heterochromatic regions. Nevertheless, the (CAG)n, (GAA)n and (CAC)n motif presented a strong accumulation pattern in the telomeric regions of *H. wetmorei*. Also, (CA)n, (GC)n and (TA)n microsatellites accumulated in telomeric regions in both *H. malcolmi* and *H. wetmorei* (Khensuwan et al. 2023). It is known that triplet sequences are able to stabilize by harping on some alternative structures generated from errors of DNA polymerase (Sinden 1999); their presence at telomeres can be related to some repair mechanism. Repeated elements have been shown to be good tools for studying biodiversity, since they can "escape" from selection pressure that works on non-repetitive regions, making them evolutionary markers for detecting recent evolutionary changes (Cioffi et al. 2012; Garrido Ramos 2017; Moraes et al. 2017). Although the "Poropuntiinae" are thought to have diverged from other cyprinids about 37 Myr ago (Yang et al. 2021), recent changes in the genomes of those related species can have occurred, given the discrepant patterns of microsatellites and NOR herein observed, as in previous investigations as well (Khensuwan et al. 2023).

Conclusions

This study applied conventional and molecular cytogenetics to study the karyotypes and chromosomal characteristics of *H. malcolmi* and *H. wetmorei*. Both species present similar morphology and a conservative 2n = 50. However, they can be distinguished based on their chromosomal morphology, NORs sites and repetitive DNAs, such as (CAG)n, (GAA)n and (CAC)n, showed specificities in their distribution among species, thus being shown as good markers and promoters of specific genomic differentiation inside the genus.

Acknowledgements

References

Almeida-Toledo LF, Foresti F, Toledo-Filho SA (2000) Karyotypic evolution in neotropical freshwater fish. Chromosome Today 13: 169–181. https://doi.org/10.1007/978-3-0348-8484-6_13

Amemiya CT, Gold JR (1988) Chromosomal NORs as taxonomic and systematic characters in North American cyprinid fishes. Genetica 76: 81–90. https://doi.org/10.1007/BF00058806

Barth A, Souza VA, Solé M, Costa MA (2013) Molecular cytogenetics of nucleolar organizer regions in *Phyllomedusa* and *Phasmahyla* species (Hylidae, Phyllomedusinae): a cytotaxonomic contribution. Genetics and Molecular Research 12(3): 2400–2408. https://doi.org/10.4238/2013.July.15.3

Bertollo LAC, Cioffi MB, Moreira-Filho O (2015) Direct chromosome preparation from freshwater teleost fishes. In: Ozouf-Costaz C, Pisano E, Foresti F, Almeida Toledo LF (Eds) Fish cytogenetic techniques. Enfield USA, CRC Press, 21–26. https://doi.org/10.1201/b18534-4

Castro J, Rodriguez S, Arias J, Sanchez L, Martinez P (1994) A population analysis of Robertsonian and Ag-NOR polymorphisms in brown trout (*Salmo trutta*). Theoretical Applied Genetics 89: 105–111. https://doi.org/10.1007/BF00226990

Cioffi MB, Bertollo LAC (2012) Chromosomal distribution and evolution of repetitive DNAs in fish. In: Garrido-Ramos MA (Ed.) Genome dynamics. Basel (Karger), 197–221. https://doi.org/10.1159/000337950

Cioffi MB, Kejnovsky E, Bertollo LAC (2011) The chromosomal distribution of microsatellite repeats in the wolf fish genome *Hoplias malabaricus*, focusing on the sex chromosomes. Cytogenetic Genome Research 132: 289–296. https://doi.org/10.1159/000322058

Cioffi MB, Martins C, Bertollo LA (2009) Comparative chromosome mapping of repetitive sequences. Implications for genomic evolution in the fish, *Hoplias malabaricus*. BMC Genetics 10(1): 34. [1–11] https://doi.org/10.1186/1471-2156-10-34

Chaiyasan P, Supiwong W, Saenjundaeng P, Seetapan K, Pinmongkhonkul S, Tanomtong A, (2018) A report on classical cytogenetics of hihgfin barb fish, *Cyclocheilichthys armatus* (Cypriniformes, Cyprinidae). Cytologia 83(2): 149–154. https://doi.org/10.1508/cytologia.83.149

Chantapan T (2015) Standardized karyotype and ideogram of cyprin id fishes (Subfamily: Cyprininae) in Thailand. Ph.D. Dissertation, Khon Kaen University, 258 pp. [In Thai]

Donsakul T, Magtoon W (2002) Karyotypes of two cyprinid fishes, *Hypsibarbus wetmorei* and *Morulius chrysophekadion*, from Thailand. Proceedings of the 28th Congress on Science and Technology of Thailand. Queen Sirikit National Convention Centre, Bangkok, Thailand, October 24–26, 2002. Bangkok, 92–101.

Donsakul T, Rangsiruji A, Magtoon W (2007) Karyotypes of four cyprinid fishes: *Poropuntius normani*, *Hypsibarbus malcolmi*, *Scaphognathops bandanensis* and *Henicorhynchus caudiguttatus* from Thailand. Proceedings of the 45th Kasetsart University Annual Conference: Fisheries. Kasetsart University, Bangkok, Thailand, January-Febuary 30-2, 2007. Bangkok, 740–748.

Fang F, Noren M, Liao TY, Kallersjo M Kullander SO (2009) Molecular phylogenetic interrelationships of the south Asian cyprinid genera *Danio*, *Devario* and *Microrasbora* (Teleostei, Cyprinidae, Danioninae). Zoologica 38: 237–256. https://doi.org/10.1111/j.1463-6409.2008.00373.x

Fricke R, Eschmeyer WNR, Van der Laan R (2023) Eschmeyer's Catalog of fishes: genera, species, references. California Academy of Sciences, San Francisco, CA, USA. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Gadgil R, Barthelemy J, Lewis T, Leffak M (2017) Replication stalling and DNA microsatellite instability. Biophysical Chemistry 225: 38–48. https://doi.org/10.1016/j.bpc.2016.11.007

Galetti Jr PM (1998) Chromosome diversity in neotropical fish: NOR studies. Italian Journal of Zoology 65(1): 53–56. https://doi.org/10.1080/11250009809386795

Garrido-Ramos MA (2017) Satellite DNA: an evolving topic. Genes 8(9): 230. https://doi.org/10.3390/genes8090230

Gold JR, Amemiya CT (1986) Cytogenetic studies in North American minnows (Cyprinidae) XII. Pattern of chromosomal nucleolus organizer region variation among 14 species. Canadian Journal of Zoology 65: 1869–1877. https://doi.org/10.1139/z86-279

Goffová I, Fajkus J (2021) The rDNA loci-intersections of replication, transcription, and repair pathways. International Journal of Molecular Sciences 22(3): 1302. https://doi.org/10.3390/ijms22031302

Kentachalee N, Pinthong K, Tongnunui S, Tanomtong A, Juntharat S, Supiwong W (2023) First report on genetic structure of *Lobocheilos rhabdoura* (Fowler, 1934) (Labeoninae, Cyprinidae) in Thailand: Insight the Basic Knowledge for Fish Breeding. Food Agricultural Sciences and Technology 9(1): 90–103.

Kerrest A, Anand RP, Sundararajan R, Bermejo R, Liberi G, Dujon B, Richard GF (2009) SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination. Nature Structural & Molecular Biology 16(2): 159–167. https://doi.org/10.1038/nsmb.1544

Khensuwan S, Sassi FDM, Moraes RL, Jantarat S, Seetapan K, Phintong K, Cioffi MB (2023) Chromosomes of Asian cyprinid fishes: Genomic differences in conserved karyotypes of 'Poropuntiinae' (Teleostei, Cyprinidae). Animals 13(8): 1415. https://doi.org/10.3390/ani13081415

Knytl M, Kalous L, Rab P (2013) Karyotype and chromosome banding of endangered crucian carp, *Carassius carassius* (Linnaeus, 1758) (Teleostei, Cyprinidae). Comparative Cytogenetics 7(3): 205–213. https://doi.org/10.3897/compcytogen.v7i3.5411

Kobayashi T, Yamada F, Hashimoto T, Abe S, Matsuda Y, Kuroiwa A (2008) Centromere repositioning in the X chromosome of XO/XO mammals, Ryukyu spiny rat. Chromosome Research 16: 587–593. https://doi.org/10.1007/s10577-008-1199-5

Kubat Z, Hobza R, Vyskot B, Kejnovsky E (2008) Microsatellite accumulation on the Y chromosome in *Silene latifolia*. Genome 51: 350–356. https://doi.org/10.1139/G08-024

Levan A, Fredga K, Sandberg AA (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201–220. https://doi.org/10.1111/j.1601-5223.1964.tb01953.x

López-Flores I, Garrido-Ramos MA (2012) The repetitive DNA content of eukaryotic genomes. Repetitive DNA 7: 1–28. https://doi.org/10.1159/000337118

Magtoon W, Arai R (1989) Karyotypes of five *Puntius* species and one *Cyclocheilichtlys* species (Pisces, Cyprinidae) from Thailand. Bulletin of the National Science Museum, Tokyo. Series A (Zoology) 15(3): 167–175.

Moraes RLR, Bertollo LAC, Marinho MMF, Yano CF, Hatanaka T, Barby FF, Cioffi MDB (2017) Evolutionary relationships and cytotaxonomy considerations in the genus *Pyrrhulina* (Characiformes, Lebiasinidae). Zebrafish 14(6): 536–546. https://doi.org/10.1089/zeb.2017.1465

Nelson JS (1994) Fish of the word. 3rd ed. Wiley, New York, 752 pp.

Oliveira EA, Bertollo LAC, Yano CF, Liehr T, Cioffi MDB (2015) Comparative cytogenetics in the genus *Hoplias* (Characiformes, Erythrinidae) highlights contrasting karyotype evolution among congeneric species. Molecular Cytogenetics 8: 56. https://doi.org/10.1186/s13039-015-0161-4

Oliveira EA, Sember A, Bertollo LAC, Yano CF, Ezaz T, Moreira-Filho O (2018) Tracking the evolutionary pathway of sex chromosomes among fishes: characterizing the unique XX/XY$_1$Y$_2$ system in *Hoplias malabaricus* (Teleostei, Characiformes). Chromosoma 127: 115–128. https://doi.org/10.1007/s00412-017-0648-3

Pelletier R, Krasilnikova MM, Samadashwily GM, Lahue R, Mirkin SM (2003) Replication and expansion of trinucleotide repeats in yeast. Molecular and Cellular Biology 23(4): 1349–1357. https://doi.org/10.1128/MCB.23.4.1349-1357.2003

Pereira AR, Reed P, Veiga H, Pinho MG (2013) The Holliday junction resolvase RecU is required for chromosome segregation and DNA damage repair in *Staphylococcus aureus*. BMC microbiology 13(1): 1–9. https://doi.org/10.1186/1471-2180-13-18

Pereira V, Tomas C, Amorim A, Morling N, Gusmao L, Prata MJ (2011) Study of 25 X-chromosome SNPs in the Portuguese. Forensic Science International: Genetics 5(4): 336–338. https://doi.org/10.1016/j.fsigen.2010.10.004

Phillips RB, Hartley SE (1988) Fluorescent banding patterns of the chromosomes of the genus *Salmo*. Genome 30: 193–197. https://doi.org/10.1139/g88-033

Phillips RB, Ihssen PE (1989) Population differences in chromosome banding polymorphisms in lake trout (*Salvelinus namaycush*). Transactions of the American Fisheries 118: 64–73. https://doi.org/10.1577/1548-8659(1989)118%3C0064:PDICPI%3E2.3.CO;2

Phillips R, Rab P (2001) Chromosome evolution in the Salmonidae (Pisces): an update. Biological Reviews 76(1): 1–25. https://doi.org/10.1017/S1464793100005613

Phimphan S, Chaiyasan P, Suwannapoom C, Reungsing M, Juntaree S, Tanomtong A, Supiwong W (2020) Comparative karyotype study of three Cyprinids (Cyprinidae, Cyprininae)

in Thailand by classical cytogenetic and FISH techniques. Comparative Cytogenetics 14(4): 597–612. https://doi.org/10.3897/CompCytogen.v14i4.54428

Pucci MB, Barbosa P, Nogaroto V, Almeida MC, Artoni RF, Scacchetti PC (2016) Chromosomal spreading of microsatellites and (TTAGGG)n sequences in the *Characidium zebra* and *C. gomesi* genomes (Characiformes: Crenuchidae). Cytogenetic Genome Research 149(3): 182–190. https://doi.org/10.1159/000447959

Piyapong C (1999) Karyotypes and distribution of nucleolus organizer regions in four cyprinid species from Thailand. M.Sc. Thesis. Chulalongkorn University, 250 pp. [In Thai]

Rainboth WJ (1996) Fishes of the Cambodian Mekong. FAO species identification field guide for fishery purposes. FAO, Rome, 265 pp.

Rainboth WJ, Vidthayanon C, Yen MD (2012) Fishes of the greater Mekong ecosystem with species list and photographic atlas. Miscellaneous Publications Museum of Zoology, University of Michigan, USA, 314 pp.

Ruber L, Kottelat M, Tan HH, Ng PKL, Britz R (2007) Evolution of miniaturization and the phylogenetic position of *Pacdocypris*, comprising the world's smallest vertebrate. BMC Evolution Biology 7(1): 38. [1–10] https://doi.org/10.1186/1471-2148-7-38

Saenjundaeng P, Kaewmad P, Supiwong W, Pinthong K, Pengseng P, Tanomtong A (2018) Karyotype and characteristics of nucleolar organizer regions in longfin carp, *Labiobarbus leptocheilus* (Cypriniformes, Cyprinidae). Cytologia 83(3): 265–269. https://doi.org/10.1508/cytologia.83.265

Saenjundaeng P, Supiwong W, Sassi F, Bertollo LA, Rab P, Kretschmer R, Cioffi MDB (2020) Chromosomes of Asian cyprinid fishes: Variable karyotype patterns and evolutionary trends in the genus *Osteochilus* (Cyprinidae, Labeoninae, "Osteochilini"). Genetics and Molecular Biology 43(4): e20200195. https://doi.org/10.1590/1678-4685-gmb-2020-0195

Sassi FDM, Moreira-Filho O, Deon GA, Sember A, Bertollo LA, Liehr T, Cioffi MDB (2021) Adding new pieces to the puzzle of karyotype evolution in *Harttia* (Siluriformes, Loricariidae): Investigation of Amazonian species. Biology 10(9): 922. https://doi.org/10.3390/biology10090922

Sassi FDM, Toma GA, de Bello Cioffi M (2023) FISH–in Fish Chromosomes. In: Liehr T (Ed.) Cytogenetics and Molecular Cytogenetics. CRC Press, 281–296. https://doi.org/10.1201/9781003223658

Schemberger MO, Nascimento VD, Coan R, Ramos É, Nogaroto V, Ziemniczak K, Vicari MR (2019) DNA transposon invasion and microsatellite accumulation guide W chromosome differentiation in a Neotropical fish genome. Chromosoma 128: 547–560. https://doi.org/10.1007/s00412-019-00721-9

Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF (2001) Genomic and genetic definition of a functional human centromere. Science 294(5540): 109–115. https://doi.org/10.1126/science.1065042

Sember A, Bertollo LA, Ráb P, Yano CF, Hatanaka T, De Oliveira EA, Cioffi MDB (2018) Sex chromosome evolution and genomic divergence in the fish *Hoplias malabaricus* (Characiformes, Erythrinidae). Frontiers in Genetics 9(1): 71. https://doi.org/10.3389/fgene.2018.00071

Sinden RR (1999) Biological implications of the DNA structures associated with disease-causing triplet repeats. The American Journal of Human Genetics 64(2): 346–353. https://doi.org/10.1086/302271

Singh M, Barman AS (2013) Chromosome breakages associated with 45S ribosomal DNA sequences in spotted snakehead fish *Channa punctatus*. Molecular Biology Reports 40(1): 723–729. https://doi.org/10.1007/s11033-012-2112-z

Sochorová J, Garcia S, Gálvez F, Symonová R, Kovařík A (2018) Evolutionary trends in animal ribosomal DNA loci: introduction to a new online database. Chromosoma 127: 141–150. https://doi.org/10.1007/s00412-017-0651-8

Stults DM, Killen MW, Pierce HH, Pierce AJ (2008) Genomic architecture and inheritance of human ribosomal RNA gene clusters. Genome Research 18(1): 13–18. https://doi.org/10.1101/gr.6858507

Stults DM, Killen MW, Williamson EP, Hourigan JS, Vargas HD, Arnold SM, Pierce AJ (2009) Human rRNA gene clusters are recombinational hotspots in cancer ribosomal DNA: Human cancer recombination hotspot. Cancer Research 69(23): 9096–9104. https://doi.org/10.1158/0008-5472.CAN-09-2680

Terencio ML, Schneider CH, Gross MC, Vicari MR, Farias IP, Passos KB (2013) Evolutionary dynamics of repetitive DNA in *Semaprochilodus* (Characiformes, Prochilodontidae): a fish model for sex chromosome differentiation. Sexual Development 7(6): 325–333. https://doi.org/10.1159/000356691

Vicari MR, Artoni RF, Bertollo LAC (2005) Comparative cytogenetics of *Hoplias malabaricus* (Pisces, Erythrinidae): A population analysis in adjacent hydrographic basins. Genetics and Molecular Biology 28: 103–110. https://doi.org/10.1590/S1415-47572005000100018

Xu D, Lou B, Bertollo LAC, Cioffi MDB (2013) Chromosomal mapping of microsatellite repeats in the rock bream fish *Oplegnathus fasciatus*, with emphasis of their distribution in the neo-Y chromosome. Molecular Cytogenetics 6: 12. https://doi.org/10.1186/1755-8166-6-12

Yang L, Sado T, Hirt MV, Pasco-Viel E, Arunachalam M, Li J, Mayden RL (2015) Phylogeny and polyploidy: resolving the classification of cyprinine fishes (Teleostei: Cypriniformes). Molecular Phylogenetics and Evolution 85: 97–116. https://doi.org/10.1016/j.ympev.2015.01.014

Yang L, Mayden RL, Sado T, He S, Saitoh K, Miya M (2010) Molecular phylogeny of the fishes traditionally referred to Cyprinini *sensu stricto* (Teleostei: Cypriniformes). Zoologica Scripta 39(6): 527–550. https://doi.org/10.1111/j.1463-6409.2010.00443.x

Yang L, Naylor GJ, Mayden RL (2022) Deciphering reticulate evolution of the largest group of polyploid vertebrates, the subfamily cyprininae (Teleostei: Cypriniformes). Molecular Phylogenetics and Evolution 166: 107323. https://doi.org/10.1016/j.ympev.2021.107323

Yang T, Liang W, Cai J, Gu H, Han L, Chen H, Yan D (2021) A new cyprinid from the Oligocene of Qaidam Basin, north-eastern Tibetan plateau, and its implications. Journal of Systematic Palaeontology 19(17): 1161–1182. https://doi.org/10.1080/14772019.2021.2015470

Yano CF, Bertollo LAC, Cioffi MB (2017) Fish-FISH: Molecular cytogenetics in fish species. In: Liehr T (Ed.) Fluorescence In Situ Hybridization (FISH). Springer, Berlin, 429–443. https://doi.org/10.1007/978-3-662-52959-1_44

Yano CF, Poltronieri J, Bertollo LAC, Artoni RF, Liehr T, Cioffi MDB (2014) Chromosomal mapping of repetitive DNAs in *Triportheus trifurcatus* (Characidae, Characiformes): insights into the differentiation of the Z and W chromosomes. PLoS ONE 9: e90946. https://doi.org/10.1371/journal.pone.0090946

Yüksel E, Gaffaroğlu M (2008) The analysis of nucleolar organizer regions in *Chalcalburnus mossulensis* (Pisces: Cyprinidae). Journal of FisheriesSciences.com 2(4): 587–591. https://doi.org/10.3153/jfscom.2008021

Wang L, Sun F, Wan ZY, Yang Z, Tay YX, Lee M, Yue GH (2022) Transposon-induced epigenetic silencing in the X chromosome as a novel form of dmrt1 expression regulation during sex determination in the fighting fish. BMC Biology 20(1): 1–16. https://doi.org/10.1186/s12915-021-01205-y

ORCID

Sudarat Khensuwan https://orcid.org/0009-0007-4019-5489
Weerayuth Supiwong https://orcid.org/0000-0002-1670-3224
Phichaya Buasriyot https://orcid.org/0000-0003-0821-7629
Thomas Liehr https://orcid.org/0000-0003-1672-3054
Alongklod Tanomtong https://orcid.org/0000-0002-8466-3594

REVIEW ARTICLE

Ancient reproductive modes and criteria of multicellularity

Ilya A. Gavrilov-Zimin[1]

I *Zoological Institute, Russian Academy of Sciences, Universitetskaya nab. 1, St. Petersburg, 199034, Russia*

Corresponding author: Ilya A. Gavrilov-Zimin (coccids@gmail.com)

https://zoobank.org/AB8B7776-4966-42E3-BACC-5F22E87362E1

Citation: Gavrilov-Zimin IA (2023) Ancient reproductive modes and criteria of multicellularity. Comparative Cytogenetics 17: 195–238. https://doi.org/10.3897/compcytogen.17.109671

Abstract

It is demonstrated that the initial method of fertilization in animals (Metazoa), embryophyte plants (Embryophyta), most groups of multicellular oogamous algae, oogamous and pseudoogamous multicellular fungi was internal fertilization (in the broad meaning) in/on the body of a maternal organism. Accordingly, during the bisexual process, the initial method of formation of a daughter multicellular organism in animals was viviparity, and in embryophyte plants and most groups of oogamous multicellular algae – the germination of a zygote in/on the body of maternal organism.

The reproductive criteria of multicellularity are proposed and discussed. In this regard, the multicellularity is considered to subdivide terminologically into three variants: 1) protonemal, the most simple, characteristic of multicellular prokaryotes, most groups of multicellular algae and gametophytes of some higher plants; 2) siphonoseptal, found among multicellular fungi, some groups of green and yellow-green algae; 3) embryogenic, most complicated, known in all animals (Metazoa), all sporophytes and some gametophytes of higher plants (Embryophyta), charophyte green algae Charophyceae s.s., oogamous species of green and brown algae, some genera of red algae.

In addition to the well-known division of reproduction methods into sexual and asexual, it is proposed to divide the reproduction of multicellular organisms into monocytic (the emergence of a new organism from one cell sexually or asexually) and polycytic (fragmentation, longitudinal / transverse division or budding based on many cells of the body of the mother organism), since these two ways have different evolutionary and ontogenetic origins.

Introduction

The origin of multicellularity in the evolution of living organisms remains one of the most important discussion topics in evolutionary biology over the past one and a half centuries. The main hypotheses explaining the sequential phylogenetic transformation of colonial protists into the first truly multicellular organisms are well known and discussed many times in specialized scientific and educational literature (see, for example, Zakhvatkin 1949, 1956; Ivanov 1968; Ivanova-Kazas 1995; Bonner 1998; Grosberg and Stratchmann 2007; Michailov et al. 2009; Knoll 2011; Herron et al. 2013; Niklas and Newman 2013; Suga and Ruiz-Trillo 2013; Umen 2014; Coates et al. 2015; Brunet and King 2017; Malakhov et al. 2019; Colizzi et al. 2020; Lamża 2023, etc.). In these hypotheses and the discussions accompanying them, the main place is given to morpho-anatomical, ontogenetic and molecular changes, without which the transition from the simple unicellular level of life organization to a higher level is impossible. At the same time, the question of how exactly the reproduction of the first multicellular organisms could be carried out is given much less attention, and some important aspects are completely overlooked. However, a clear answer to this question is necessary to understand the entire course of the subsequent evolution of reproductive systems. In addition, as will be shown below, the features of reproduction can be considered as the important criteria for multicellularity itself.

The traditional, well-known division of reproduction modes into two large groups, sexual and asexual, has an almost universal meaning, since it is to some extent applicable to all living systems, with the exception of only prokaryotic organisms and viruses. To avoid confusion, it should be noted right away that asexual and sexual methods of reproduction are not always accompanied by the increasing of a population. For example, in higher plants (Embryophyta), as well as in most groups of algae and fungi, producing of numerous descendants occurs primarily with the asexual formation of spores, while as a result of the sexual process, only one daughter organism (usually a sporophyte) often develops on one maternal organism (usually a gametophyte), that is, there is no increase in the number of individuals. The sexual process in prokaryotic organisms and in some protists is not at all directly connected with reproduction.

Significant terminological confusion also occurs when discussing variants of parthenogenesis, i.e. development of an organism from a gamete without its fusion with another gamete. In recent decades, especially in the English-language literature (see, for example, Heesch et al. 2021), it has become commonplace to attribute parthenogenesis to asexual reproduction. With this approach, the difference between asexual and sexual reproduction is made dependent on a random event (fusion of gametes), which may not occur in the life cycle of an individual for external reasons that do not depend on its morphology, physiology, lifestyle, taxonomic and phylogenetic position. That is, the classification of a biological phenomenon (reproduction) in this case is made dependent on random non-biological causes. This approach could theoretically be justified by the homology and great similarity between the development of the unfertilized gamete and the spore in many simply constructed organisms. However, in all

higher plants, oogamous algae, and all animals, gametogenesis usually differs sharply from the processes of asexual reproduction and is associated with the spatial and functional separation of the germ cell line from somatic ones. In this regard, the traditional approach to understanding parthenogenesis as a variant of sexual reproduction seems more convenient, since parthenogenetic offspring arise from an extremely specialized haploid germ cell – the gamete, which, moreover, in many cases merges with one or another other product of gametogenesis to restore its diploidy (for example, with polar bodies). In plant organisms, parthenogenesis itself should, of course, be distinguished from other variants of apomixis, in which the embryo arises not from the egg, but from other cells of the embryo sac, nucellus, or integument (see: Yakovlev 1981: 7–8; Reproductive Systems 2000: 142–218).

In addition, when considering methods of reproduction of multicellular organisms, it is important not to lose sight of the following aspect. A daughter multicellular organism can arise from a single cell of the mother's body (spore, zygote, haploid gamete, parthenogenetic egg with restored diploidy, or simply a separate somatic cell that has retained totipotency [that is, the ability to produce various types of differentiated cells]) or simultaneously from many mother cells (with various variants of budding, fragmentation, simple division of the body into two or many parts). According to this criterion, the reproduction of multicellular organisms can be divided into monocytic and polycytic; the second term only partly overlaps with the concept of "vegetative reproduction", since in the botanical literature, simple mitotic division of unicellular algae is also called vegetative (see, for example, Belyakova et al. 2006b) and various cases of budding based on one initial meristematic cell (Reproductive Systems 2000: 342). In different senses, vegetative reproduction is also mentioned in the zoological literature (Ivanova-Kazas 1977). The term "blastogenesis" is closer in meaning to polycytic reproduction, which is understood as the opposite of embryogenesis (Ivanova-Kazas 1977: 227) and corresponds to polycytic budding (see below). As will be shown below, the division of reproduction into monocytic and polycytic is no less important for understanding the evolution of reproduction and self-reproduction than the criterion for the presence/absence of gamete fusion.

Numerous taxonomic names of organisms are used in the analysis below. It is important for the reader who does not have a serious personal experience of taxonomic work to take into account that there is no single universal system of living nature and a universal method of taxonomic constructions. For any group of organisms, the scientific literature presents competing views of various specialists and scientific schools on the phylogeny of the corresponding group and its "internal" classification. At the same time, phylogenetic schemes and taxonomic systems published later in date are by no means necessarily more correct or more reasoned than those published earlier. In this article, I do not have the opportunity to discuss any particular aspects of phylogenesis, the ideological basis of numerous classification schemes, contradictions between evolutionary and cladistic systematics, the suitability/unsuitability of various computer-molecular approaches, etc. Solely for practical convenience, I use the names of algal taxa appearing in the AlgaeBase database (https://www.algaebase.org/), since this

database compiles all nominal taxa of algae (as well as cyanobacteria) at the same time and reveals the corresponding nomenclature of names. The use of AlgaeBase does not mean my automatic agreement with all classification constructions implemented in this database. The same applies to the use of the names of higher taxa of heterotrophic protists and invertebrate animals, the classifications of which differ quite significantly in the works of different authors published in recent decades. In general, I follow the approach used in one of the most famous modern manuals on invertebrate zoology, a two-volume edition edited by Westheide and Rieger (Westheide and Rieger 2004). Unlike later papers (e.g., Dunn et al. 2014), which claim to reconstruct the phylogeny and provide a general classification of animals, this fundamental guide differs in that it is based primarily on easily verifiable and well-studied phenotypic characters of organisms. When using the names of higher taxa of terrestrial plants (Embryophyta) and fungi, I am guided by the multi-volume monograph "Botanica", prepared by a team of specialists from the Faculty of Biology of Moscow State University (Belyakova et al. 2006a, b; Timonin 2007; Timonin and Filin 2009; Timonin et al. 2009).

Reproductive criteria of multicellularity

For further discussions, it is necessary to clearly define the range of organisms that can be considered multicellular. Unfortunately, the border between the coloniality of unicellular protists and simple forms of multicellularity is understood in the scientific literature very vaguely. With an expanded approach to this issue (for example, Grosberg and Stratchmann 2007), multicellular organisms, in addition to animals, higher plants and a number of groups of algae, also mean some groups of slime molds and fungi, as well as a number of groups of Prokaryota.

In addition, there is no clear unequivocal separation of different types of multicellularity. Usually, one speaks only of simple and complex multicellularity (Knoll 2011; Niklas and Newman 2013), implying the presence of differentiated cells and tissues by the latter. However, the degree of differentiation varies greatly from one taxon to another (and even between individual stages of the life cycle of the same species of organisms) and demonstrates numerous chaotic transitions from simpler to more complex options and back.

In a broad interpretation, "clonal" and "aggregative" multicellularity are also distinguished (Grosberg and Stratchmann 2007; Coates et al. 2015; Lamża 2023), meaning by the latter the formation of cell clusters from the original free-living unicellular organisms. This approach seems to me unfortunate, since it does not allow any clear distinction between the various colonial prokaryotes, colonial fungi and algae, on the one hand, and the multicellular representatives of these same groups, on the other hand.

I consider it logical to proceed from the fact that a unitary multicellular organism, unlike a colonial one, obligatorily develops as a multicellular organism and reproduces itself only after it reaches the multicellular «vegetative» stage of ontogenesis. That is, the life cycle of a unitary multicellular organism is as follows (Fig. 1). In such a cycle,
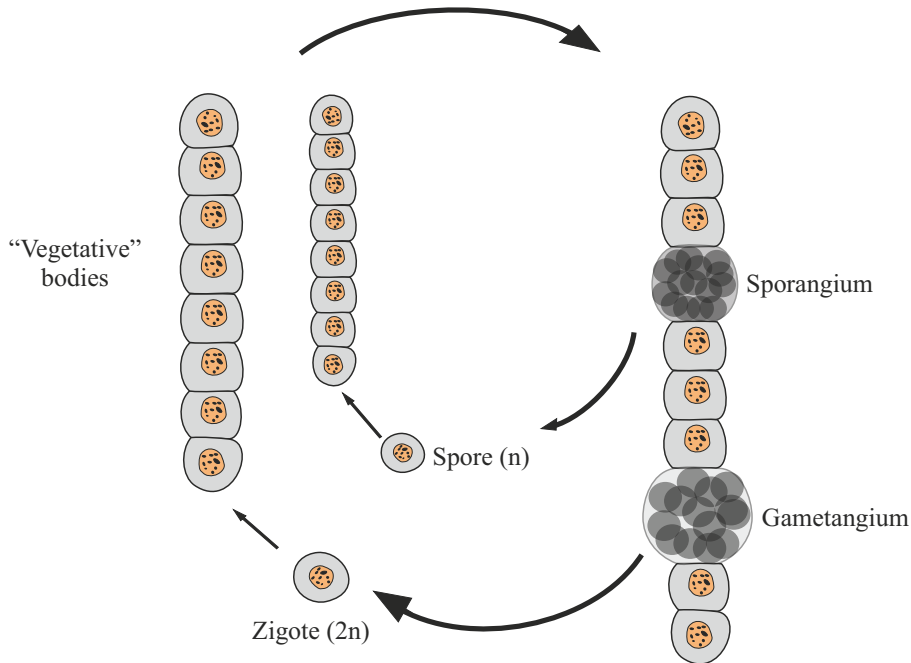
Figure 1. Generalized scheme of the life cycle of a multicellular organism (protonemal multicellularity).

the only unicellular (and mononuclear for eukaryotes) stage is the spore, gamete, or zygote that has no an independent life (i.e., nutrition, reproduction). A multicellular body of a unitary organism obligately grows from a spore, zygote or parthenogenetic gamete. This first reproductive criterion for multicellularity avoids ambiguity in the understanding of coloniality vs. unitary multicellularity and adequately assess the evolutionary consequences of the transition from one level of life organization to another. In particular, different variants of colonies in archaea (Archaea), myxobacteria (Mixococcales) and slime molds (Myxomycota, Acrasiomycota), even in the most complex cases, are only secondary accumulations of independent cells homogeneous in structure or multinucleated plasmodia, pseudoplasmodia, etc. From spores and/or zygotes of these organisms, daughter independent unicellular organisms are formed, which then gather into a new colony, or the zygote gives rise to a multinuclear plasmodium (see, for example, Novozhilov and Gudkov 2000: 417–443).

A similar situation occurs in the case of the formation of various specialized colonies (coenobia) of unicellular algae (for example, *Coelastrum* Nägeli, 1849, *Scenedesmus* Meyen, 1829, *Sphaerocystis* Chodat, 1897 and many others, especially among green and diatom algae), which are the result of secondary accretion or immersion in a common mucosal capsule of initially independent, self-feeding and reproducing cells. Inside each cell of the coenobium, small zoospores are again formed, which coalesce into a tiny daughter coenobium inside the mother cell, and then are released due to the rupture of the wall of this cell (Matvienko 1977: 271).

I also do not consider as multicellular organisms various multinucleated coenocytes (= somatella, cytoids, polycystids, etc.), known in some complexly organized ciliates, opalines, sporozoans, dinoflagellates, foraminifera and other protists. All these organisms do not meet the first reproductive criterion of multicellularity formulated above. The bodies of some "colonial" ciliates, for example, from the genus *Zoothamnium* Bory de St. Vincent, 1824, formed as a result of incomplete monocytic budding. Nevertheless, the resulting "colony" remains a de facto unicellular formation, within which there are no partitions, and all parts of which are connected by cytoplasmic strands caused by the so-called spasmonemes (Foster et al. 1978). There is no division into cells and inside multinucleated bodies (cenocytes) of parasitic dinoflagellates of the genus *Haplozoon* Dogiel, 1906 (see Angel et al. 2021), which were earlier erroneously identified as primary multicellular organisms (see, for example, Ivanov 1968).

Some difficulty can be caused by the application of the first reproductive criterion in relation to various cases of asexual reproduction at the initial stages of development of a multicellular organism. So, for example, in some cnidarians (Cnidaria) under experimental conditions, individual blastomeres retain the ability to give rise to independent embryos (Zakhvatkin 1949: 217). In a number of multicellular green algae (Chlorophyta) from the orders Ulotrichales, Sphaeropleales, Oedogoniales, and simply organized Charophyta s.l. from the order Coleochaetales the so-called "unicellular sporophyte" is preserved in the life cycle; it is a zygote, which is covered with a protective membrane and, after a dormant period, divides meiotically (and then mitotically), giving rise to 4–32 haploid zoospores (Belyakova et al. 2006b: 221, 267). In many red algae (Rhodophyta), the zygote gives rise to the so-called "gonimoblast filaments" (see more details below). In all these cases, no separate unicellular cycle of nutrition, development, and reproduction arises, since the mentioned zygotes are not independent organisms, and the products of their division obligatory grow into multicellular bodies.

Regular polyembryony, which occurs in a number of groups of highly developed animals and plants, is all the more not an example of unicellular reproduction, since it is realized on a multicellular basis (with the exception of random developmental anomalies in some individuals). First, a multicellular body of the embryo begins to form from a zygote or a parthenogenetic egg, and only then it is divided into several or many daughter embryos (Ivanova-Kazas 1977: 199–213, 1995: 480), i.e., in fact, we are talking about some kind of monocytic or polycytic budding (see more details below) in all studied examples of regular polyembryony. In higher plants, "polyembryony" is often understood not as the division of one embryo into several daughter ones, but as the appearance of many embryos and embryoids from different cells of the embryo sac, nucellus, and ovulum (Reproductive Systems 2000: 401).

A certain difficulty is also caused by the understanding of multicellularity in secondarily simplified parasitic animals – orthonectids (Orthonectida), in which one of the stages of the life cycle is a multinuclear "plasmodium", capable of reproducing by monocyte budding. However, inside such a plasmodium, in addition to trophic nuclei,

there are also generative nuclei with isolated sections of the cytoplasm, which are agametes (Malakhov 1990: 49; Slyusarev 2008). Thus, the body of these organisms is not a simple plasmodium, known in different protists, but a system of small cells located inside another, larger cell - a phenomenon known for a number of groups of animals and higher plants (see more details below).

The second reproductive criterion of multicellularity determines exactly how a multicellular body reproduces itself by the monocyte method of forming a daughter organism and allows us to divide all known ways of implementing obligate multicellularity into three fundamentally different variants.

The simplest and most archaic variant is protonemal multicellularity, in which a spore or zygote divides monotomically (by mitosis or simple cytokinesis), forming a single filament, a protonema (Fig. 1).

Monotomic division implies the obligatory growth of daughter cells after their division. As a result, a multicellular structure is formed from cells of approximately the same size, quite similar to the original cell or even exceeding its size. Such a single-row thread can then grow, branch many times, intertwine, forming a multilayer body (thallus). Protonemic multicellular organisms include the following groups:

1.   Multicellular species of cyanobacteria (Cyanobacteria), actinobacteria (Actinobacteria), caryophane bacteria (Caryophanales) and some other prokaryotic groups. Cases of palintomy occurring in prokaryotes (for example, in the cyanobacteria *Gloeocapsa* Kützing, 1843, *Mycrocystis* Kützing, 1833, etc.) lead to the formation of independent daughter cells "nanocytes", while multicellular bacterial thalli are formed from spores (akinetes) in a monotomic way (see, for example, illustrations in Kaplan-Levy et al. 2010).

2.   Some genera of golden algae (Chrysophyceae), for example, *Hydrurus* Agardh, 1824, *Nematochrysis* Pascher, 1925, *Phaeodermatium* Hansgirg, 1889, etc.

3.   Separate genera of yellow-green algae (Xanthophyceae), such as *Tribonema* Derbès et Solier, 1851, *Xanthonema* Silva, 1979, *Heteropedia* Pascher, 1939, *Heterococcus* Chodat, 1908, etc.

4.   Some genera of pheotamniophic algae (Phaeothamniophyceae), for example, *Phaeothamnion* Lagerheim, 1884 and possibly *Sphaeridiothrix* Pascher et Vlk, 1943.

5.   Isogamous and heterogamous genera of brown algae (Phaeophyceae), for example, from the orders Discosporangiales, Sphacelariales, Ectocarpales, etc., as well as the monotypic genus *Schizocladia* Henry et al., 2003, which the authors of this taxon propose to consider as an independent class Schizocladiophyceae, sister to brown algae.

6.   Most multicellular red algae (Rhodophyta), with the exception of a number of highly developed genera (see below), in which an embryogenic variant of the development of bodies from carpospores and tetraspores is observed.

7.   Obligate multicellular representatives of green algae (Chlorophyta s.l.) that meet the first reproductive criterion of multicellularity. For example, *Microthamnion* Nägeli, 1849 (Trebouxiophyceae: Microthamniales), *Schizogonium* Kützing, 1843, *Prasiola* Meneghini, 1838, *Raphidonema* Lagerheim, 1892 (Trebouxiophyceae:

Prasiolales), *Protococcus* Agardh, 1824 (Chlorophyceae: Chlamidomonadales[*]), some genera of Sphaeropleales[*], most representatives of Ulvales, Ulotrichales, Trentepohliales, Chaetophorales, and Oedogoniales.

8. Obligate multicellular representatives of charophyta algae (Charophyta s.l.) from the classes Klebsormidiophyceae, Zygnematophyceae, and Coleochaetophyceae. To the contrary, highly organized charophyceous algae (class Charophyceae s.s.) develop according to the type of embryogenic multicellularity (see below).

9. The gametophytes of many genera of higher plants, especially bryophytes (Bryomorphae) and ferns (Pteridophyta), but in some cases also Lycopodiophyta, retain the simple protonemal character of spore germination. On the protonema, by budding, more complex bodies of gametophytes, differentiated into tissues and organs, can subsequently form. However, in other genera of the same plant groups, spores undergo palintomic/syntomic cleavage and develop according to the type of embryogenic multicellularity (see below).

The second variant is siphonoseptal multicellularity (Fig. 2). Here, the zygote or spore initially undergoes multiple karyokinesises without division of the cytoplasm and forms a multinucleated cell, i.e. cenocyte. Further, this cell grows apically, sometimes reaching macroscopic dimensions of several tens of centimeters, and inside such a body, called the term "siphon", regular or irregular partitions (septae) appear, dividing this siphon into multi-core compartments or clades (from the Greek "κλάδος" – a branch) with a different, less often the same, number of nuclei. Septae are formed by centripetal ingrowth of the membrane and cell wall into the inner cavity of the cell (Fritsch 1929; Egerod 1952; Enomoto and Hirose 1971; McDonald and Pickett-Heaps 1976; Liliaert et al. 2007; Okuda et al. 2016). This variant of body formation is well known in a number of genera of green algae of the order Siphonocladales and some other not closely related genera of green algae (see below). However, in fact, the same principle of the formation of a multinuclear thallus, divided into sections by septae, also takes place in various multicellular fungi and fungi-like organisms, including those that form septae only to separate sporangia and gametangia from a multinuclear hypha. The latter is typical, for example, for many oomycetes (Oomycota) and chytridiomycetes (Chytridiomycota). For this reason, I propose to understand siphonoseptality as a variant of multicellularity that arose independently in different groups of fungi and algae. A peculiar formation of irregular "septae" growing centripetally is also known during the formation of colonies in some mycobacteria (Dobrovolskaya 1974: 299).

Unfortunately, the ultrastructural and biochemical mechanisms of septa formation in multicellular algae, fungi, and, especially, prokaryotes, remain insufficiently studied, and the available knowledge is limited to single model objects (Barr and Gruneberg

---

[*]   In green algae of the genus *Volvox* Linnaeus, 1758, s.l. (Chlorophyceae: Chlamidomonadales) embryogenic multicellularity is realized, while in the genus *Sphaeroplea* Agardh, 1824 (Chlorophyceae: Sphaeropleales) siphonoseptal multicellularity occurs (see below).
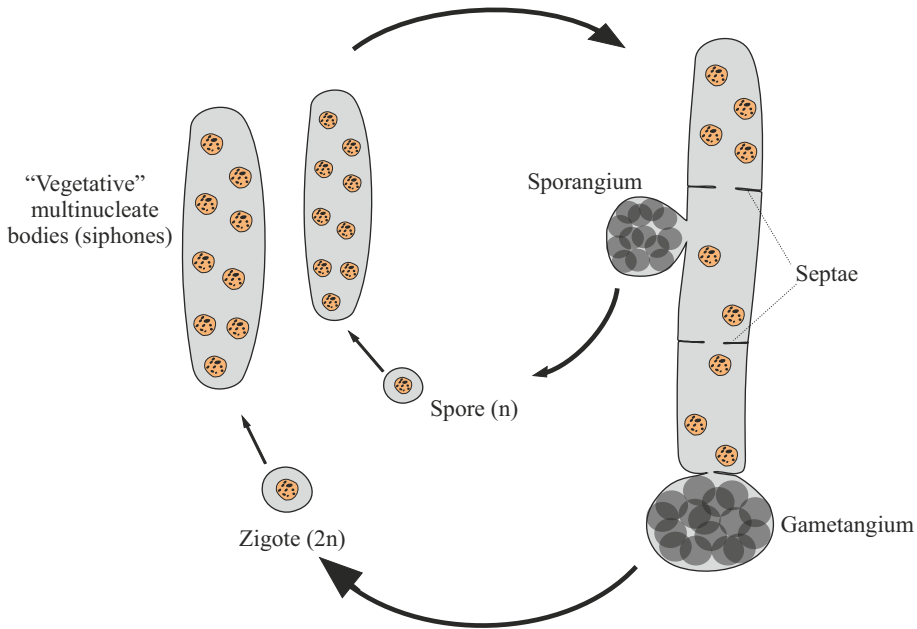
Figure 2. Generalized scheme of the life cycle in siphonoseptal multicellular organisms.

2007; Seiler and Heilig 2019). In addition, in some siphonoclad algae (*Siphonocladus* Schmitz, 1879, *Dictyosphaeria* Decaisne, 1842, *Cladophoropsis* Børgesen, 1905, *Boodlea* Murray et De Toni, 1889, *Struvea* Sonder, 1845, and *Chamaedoris* Montagne, 1842), instead of the formation of septae, a special "segregative" division of body into separate parts occurs, and these parts then fuse again (Egerod 1952; McDonald and Pickett-Heaps 1976; Liliaert et al. 2007; Okuda et al. 2016). In fact, such bodies are not multicellular, but are just colonies of cenocytes, each of which, having separated, can give rise to a new organism.

It should be noted that the structure of the septate bodies of fungi and algae is not similar to the complicated construction of some protists (Protista), for example, gregarine (Gregarinea). In the latter, a single cell is sometimes divided into communicating parts by a "tangle of thin fibrils" (Simdyanov 2007: 50), while in parasitic dinoflagellates of the genus *Haplozoon* Dogiel, 1906, a single coenocyte is partially divided due to "alveolar vesicles" (Angel et al. 2021).

Siphonoseptal multicellularity is characteristic of the following groups:

1. A number of genera of Ulvophyceae green algae from the order Siphonocladales (for example, *Anadyomene* Lamouroux, 1812, *Cladophora* Kützing, 1843, *Valonia* Agardh, 1823, etc.), individual representatives of the related orders Dasycladales and Siphonales, in which septa are formed during the separation of rhizoids, sporangia and gametangia (for example, *Bryopsis* Lamouroux, 1809, *Derbesia* Solier, 1846,

Pseudobryopsis Berthold, 1904, etc.), as well as the genus *Sphaeroplea* Agardh, 1824 from the order Sphaeropleales (see Fritsch, 1929).

2.   Some genera of yellow-green algae (Xanthophyceae or Tribophyceae) from the order Vaucheriales. The multinuclear branching filaments of these algae usually lack septae, but their sporangia and gametangia are separated by septae.

3.   Various groups of multicellular fungi and fungi-like organisms (as Oomycota, Chytridiomycota, etc.). In some fungi, for example, powdery mildew ascomycetes of the order Erysiphomycetes, there is a regular formation of septae with successive formation of mononuclear compartments of the hyphae (Belyakova et al. 2006a: 240). In many cases, especially during the formation of fungal "fruiting bodies", false tissues are formed due to close fusion and even anastomoses between hyphae. This phenomenon is in many ways reminiscent of the secondary fusion of multinucleated cenocytes in siphonoclad algae, which are characterized by segregative division of the original cell.

4.   It is possible that siphonoseptal multicellularity is also present in some ichthyosporids (Ichthyosporea), which are considered a group close to fungi and animals. At least some species of ichthyosporids form multinucleated thalli separated by septae, or such septae separate sporangia from the main "vegetative" body (Karpov 2011: 342–369). On the other hand, some ichthyosporid species have been suggested to have syntomic cell division (Suga & Ruiz-Trillo 2013). In general, ichthyosporids remain a poorly studied group, and the presence of a sexual process in them is assumed, but not proven.

Finally, the third and most complicated variant is embryogenic multicellularity (Fig. 3). It arises on the basis of obligate accumulative oogamy or accumulative aplanosporia, in which the gamete/spore exceeds in size (sometimes hundreds and even thousands of times (Ivanova-Kazas 1975: 39)) the original mother cells. As a result of palintomic or syntomic divisions, an embryo or embryoid is formed from an oogamete/spore (see below). Actually, only with this variant of reproduction for the first time in the evolution of living systems does the embryo appear as a biological phenomenon. In asexual monocytic reproduction, the analog of the oogamete is a large, immobile spore, the aplanospore, which gives rise to the embryoid. The term "embryoid" is widely and very ambiguously used in botany (less often in zoology) to refer to a variety of germ-like bodies arising from somatic cells (Reproductive Systems 2000: 334). I consider it expedient to understand by embryoids only cases of complete analogy with the sexual embryo: the emergence of the body from a single cell, enlarged in size, undergoing palintomic or syntomic divisions. The remaining cases of the emergence of daughter bodies from somatic cells I refer to budding.

The embryogenic variant of multicellularity is observed in the following organisms.

1.   All animals (Metazoa) as a holophyletic group that originally arose on the basis of embryogenic multicellularity.

2.   Sporophytes of all higher plants (Embryophyta).
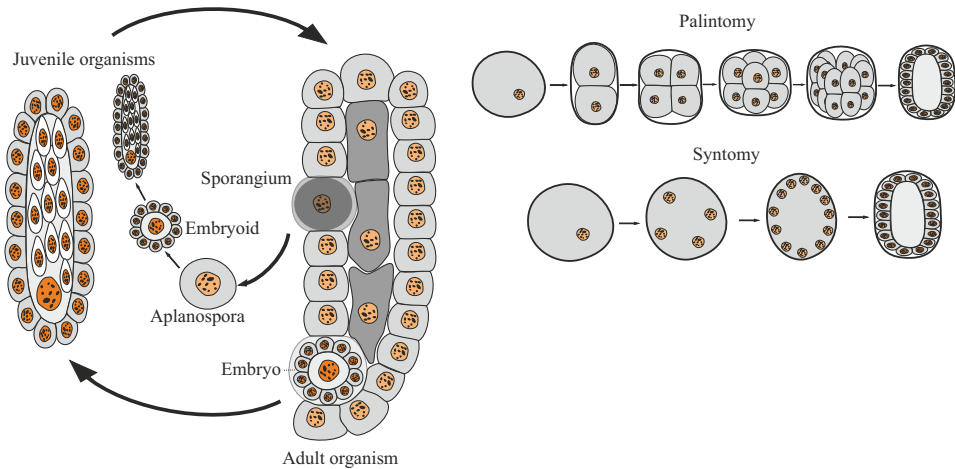
Figure 3. Generalized scheme of the life cycle and initial stages of development in embryogenic multicellular organisms.

3. Gametophytes of a number of genera of higher plants (Bryomorphae, Lycopodiophyta, and Pteridophyta), in which spores undergo palintomic/syntomic fragmentation inside their shell, often still inside sporangia (Fig. 4). In bryophytes (Bryomorphae), such spores give rise to a multicellular embryoid, from which a more or less large gametophyte then grows (see review in Nehira 1983). Such an embryoid looks quite similar to the embryos arising from the zygote and giving rise to the sporophyte generation. In Lycopodiophyta, gametophytes are microscopic organisms, in most cases formed as a result of palintomic or syntomic spore cleavage, while in Selaginellopsida and Isoetopsida gametophytes do not leave the spore shell at all (Filin 1978; Timonin and Filin 2009: 181–221) (Fig. 4). Among ferns (Pteridophyta), palintomic/syntomic division of the spore is characteristic of heterosporous ferns, while homosporous ferns retain the protonemal character of gametophyte development (Nayar and Kaur 1971; Timonin and Filin 2009: 221–312).

4. Charophyceae s.s. in the traditional narrow sense.

5. Oogamous genera of brown algae (Phaeophyceae) from the orders Fucales, Desmarestiales, Dictyotales, Laminariales, Chordales, Tilopteridales, Sporochnales, etc. (see the summary table of such genera in Luthringer et al., 2014), characterized by complex differentiation of cells and tissues, like the sporophytes of higher plants. Some of these genera (for example, *Fucus* Linnaeus, 1753, *Sargassum* Agardh, 1820, etc.) have a diplontic life cycle with gametic meiosis. That is, the reduction in the number of chromosomes occurs during the formation of gametes, similar to how it takes place in the life cycle of animals; there is no haploid generation in such a cycle. Other genera (for example, *Dictyota* Lamouroux, 1809, *Padina* Adanson, 1763, etc.) demonstrate a haplodiplontic life cycle with isomorphic generations, i.e. gametophytes are morpho-anatomically quite similar to sporophytes.
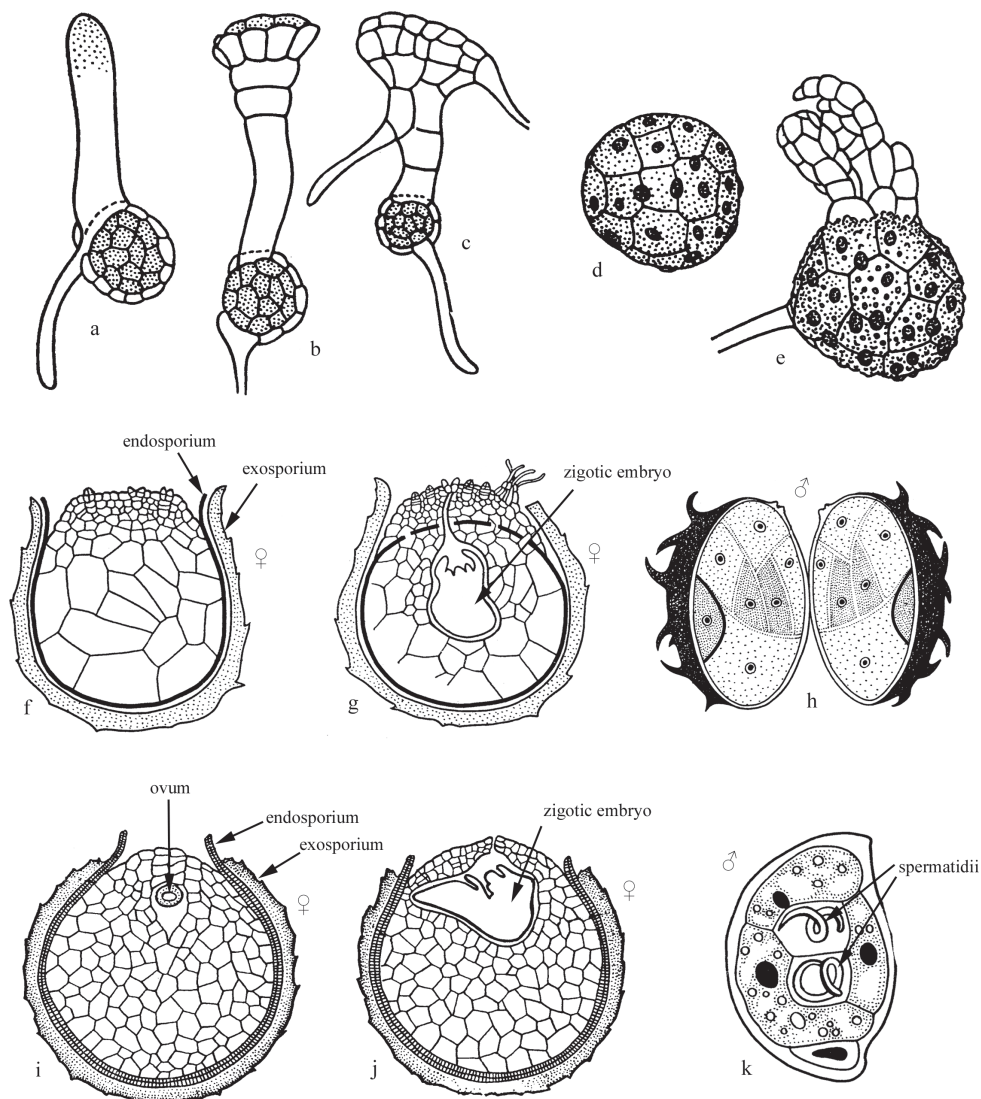
Figure 4. Embryoid gametophytes of higher plants. **a–c** *Reboulia hemisphaerica* (Linnaeus, 1753) (Marchantiophyta) **d, e** *Frullania muscicola* Stephani, 1894 (Marchantiophyta) **f–h** *Selaginella* spp. (Lycopodiophyta); **i–k** *Isoetes* sp. (Lycopodiophyta) **a–e** after Abramov and Abramova 1978, with changes **f–k** after Filin 1978, with changes.

In the third group of genera (for example, *Himantothallus* Scottsberg, 1907, *Desmarestia* Lamouroux, 1813, *Laminaria* Lamouroux, 1813, etc.), heteromorphism of generations is observed (Petrov 1977: 143–192; Luthringer et al. 2014). In these cases, sporophytes usually have an embryogenic origin, while strongly reduced filamentous gametophytes develop from a protonema or even represent a single cell. In addition, examples of irregular alternation of haploid and diploid generations,

parthenogenetic germination of gametes and the formation of microscopic protonemal sporophytes ("plethysmothallus") are known in brown algae, capable of producing not only spores, but also directly give rise to a macroscopic thallus (Petrov 1977: 143–192). All this confusing picture of the reproductive strategies of Phaeophyceae probably indicates the multiple independent origin of oogamy and embryogenic multicellularity in them during the haploid and/or diploid phases of the life cycle. Some authors (Heesch et al. 2021) make unexpected suggestions about secondary transitions from oogamy to heterogamy and isogamy in brown algae. However, it should be noted that these hypotheses are based solely on the belief in the infallibility and universality of molecular statistical cladism as a method of phylogenetic reconstructions.

It is interesting that in a number of works on various genera of brown algae, for example, in the articles by Nanda (1993), Edwards (2000), Kawai et al. (2001), and Bogaert et al. (2017), the initial stages of development of these algae are directly called embryonic, i.e. the similarity of the division of their zygotes with the embryonic development of higher plants and animals was noted.

6.   Some genera of red algae (Rhodophyta). In species some highly developed genera, for example, *Corallina* Linnaeus, 1758, *Dumontia* Lamouroux, 1813, *Jania* Lamouroux, 1812, *Amphiroa* Lamouroux, 1812, *Gracilaria* Greville, 1830, etc., palintomic divisions of aplanospores ("tetraspores" and "carpospores") is observed with subsequent formation a kind of hemispherical multicellular disk (Chemin 1937; Jones and Moorjani 1973; Michetti et al. 2013; Wai 2018; etc.). This structure is quite consistent in origin with the embryoids formed during asexual reproduction in other groups of plant organisms. Some authors (Chemin 1937: 369) have even compared this development with the formation of the embryonic morula of animals.

7.   Oogamous species of green algae of the genus *Volvox* Linnaeus, 1758, s.l., with differentiation of cells connected by plasmodesmata. The embryogenic origin of the multicellularity of oogamous *Volvox* spp. is well known in the literature and described in detail in many works, for example, by Zakhvatkin (1949: 220–232). A review of more recent data can be found, for example, in Desnitsky (2018).

From the above list of organisms, it can be seen that embryogenic multicellularity did not arise on the basis of prokaryotic cells. This fact, of course, is not accidental and is probably due to the fact that prokaryotic cells are not capable of providing effective intercellular transport of substances and, accordingly, of the formation of differentiated tissues. As a result, prokaryotes do not have examples of the embryonic development required for initial cell differentiation. Moreover, due to the absence of the endoplasmic reticulum, the transport of substances within prokaryotic cells is limited by the possibilities of diffusion, which imposes significant restrictions on cell size. Large sizes (sometimes up to 0.75 mm in diameter) of cells in some prokaryotes, for example, in the bacterium *Triomargarita namibiensis* Schulz et al., 1999, are explained by the fact that the entire central part of such cells is occupied by vacuoles, while the cytoplasm forms only a thin peripheral layer (Schulz et al. 1999).

Monocytic reproduction of protonemal multicellular organisms

Monocytic bisexual reproduction in protonemal multicellular organisms can proceed according to the type of isogamy, heterogamy, oogamy, or analogs of oogamy, whereas asexual monocytic reproduction can proceed according to the type of zoosporia or aplanosporia. Evolutionary models for the emergence of gamete diversity (anisogamy) from the initial isogamous sexual process have been repeatedly proposed in the specialized literature (Parker at al. 1972; Bell 1978; Bulmer and Parker 2002; Umen and Coelho 2019, etc.; see also the review by Blute 2012) and therefore there is no need to dwell on the discussion of this issue here. In general, there is no doubt that the appearance of anisogamy, with rare exceptions, directly correlates with an increase in the complexity of the body of an organism and, in particular, with the appearance of multicellularity (Bell 1978).

The various evolutionary transformations within the broadly understood oogamy deserve more detailed consideration, since, as will be shown below, oogamy is a necessary prerequisite for the transition to complex forms of multicellularity. The oogamous sexual process (or its analogues) in protonemal multicellularity is still carried out in an extremely achaic way, since in this case the oogamete (with rare exceptions) does not accumulate nutrients for further development, but remains comparable in volume to usual somatic cells or even turns out to be significantly smaller than the latter. As a result of this, the further development of the parthenogenetic or fertilized oogamete (zygote) inevitably occurs through monotomic germination, i.e. successive division and growth of daughter cells forming a filamentous structure (protonema).

It should be noted that examples of archaic oogamy are already found in unicellular and unicellular-colonial organisms. Thus, some genera of colonial diatoms (Diatomophyceae), for example, the so-called centric diatoms (orders Thalassiosirales, Coscinodiscales, Melosirales, Chaetocerotales) and pennate diatoms of the genus *Rhabdonema* Kützing, 1844, demonstrate oogamy, in which germ cells are smaller than somatic ones (Belyakova et al. 2006b: 85–93; Kaszmarska et al. 2013; Davidovich 2019: 31, 62). A similar archaic oogamy is known in some unicellular Trebouxiophyceae algae (Gonzalves and Mehra 1959). In most of the studied species of gregarine (Gregarinea), during sexual reproduction, two parental haploid cells unite, forming the so-called syzygy, and become covered by a common membrane (Fig. 5). Inside the shell of the syzygy, each parent cell divides by syntomy (schizogony) and forms gametes. The latter can be the same in size and functionality (isogamy) or differ significantly (anisogamy). In different genera of gregarines, immobile "female" gametes can, at the same time, be larger or smaller than mobile "male" gametes with flagella (Zakhvatkin 1949: 197; Grassé 1953; Simdyanov 2007: 26, 52–61). In most cases, the resulting gametes are many times smaller than the original parental cells, or slightly smaller (when a single zygote is formed inside the syzygy), but never exceed them in size. The fusion of gametes occurs inside the shell of the syzygy. Each resulting zygote is surrounded by its own protective shell and becomes an "oocyst". Subsequently, the "oocyst" undergoes two meiotic divisions, and the resulting haploid cells give rise to a new generation of unicellular or polycystid gregarines (Simdyanov 2007: 33, 50).
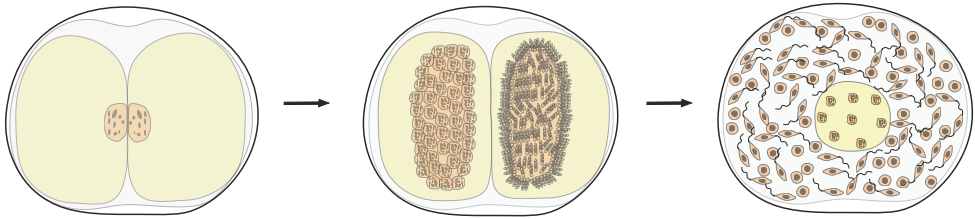
Figure 5. Formation of syzygy and copulation in the gregarine *Stylocephalus longicollis* (Stein, 1848).

In the related group of coccidia (Coccidea), the oogamete is formed directly from the haploid parent cell (merozoite) without division of the latter, and biflagellated (rarely non-flagellated) male gametes arise as a result of syntomic division of the merozoite. The possibility of fusion of gametes in this case is achieved by the fact that the parent cells are in close proximity to each other inside the body of the host organism. Meiosis in the life cycle of coccidia, as in gregarines, occurs in the "oocyst" formed from the zygote (Beyer 2007: 149–248).

Some highly developed ciliates that form "colonies" by incomplete budding (Fursenko 1924; Ivanov 1968: 30–31) demonstrate a kind of analogue of oogamy, in which the "macrogamete" (macrozooid) remains motionless, and the mobile small "microgamete" (microzooid) swims up and carries out "fertilization" (Fig. 6).

Relatively few examples of archaic oogamy (without an increase in the size of the gamete) are known among protonemal multicellular organisms. For example, such oogamy has been well studied in green algae of the genus *Prasiola* Meneghini, 1838 (Trebouxiophyceae: Prasiolales). In the upper part of their multicellular diploid thallus, meiotic divisions occur and biflagellated spermatozoa and non-flagellated oogametes (ova) are formed. Female gametes are about twice as large as male, but smaller than the original diploid cells of thallus. They are released due to the destruction ("dissolution") of the lower cell walls of thallus and end up in a bubble-like space bounded by the persistent outer common shell of the thallus ("persisting bladder-like coating lamella"). At the same time, hundreds or even thousands of heterosexual gametes are released into this space and fertilization occurs. A protonema grows from the zygote, and a new diploid thallus grows from it (Friedmann 1959; Cole and Akintobi 1963). Thus, there is hermaphroditism and self-fertilization in a closed space, which resembles the corresponding processes in various intracavitary parasitic organisms.

Even rarer in protonemal multicellular organisms, accumulative oogamy occurs, in which an increase in the volume of the egg takes place in comparison with the cells of the "vegetative" body that preceded it (Fig. 7). This variant is known in a number of multicellular green algae (Chlorophyta) from the orders Ulotrichales, Oedogoniales and in simply organized members of charophyta algae (Charophyta s.l.) of the order Coleochaetales. In their life cycle, the so-called "unicellular sporophyte" is preserved, which is a zygote, covered with a protective membrane and, after a dormant period, divides meiotically and then mitotically, giving rise to 4–32 haploid zoospores (Vinogradova 1977b: 282, 285; Belyakova et al. 2006b: 221, 267). In Oedogoniales, special "androspores"
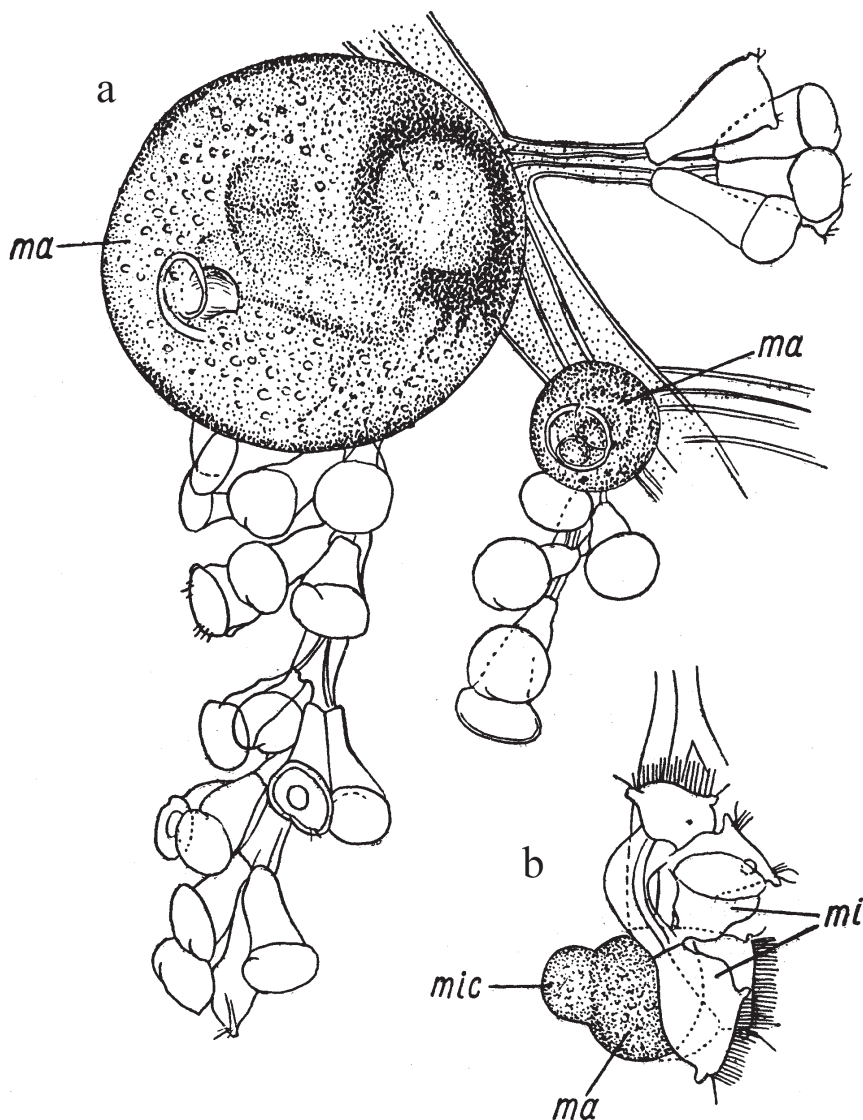
Figure 6. Analogy of the oogamous sexual process in the ciliate *Zoothamnium arbuscula* Ehrenberg, 1839 **a** colony with macrozooids (ma) **b** conjugation (mac – macroconjugant, mic – microconjugant, mi – microzooid). After Fursenko 1924 and Ivanov 1968.

settling on the oogonium or cells adjacent to it form peculiar dwarf gametophytes – "nanandria", the upper cells of which function as antheridia (Vinogradova 1977b: 293; Belyakova et al. 2006b: 254) (Fig. 7). A similar process is observed in *Cylindrocapsopsis* Iyengar, 1957 (Chlorophyceae: Sphaeropleales) (Vinogradova, 1977b: 294).

A peculiar analogy of archaic oogamy among protonemal organisms occurs in red algae (Rhodophyta) (Fig. 8). Their male gametes (sperms) are devoid of flagella and are
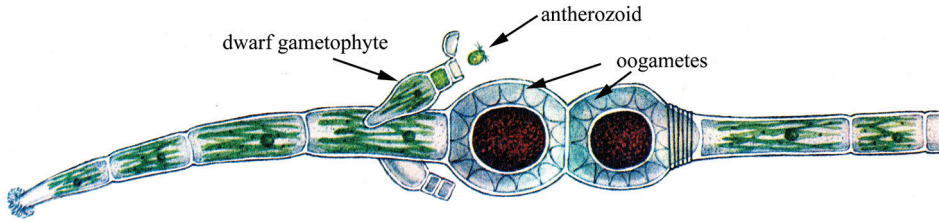
Figure 7. Accumulative oogamy in protonemal multicellularity in *Oedogonium stellatum* Wittrock ex Hirn, 1900 (after Vinogradova 1977b, with modifications).

passively transferred to the female genital organs (carpogons). There is no female gamete as such. The male gamete fuses with the carpogon nucleus. The fusion nucleus then grows into a diploid gonimoblast ("carposoporophyte"). Carposporangia producing spores are formed on the gonimoblast. These spores form the "second diploid generation – the tetrasporophyte" (Vinogradova 1977a).

The increase in the number of individuals in protonema-multicellular organisms occurs mainly during the production of spores. In fact, the spore in archaic organisms is quite homologous to the unfertilized gamete, which was convincingly shown, for example, in the fundamental work of Zakhvatkin (1949; 1956). In many algae, unfertilized gametes, including flagellar ones, can develop into new thalli – see, for example, Smith 1947; Belyakova et al. 2006b: 225–226, 234.

It is well known that, similar to the evolutionary transition from small mobile gametes to large immobile gametes, in various groups of organisms there is a transition from small zoospores to immobile aplanospores, which in many cases do not exceed
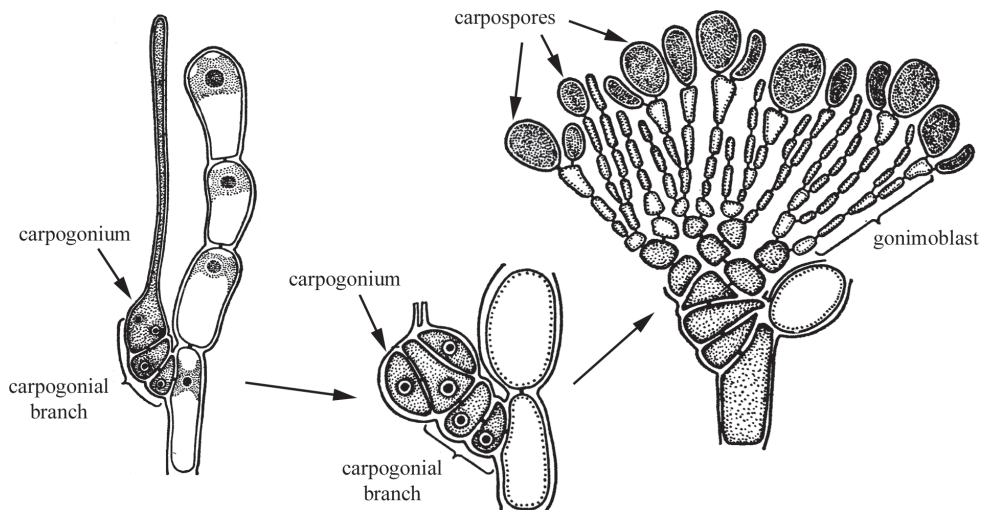
Figure 8. Scheme of development of the "carposporophyte generation" of floridian red algae (Rhodophyta: Florideophyceae). After Vinogradova 1977a, with changes.

ordinary somatic cells in volume, but in a number of organisms they accumulate nutrients and increase significantly in size. At the same time, the production of a large number of small spores is typical for most protonemal multicellular organisms. Particularly impressive examples are demonstrated by some genera of red algae: each sporophyte produces about 12 million carpospores, and one tetrasporophyte produces 100 million tetraspores (Vinogradova 1977a: 212).

Homologous to the process of sporulation can be considered monocytic budding of protonemal multicellular organisms. Thus, in one of the isogamous genera of brown algae, *Sphacelaria* Lyngbye, 1818, vegetative reproduction is carried out by multicellular structures formed at the ends of branches (Petrov 1977: 162). Thus, each such structure arises from a single apical cell and is similar to a multicellular spore that began its development while still on the mother's body (Fig. 9). Similarly, monocytic brood buds are formed in the gametophytes of some ferns (Gladkova 1978: 222–223).
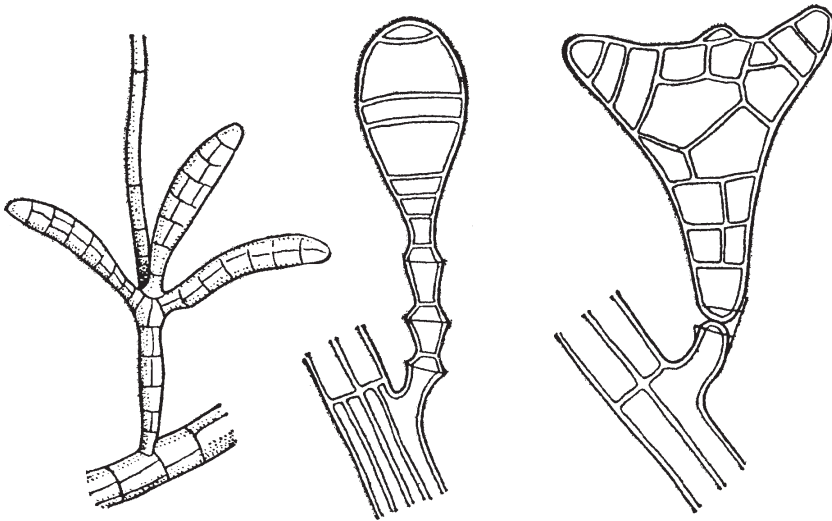
Figure 9. Monocytic budding in protonemal multicellularity, on the example of species of the genus *Sphacelaria* Lyngbye, 1818 (after Petrov 1977, with changes).

Monocytic reproduction of siphonoseptal multicellular organisms

The sexual process is predominantly isogamous or heterogamous. However, in some groups, archaic oogamy occurs (without an increase in the size of the eggs in comparison with the original cells of the mother's body), analogues of oogamy, or somatogamy (fusion of two somatic cells).

Thus, in Chytridiomycota of the order Monoblepharidales, the multinuclear mycelium usually does not contain septa, but zoosporangia, oogonia, and antheridia are separated from the body by septae (Belyakova et al. 2006a: 157). Each oogonium produces one or more small ova (Fig. 10). Uniflagellated spermatozoa fertilize the egg
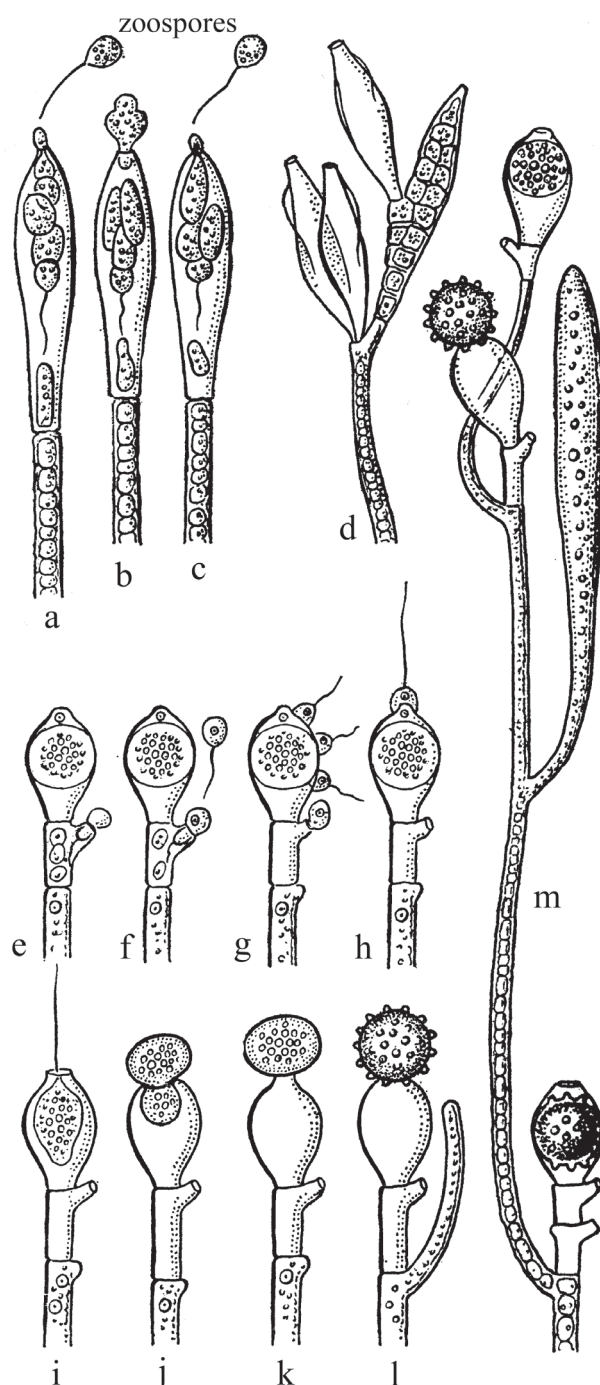
Figure 10. Asexual and sexual reproduction in siphonoseptal multicellularity on the example of *Monoblepharis* spp. **a–c** asexual reproduction by zoospores **d** branching of zoosporangia **e–l** successive stages of the sexual process **m** section of the siphon-septal body with genital organs and zygotes (after Sparrow 1933 and Sizova 1976, with changes).

inside the oogonium. The zygotes retain amoeboid movement, or they can move at the expense of one flagellum left from the fusion with the spermatozoon. After leaving the oogonium or inside it, the zygote is covered with protective layers and is at rest for some time. Subsequently, a multinucleated hypha develops from such a zygote (Sparrow 1933; Sizova 1976: 32–34).

In some genera of yellow-green algae (Xanthophyceae or Tribophyceae) from the order Vaucheriales, multinuclear branching siphons are usually devoid of septae, but sporangia and gametangia are separated by septae. In *Vaucheria* de Candolle, 1801 and *Pseudodichotomosiphon* Yamada, 1934, a single small egg cell is formed in the spherical oogonium, while numerous biflagellated spermatozoa are formed in the antheridia. The sperm enters the oogonium through a pore in the membrane. The diploid zygote, after a dormant period, germinates into a new thallus (Kobara and Chihara 1984; Belyakova et al. 2006b: 109–110).

In green algae of the genus *Sphaeroplea* Agardh, 1824 (Sphaeropleales), multinuclear siphonal bodies are separated by centripetally formed septae, but no special organs of sexual reproduction are formed. Small eggs and spermatozoa are formed in any of the segments of the body (Fritsch 1929). Fertilization is internal. The zygotes dress in dense shells and leave the mother thread only after the destruction of the latter. During the germination of the zygote, 4 zoospores are formed, each of which then gives rise to a new siphonoseptal plant (Vinogradova 1977b: 294).

An analogue of archaic oogamy in siphonoseptal multicellular organisms can be considered sexual reproduction of multicellular ascomycetes (Ascomycota), basidiomycetes (Basidiomycota), oomycetes (Oomycota) and some others. True oogamy is completely absent in these organisms. Instead, there are different variants of fusion of hyphae segments that function as unicellular gametangia, as well as the formation of male gametes (spermatia) in the absence of female ones, or vice versa (Belyakova et al. 2006a: 78).

Asexual monocytic reproduction of siphonoseptal multicellular organisms is usually carried out by small unicellular zoospores or immobile, including multicellular, aplanospores (ascospores, basidiospores, etc.). The formation of unicellular or multicellular conidia in many groups of fungi is considered a special variant of spore formation (Belyakova et al. 2006a: 77).

Monocytic reproduction of embryogenic multicellular organisms

Monocytic reproduction of embryogenic organisms in all known cases is associated either with accumulative oogamy (during the sexual process or parthenogenesis), or with accumulative aplanosporia (during asexual reproduction). In both cases, a new organism begins to develop from one immobile cell, which is larger than the usual somatic cells of the organism. The division of such a cell occurs according to the type of palintomy, i.e. rapidly recurring karyokinesis and cytokinesis, without a period of growth of daughter cells, or by the type of syntomy, i.e. rapidly recurring karyokinesis followed by simultaneous division of all cytoplasm of the mother cell into numerous

compartments (Fig. 3). As a result of such divisions, a special phase of ontogenesis appears — the embryo, as a fundamentally new biological phenomenon or an analogue of the embryo (embryoid), when it comes to development from a spore. Embryos (embryoids), unlike seedlings of protonemal and siphonoseptal multicellular organisms, are not capable of independent life; they are completely dependent on the supply of nutrients from the mother's body or on the reserves of substances accumulated by the mother in the form of a "yolk" inside the egg. Thus, the following definitions of terms can be proposed:

Embryo — the initial stage of development of embryogenic multicellular organisms, from the first division of the zygote (or parthenogenetic egg) to the beginning of independent life (exit from the shell of the zygote (egg) or separation from the mother's organism).

Embryoid — an analogue (in some cases also a homologue) of an embryo, the initial stage of development of embryogenic multicellular organisms during asexual monocytic reproduction, from the first divisions of the original cell (spore or spore-like cell) to the beginning of independent life.

Probably the most simple embryogenesis is saved in Charophyceae algae. Their zygote ("oospore"), while still inside its shell, undergoes syntomic divisions of meiosis, resulting in the formation of a four-nuclear cell. One of these haploid nuclei is separated by a septum, undergoes another palintomic division, and gives rise to root and stem cells. The remaining trinuclear cell performs the function of storing nutrients (Gollerbach 1977: 347). Thus, the embryonic stage is represented here by only a few cells. After reaching this stage, the zygote shell opens and intensive postembryonic growth of the root and stem begins in a monotomic way.

Oogamy and embryogenesis are most complex in animals (Metazoa). In them, unlike plant and fungal organisms, the process of female meiosis is in the nature of unequal division, resulting in the formation of one egg and several reductional (polar) bodies. The evolutionary meaning of this phenomenon is not clear, since it is not known in other embryogenic multicellular organisms. It is worth mentioning that in some animals, for reasons that are not entirely clear, male meiosis is also unequal (Hodgson 1997; Swallow and Wilkinson 2002; Gavrilov-Zimin et al. 2015; Gavrilov-Zimin 2018: 25–33).

In addition, a special variant of embryogenesis "a cell within a cell", known for a number of secondarily simplified parasitic animals (myxosoporidia, cnidarians of the genus *Polypodium* Ussov, 1885, orthonectids, dicyemids), as well as for angiosperms (Magnoliopsida) looks unclear in terms of evolutionary meaning and causes of occurrence.

For a century and a half, myxosporea (Mixozoa) were considered by most zoologists as one of the groups of protists (Pugachev, Podlipaev 2007: 1045–1080), although the hypothesis of their belonging to multicellular animals (Metazoa) was first put forward as early as 1899 (Štolc 1899). In recent years, on the contrary, the point of view has become generally accepted that myxosporea are descendants of cnidarians (Cnidaria), greatly simplified due to parasitism. As arguments for the multicellular-

ity of myxosporea, their structural and biochemical features, as well as comparison of nucleotide sequences of genes are given (Siddall et al. 1995; Foox and Siddall 2015). The reproductive features of myxosporea and the initial stages of their ontogeny, at first glance, seem unique and mysterious (see reviews: Feist et al. 2015; Gruhl and Okamura 2015). However, for all its aberration, the reproduction of myxosporea fundamentally corresponds to the complex embryogenic reproduction of multicellular organisms: 1) during meiosis, non-flagellated haploid gametes are formed, which can be considered eggs, since meiosis proceeds according to the female type, with the formation of polar bodies; 2) diploidy is restored parthenogenetically, although the details of this process remain largely unclear; 3) the parthenogenetic zygote splits palintomically, like the zygote of embryogenic multicellular organisms, although without the usual stages of blastulation and gastrulation; 4) as a result a multicellular (with cell differentiation) dispersal stage of the life cycle, called "actinospore" is formed; 5) inside the body of this dispersal stage, the so-called "sporoplasms" are formed asexually, giving rise to asexual generations.

In an even more aberrant, but still insufficiently studied way, the reproductive system functions in another parasitic representative of the coelenterates, *Polypodium hydriforme* Ussov, 1885 (Cnidaria: Polypodiozoa). The sexual generation of this species is a free-swimming, dioecious freshwater jellyfish. Their female gonads are appeared during ontogenesis, but do not function. The male gonads produce non-flagellate, binuclear gametes that inexplicably enter the oocytes of the host organism (fish). Further, these gametes, without fertilization, proceed to unequal cleavage, as a result of which a kind of morula is formed, which is placed inside a large polyploid cell called a trophamnion. Embryogenesis lasts several years, in accordance with the development of the host's oocytes, and ends with the formation of a larva that looks like an inside-out planula. Numerous "buds" are formed on the body of this larva and the whole organism takes the form of a stolon. Shortly before host spawning, the stolon inside the oocyte turns inside out and acquires the normal position of the cell layers for the coelenterates. The release of stolons from the eggs of the host occurs in the reproductive ducts of the fish. After entering the water, the polypodium stolon undergoes fragmentation with the formation of daughter medusoid forms (Raikova 1994).

The body of dicyemides (Dicyemida) is arranged in an extremely simplified way, consists of only 8–40 cells and does not have any tissues, organs and intercellular cavities. The total number of cells is determined and characteristic of each species. The body of adult worm-like stages (nematogen and rhombogen) is formed by one layer of integumentary (somatic) ciliated cells and one (rarely several) large internal axial (generative) cell (Fig. 11) with a polyploid nucleus. In the internal cytoplasmic chambers of this cell, smaller cells are located — axoblasts, which give rise to individuals of the next generations (McConnaughey 1951; McConnaughey and McConnaughey 1954; Ivanova-Kazas 1975; Malakhov 1990; Furuya and Tsuneki 2003; Furuya et al. 2003; Noto et al. 2003). Reproduction is carried out by parthenogenetic and bisexual methods. During parthenogenesis, the axoblast undergoes irregular mitotic divisions such that the larger cell (macromere) becomes the new axial cell, and the micromere

Transverse section of nematogen
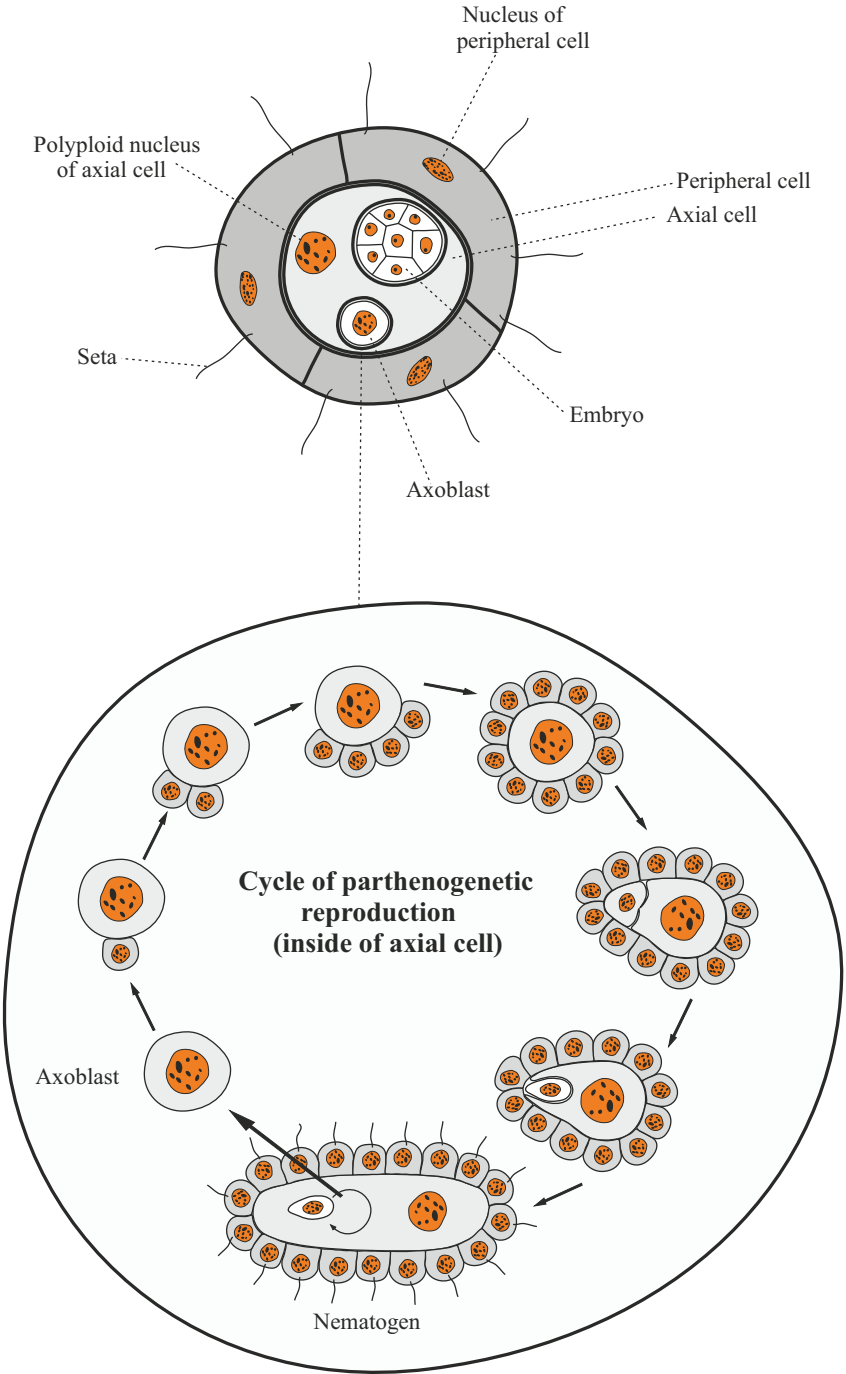
Figure 11. The structure of the nematogen and the cycle of parthenogenetic reproduction of Dicyemida.

continues to divide, resulting in the formation of the next generation nematogen cover layer. Then the macromere stretches out and undergoes another uneven division; the smaller of the two daughter cells becomes a new axoblast and invades the cytoplasm of the larger cell. After these processes, the young nematogen leaves the parent individual, squeezing between its cells. In different species of dicyemides, from one or two to more than a hundred daughter nematogens can simultaneously develop in the axial cell (McConnaughey 1951; McConnaughey and McConnaughey 1954; Ivanova-Kazas 1975; Malakhov 1990; Furuya and Tsuneki 2003; Furuya et al. 2003). Thus, in dicyemids, the initial stage of reproduction is associated with a single cell (axoblast), which can be called an agamete or pseudogamete. In this case, one can speak of a special variant of ameiotic parthenogenesis, in which not only fertilization, but also reductional division of oocytes is absent (see Ivanova-Kazas 1975: 100). An adult dicyemid that performs sexual reproduction is called a rhombogen. Morpho-anatomically, this individual does not differ much from a nematogen, but in its axial cell, axoblasts form hermaphrodite "gonads", called infusorigens. In this case, after the first unequal division of the axoblast, the micromere separates from the macromere, loses its cytoplasm and remains in the axial cell of the rhombogen in the form of a free nucleus called the paranucleus (Fig. 12). As a result of the accumulation of paranuclei originating from many embryos, the axial cell of the rhombogen gradually becomes multinucleated. As a result of subsequent unequal divisions of the macromere, micromeres are again separated from it, which give rise to oogonia and spermatogonia, and the original macromere itself becomes an axial cell of the infusorigen. Oogonia are located along the periphery, and spermatogonia are inside the axial cell as a result of the invagination of one of the micromeres into its cytoplasm and its subsequent divisions. Spermatogonial cells, after passing through meiotic divisions, lose most of their cytoplasm and form non-flagellate sperms from them. These sperms penetrate into the oocytes located near the axial cell and fertilize them. From the zygote, a mobile larval stage is formed — infusoriform, which leaves the maternal rhombogen, and then the kidney of the host organism (mollusk) and enters the external aquatic environment (McConnaughey 1951; McConnaughey and McConnaughey 1954; Ivanova-Kazas 1975; Malakhov 1990; Furuya and Tsuneki 2003; Furuya et al. 2003). The way infusoriform penetrates into a new host mollusk remains still insufficiently studied, but it is assumed that the so-called "urn cells" of infusoriform give rise to a two-cell "embryo" that grows into a founder nematogen, and that, in turn, produces new nematogens or rhombogens. From the above information, it becomes clear that the development of new nematogens, rhombogens and infusoriform larvae of dicyemides occurs entirely within the maternal organism, from which they receive the necessary nutrition. Thus, dicyemides as a whole, as a taxonomic group, should be considered viviparous organisms.

In orthonectids (Orthonectida), the main stage of the life cycle (Fig. 13) is a multinuclear plasmodium located in the tissues of the host organism. Plasmodium does not have a definite shape, and daughter plasmodia can be formed from different parts of its surface by simple budding. Inside the plasmodia, in addition to the trophic nuclei, there are generative nuclei with isolated sections of the cytoplasm, which are agametes.
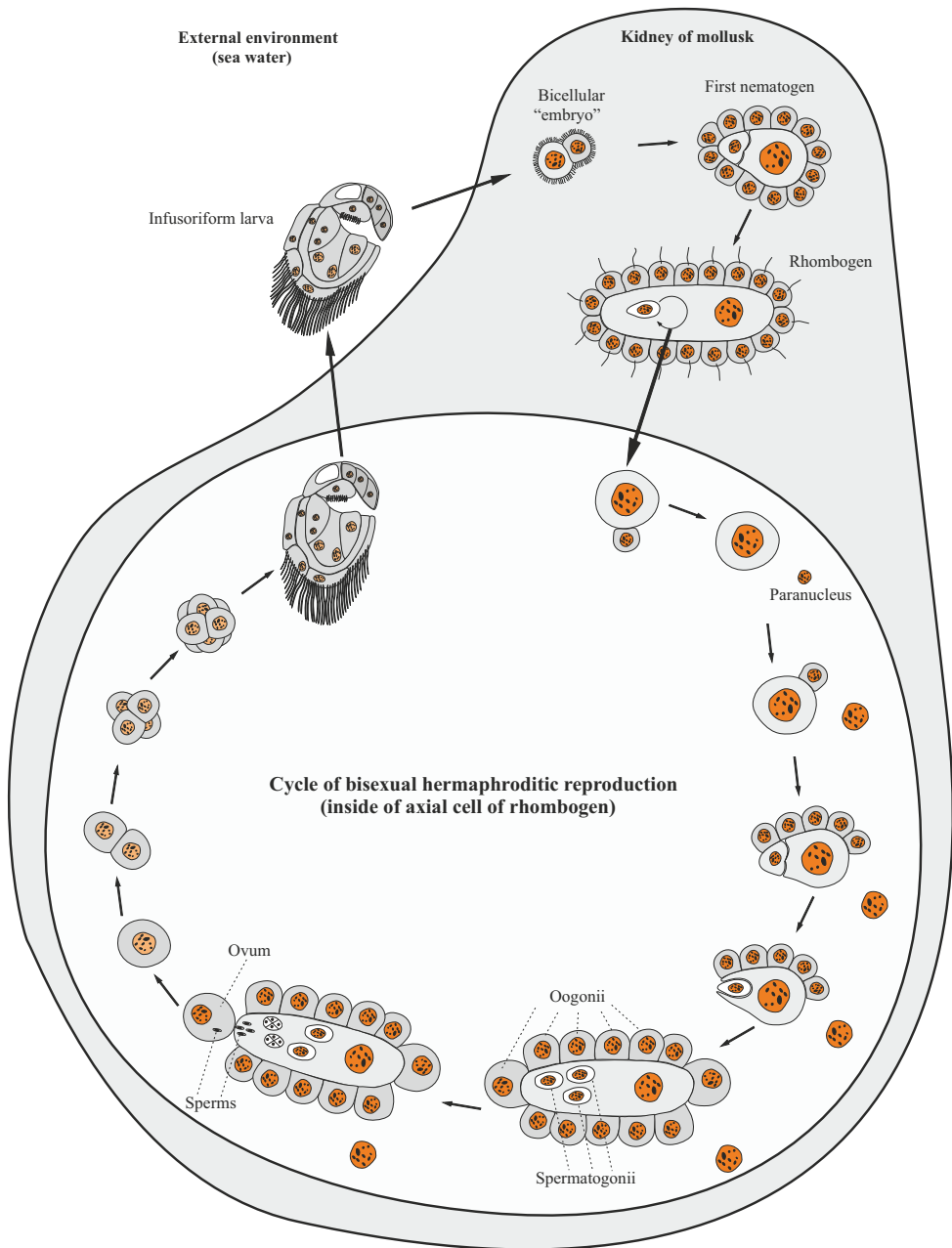
Figure 12. The cycle of sexual hermaphrodite reproduction of Dicyemida.

These agametes without fertilization undergo uneven fragmentation with subsequent gastrulation by the type of delamination (Ivanova-Kazas 1975: 92; Malakhov 1990: 49) and the formation of adults, bypassing the larval stage. As a result of this kind of parthenogenesis, males and females of the next (bisexual generation) are formed inside
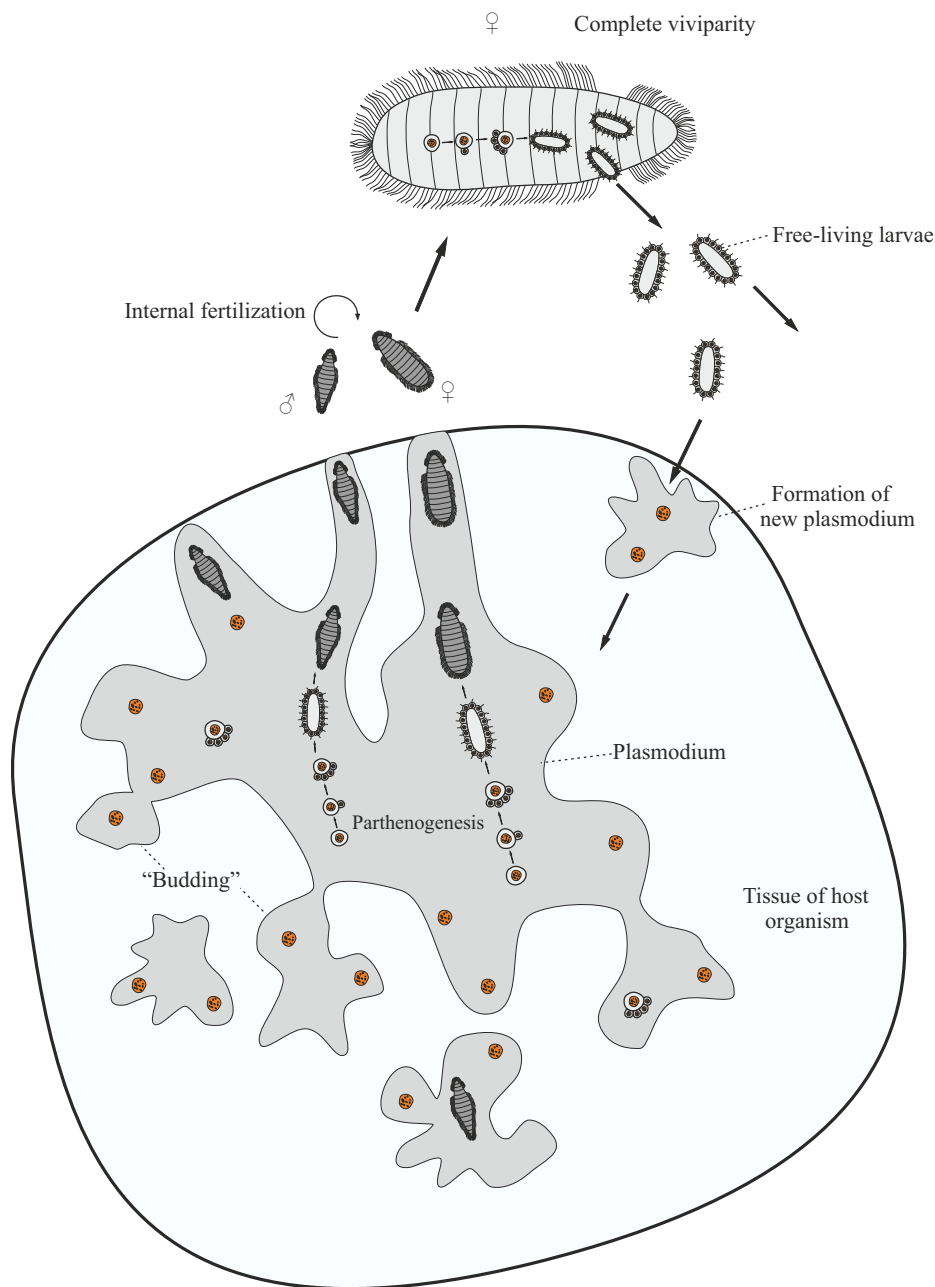
Figure 13. Generalized scheme of the life cycle of Orthonectida.

the plasmodia, and in different species, individuals of both sexes can form inside one plasmodium or in different plasmodia. Sexual individuals of orthonectids are not capable of self-feeding, but have significant mobility due to numerous setae covering the

body. Eggs and spermatozoa are formed in sexual individuals even while they are in the maternal plasmodium (Slyusarev 2008: 421). Individuals ready for bisexual reproduction emerge from the plasmodia along special outgrowths directed to the surface of the host's body and enter the surrounding sea water, where fertilization takes place: spermatozoa enter the water through a special genital opening on the body of the male and then enter the body of the female through her genital opening. Fertilized eggs undergo unequal cleavage, resulting in the formation of a morula-like embryo. From the embryo, a small (body size is about 15 microns), covered with setae, larva is formed, which leaves the mother's body through the genital opening. After swimming freely in sea water, the larva enters the body of a new host, where its outer ciliated layer of cells is shed, and one or more next-generation plasmodia are formed from internal cells (Ivanova-Kazas 1975: 92; Malakhov 1990: 49; Slyusarev 2008). Thus, orthonectids demonstrate asexual reproduction (by budding of plasmodia), ameiotic parthenogenesis, and bisexual reproduction, in which individuals of the new generation develop completely inside the maternal organism (complete viviparity).

The formation of the embryo sac in angiosperm plants (Magnoliopsida) and the processes of embryo and endosperm formation occurring inside this single multinucleated cell are so well known that there is no need to dwell on them in more detail here. However, it is worth noting the remarkable and rather strange circumstance that among plants, embryogenesis according to the "cell within a cell" type appears only in this youngest, evolutionarily advanced group, while the analogous examples listed above among animals are characteristic only of very simply organized groups that have passed to parasitism.

Synchronization of copulative processes

Based on the well-known evolutionary advantages of sexual reproduction over asexual reproduction (see, for example, Felsenstein 1974), in order to increase variability and biological diversity, it is theoretically more advantageous for a living organism to reproduce by gametes and their fusion products (zygotes) than by spores. It is well known that the onset of the reproductive period in the life cycle of a particular species of living organisms is somehow coordinated with various environmental factors (see, for example, a review on algae in Brawley and Johnson 1992). However, for reproduction by zygotes, it is necessary not only to synchronize the period of gametogenesis, but also the very moment of gamete release in different individuals of the population, so that "male" and "female" gametes meet in a certain place in space at the same time. The distances that the gametes themselves are able to overcome solely due to cellular movements are very short. This issue has been studied in detail, for example, in diatoms (Davidovich et al. 2012; Davidovich 2019: 151–162). Their "male" gametes, due to flagellar activity, amoeboid movement, or the formation of special cytoplasmic filaments, can move at distances only several times (rarely ten times) greater than the diameter of the gametes themselves. The movement of "male" gametes is chaotic in this case, and the "female" gametes of diatoms are not at all capable of active movement. Thus, for copulation it is necessary that the pa-

rental individuals are in close proximity to each other. Often, even as a result of the close but unfortunate location of parental individuals, their gametes cannot merge with each other (Davidovich 2019: 152). However, in general, for microscopic unicellular organisms that form dense populations, the synchronization of gamete release does not seem to be a significant problem (Brawley and Johnson 1992). The gametogenesis of unicellular organisms is just a direct transformation of an "adult" unicellular organism into one or several gametes, takes a relatively short time, comparable to the viability of gametes, and directly depends on the onset of external factors (the same for all individuals of the population), and the resulting gametes usually similar in size to adults. Thus, at the same point in space at the same time, numerous gametes capable of fusion inevitably appear. In addition, some unicellular organisms form a "syzygy" before copulation, within the closed shell of which fusion takes place (Fig. 5), or the union of gametes occurs in the cramped space of the internal cavities of the body of the host organism (in parasitic life cycles).

The situation is quite different in multicellular organisms. First, due to the increase in the size of their bodies, each multicellular organism occupies a place in space that is many times greater than the size of the gametes it produces. Secondly, before the start of gametogenesis, such an organism must reach the complex multicellular stage of the "vegetative" body (see the first reproductive criterion of multicellularity above), which creates a certain (often very significant) individual variation in terms of readiness for the sexual process and maturation of gametes. Thirdly, the appearance of oogamy in multicellular organisms leads to the fact that only male gametes retain their own mobility (and sometimes it is lost in gametes of both sexes). Considering all these features, it is possible to achieve cross-fertilization of gametes at the multicellular level of organization in the following ways: 1) by keeping immobile female gametes in the body of the mother's organism until they are found by spermatozoa (internal fertilization in the broad sense); 2) forcibly ejecting female gametes into the external environment synchronously with the ejection of spermatozoa by other individuals of the population (external fertilization); 3) providing a passive release of numerous gametes into the external environment at a strictly defined time (also external fertilization).

The first way, undoubtedly, turns out to be technically simpler and is implemented independently in the vast majority of groups of archaic multicellular organisms. Thus, in the vast majority of sponges, in trichoplax, in archaic turbellarians, in extremely simplified orthonectids and dicyemids, in multicellular fungi, in volvox, in most oogamous multicellular algae, as well as in all higher plants, internal fertilization of the egg occurs, and the initial stages of embryogenesis take place inside the body of the mother organism, or the zygote becomes a resting stage and finds itself in the external environment after the death and disintegration of the mother's body. A clear understanding of this circumstance allows us to answer the age-old question of classical biology about whether for animals and other multicellular organisms the original method of reproduction was external fertilization with the corresponding complete development of the daughter organism in the external aquatic environment. In many old and modern general theoretical works, this was taken for granted so much that it was not even specially argued (see, for example, Franz 1924; Ivanov 1968; Ivanova-Kazas 1995; Mikhailov et

al. 2009; Malakhov et al. 2019, etc.). According to such ideas, in hypothetical diagrams of the origin of the first Metazoa from colonial choanoflagellates (Choanoflagellata), it is usually drawn that the eggs somehow get into the water and are fertilized by spermatozoa there (Fig. 14). However, already from the fact that all animals and, in general, all organisms with embryogenic multicellularity are characterized by obligate oogamy, it follows that the female gametes themselves cannot get out of the multicellular body or even from the colony of protists in any way, but the genital ducts, muscles and nervous system, which would regulate the forced release of gametes, are absent yet in archaic organisms; all these apomorphies appear in evolution, starting from the level of organization of coelenterates and above, i.e. in Eumetazoa (for a more detailed analysis, see Gavrilov-Zimin 2022). It remains theoretically possible to passively release oogametes through a simple rupture of the body wall or the cell wall of some conditional "gametangium", according to the principle of opening an abscess. In this way, for example, various gametangia and sporangia are opened in lower and higher plants. However, this method is not entirely suitable for ensuring fertilization, since each abscess or sporangium is opened at the moment when it is ripe, not in accordance with other abscesses on the body of the same organism, and even more so of another organism. Therefore, contrary to popular belief, all the most archaic multicellular organisms are characterized by internal fertilization of the egg directly inside the body or on the body of the mother's organism. This means the inevitability of initial viviparity in Metazoa and sporophyte germination on gametophyte in plants (see: Gavrilov-Zimin 2022). That is, the scheme for the origin of animals should look like this (Fig. 15). At first, in some choanoflagellates, the immobile zygote underwent divisions, remaining inside the colony, and the products of this division, mobile zoospores, left the colonies, floated away and gave rise to new organisms. Then the colonies began to grow, the mobility of a single zoospore was no longer enough to leave the huge colony and sail somewhere to a new place. Under these conditions, the appearance of a synzoospore (i.e., a product of zygote cleavage that has not disintegrated into parts), a hypothetical stage of development used by Zakhvatkin (1949; 1956) in his theory of the emergence of multicellular Metazoa, turns out to be logical. In fact, the synzoospore was the first larval stage in the evolution of organisms, which was unable to feed itself yet, but ensured distribution. Such variant of reproduction/development is known, for example, in modern sponges (Porifera).
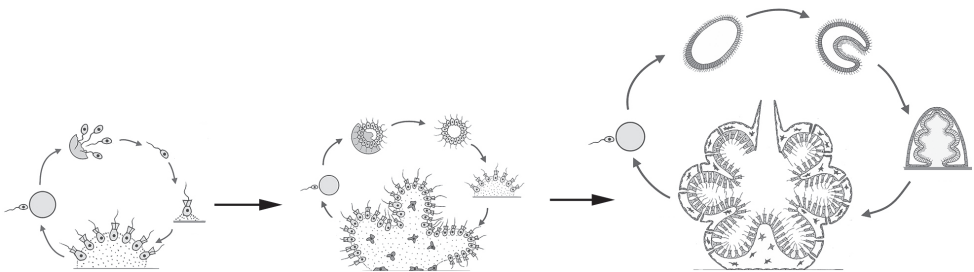
Figure 14. Graphical interpretation of the "sedentary" hypothesis of the origin of multicellular animals under the assumption of the initial external fertilization (according to Malakhov et al. 2019, with changes).
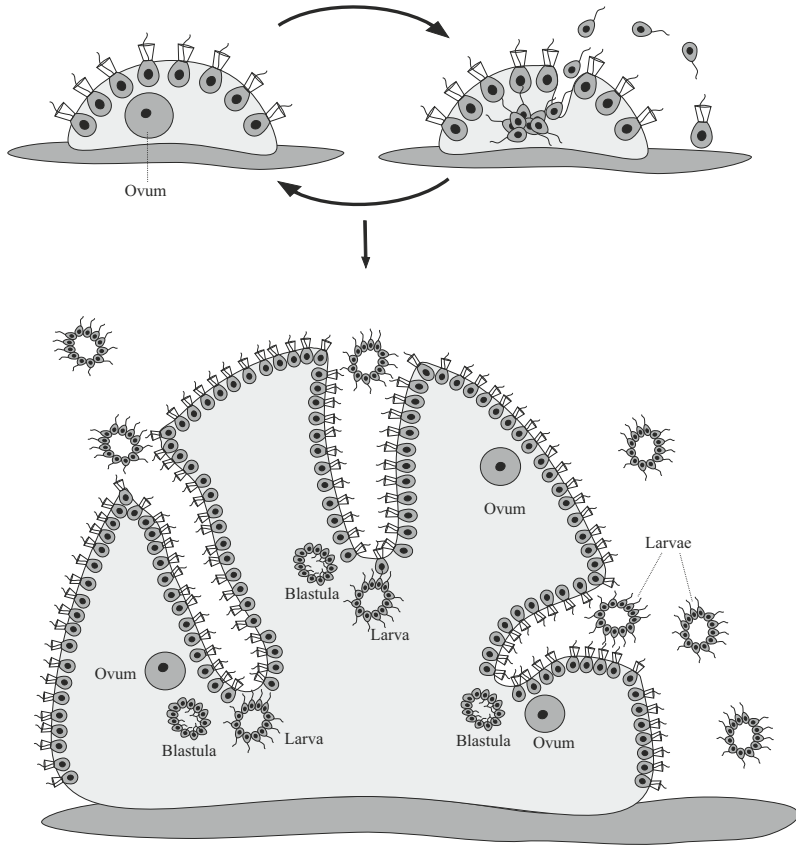
Figure 15. Scheme of the origin of multicellular animals (Metazoa), based on the hypothesis of primary viviparity. Maximal figure size, please!

In some oogamous multicellular algae, due to the simplicity of their structure, it turns out to be rather difficult to draw a clear line between fertilization inside the mother's body and on its surface. Thus, in Laminariales brown algae, the egg is released from the oogonium before fertilization, but remains attached to its edges. The zygote germinates without detaching from the maternal gametophyte. If, due to random events, the egg or zygote loses its connection with the mother plant, then differentiation processes are disrupted during germination and the resulting defective thallus soon dies (Belyakova et al. 2006b: 133–134).

External fertilization is well known and studied in many groups of marine and freshwater animals, in which this process is ensured by the presence of the nervous system, sensory organs, muscles, and various genital ducts. Receiving a certain signal (visual, tactile, chemical) from each other, sexual partners implement a forced synchronous release of gametes into the external aquatic environment. However, for the most archaic animals and plants, the only possible way is the passive release of gametes,

in particular, through the rupture of the shells of the "gametangia", synchronized by external causes. A comparative analysis of the reproductive strategies of various multicellular organisms shows that it is extremely difficult to achieve synchrony with the passive variant. This path was realized only in a few small groups of marine organisms, strictly synchronized in their reproductive activity with the lunar cycles and/or the corresponding periodicity of tides. In this case, the passive release of gametes is technically provided in two different methods, but both of them are associated with significant limitations and remain evolutionarily dead ends.

The first method is known in some highly developed Demospongiae, which are built according to the progressive "leucon" type and reach large body sizes. The structure of the body allows these sponges to regulate the flow of water passing through the body and carrying out a large number of immobile eggs and motile spermatozoa (Reiswig 1970a, b; 1976; Ereskovsky 2005: 55–59). However, in most cases it remains unknown whether the release of eggs from the body occurs before fertilization or after fertilization (Reiswig 1976: 104). Sponges have no formed gonads; female germ cells are located diffusely or in small groups among the somatic cells of the body, and spermatozoa are collected inside temporary formations – spermatocysts (Ivanova-Kazas 1975; Ereskovsky 2005). This circumstance undoubtedly facilitates the task of excreting gametes with water flows, since there is no need to open the shell of the conditional "gametangium", which is present in most other multicellular organisms. Pheromones probably act as an additional regulator of the synchronicity of gamete release in sponges.

The second way of passive synchronous release of gametes is implemented in a number of genera of brown and green algae. They are unable to regulate water flows, but their reproduction is coordinated in a complex way with the lunar cycle and tidal rhythms (Smith 1947; Brawley 1992; Brawley and Johnson 1992; Feis 2010; Heesch et al. 2021). This synchronization is best studied in various *Fucus* spp. Unlike the vast majority of other plants, meiosis in *Fucus* spp. occurs during the formation of eggs and spermatozoa, there is no gametophyte stage and the haploid phase is represented only by gametes (Fig. 16), similar to how it occurs in the life cycle of animals. The immobile eggs are released into the water simply through a break in the wall of the gametangium and then settle to the bottom, while the motile spermatozoa find them due to pheromones that act at a distance of only a micrometer to a millimeter (Serrão et al. 1996; Feis 2010). That is, fertilization is possible only between algae located next to each other. The synchronism of the release of gametes is achieved due to the fact that the gametangia dry up at low tide, and then massively burst upon repeated wetting and/or changes in salinity during high tide (the so-called osmotic stress). At the same time, it is necessary that the tide be calm, without strong waves that can spread the gametes in different directions. It is clear that this method of bisexual reproduction, based on a combination of many specific external causes, is not suitable for most other living organisms. However, the superficial similarity of the reproductive biology of fucuses with the reproductive behavior of oviparous animals even gave rise at one time to the hypothesis of the origin of the Metazoa directly from fucus-like ancestors (Franz 1924). Extremely accurate synchronization of maturation and excretion of gametes due to
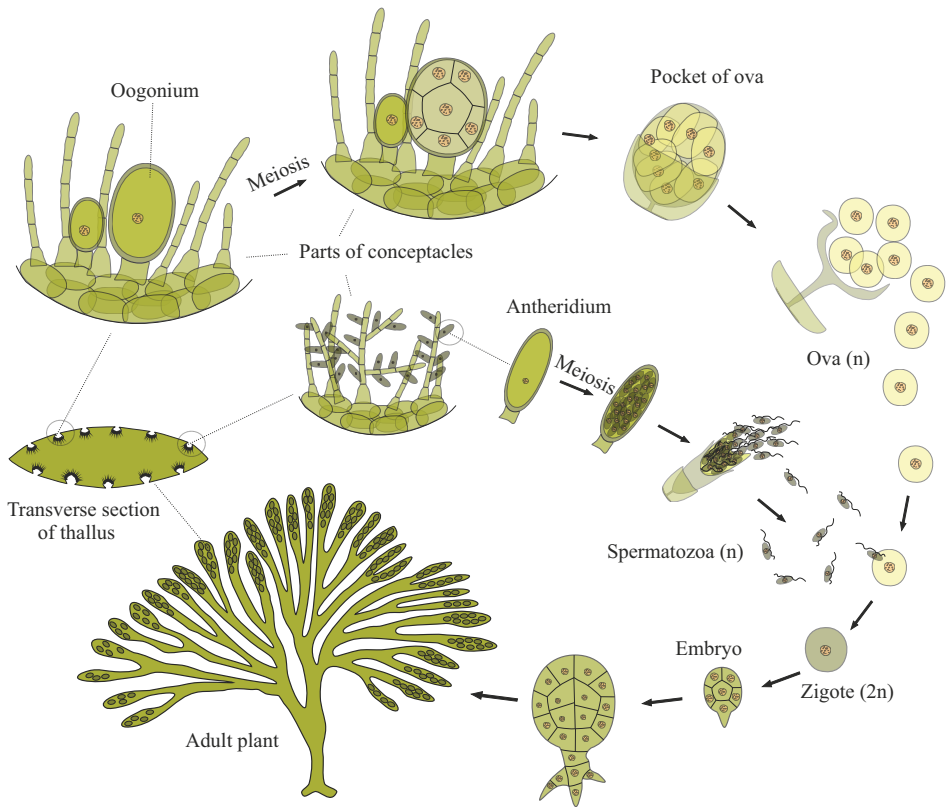
Figure 16. Generalized diagram of the life cycle of *Fucus* spp.

tidal rhythms is also known in those algae that retain the alternation of gametophyte/sporophyte in their cycle. In this regard, various species of brown algae of the genus *Dictyota* Lamouroux, 1809 (see Hoyt 1927; Bogaert et al. 2020) and green algae of the genus *Ulva* Linnaeus, 1753 (see Smith 1947) are the most studied.

In laboratory conditions, it is very often possible to achieve synchronous opening of gametangia due to a sharp change in illumination (see review in Brawley and Johnson 1992: 237–238), but such studies were again carried out on algae, whose reproduction under natural conditions is confined to tidal cycles.

There does not seem to be any other effective means of precise synchronization of gamete release, apart from tidal, in multicellular plants. Understanding this, one can offer an explanation for why plants do not have egg-laying, similar to that of animals, and why asexual reproduction with sporophyte/gametophyte alternation absolutely predominates in plants, despite the obvious evolutionary advantage of bisexual reproduction and the diploid state of the multicellular body. The answer lies in the fact that plants are not able to independently release eggs into the external environment synchronously with spermatozoa. Their eggs in the vast majority of cases remain on the mother's body, wait until spermatozoa (or sperm) reach them in one way or another, and then

germinate inside or on the body of the mother's body. In this case, reproduction and distribution are not provided by gametes or zygotes, but by spores, since no synchronization is required for this at all. Up to the highest stages of plant evolution, they fail to switch to normal independent sexual reproduction, and most flowering plants in their sexual process are also completely dependent on animals, especially pollinating insects. There are examples of plant gamete transfer in some marine plants, for example, in some red algae, for which crustaceans act as pollinators (Ollerton and Ren 2022).

In animals, on the contrary, bisexual reproduction absolutely predominates, and synchronization of the release of gametes is achieved at fairly early stages of their evolution, starting with the most complexly organized sponges and coelenterates. The latter develop a simple nervous system, gonads, and musculature, in particular, a muscular intestine/stomach, through which, in the simplest case, sexual products are excreted. Some ctenophores (Ctenophora) even have specialized reproductive ducts (Beklemishev 1964: 334). As a result, both among cnidarians (Cnidaria) and among ctenophores (Ctenophora), external fertilization with the development of eggs in the external environment prevails in the vast majority of species, and only a few species retain viviparity or ovoviviparity. All further evolution of the reproductive sphere of animals is the constant improvement of the reproductive ducts, gonads, external ovipositors and copulatory organs, and the very methods of laying eggs protected by shells into the external environment. Separate aberrations of the reproductive system, leading in some groups, small in diversity, to secondary vivparity, I have considered in detail in a special article (Gavrilov-Zimin 2022).

Policytic reproduction

The simplest version of polycytic reproduction, which consists in restoring the whole body from separate fragments, is observed in almost all archaic multicellular organisms and probably represents the original (plesiomorphic) method of polycytic reproduction for most phylogenetic lines. Despite its extreme archaism, the ability to restore the whole body from fragments is retained during the entire further evolution in most groups of plants, including the most highly developed angiosperms (Magnoliopsida), as well as in most fungi. On the contrary, among animals, this method remains possible only in organisms that are at a relatively low level of morpho-anatomical organization: sponges (Porifera), coelenterates (Coelenterata), various taxa of flatworms (Plathelminthes), some nemerteans (Nemertini), and annelids (Annelida). A somewhat more complicated version of fragmentation can be considered the division of the body in two by lacing or splitting. Such methods are known, for example, in trichoplax, some coelenterates and flatworms. At the same time, division without previous morphogenetic preparation (architomy) and division after preliminary doubling of body parts (paratomy) are distinguished – see, for example, Zakhvatkin (1949: 171).

An apomorphic feature inherent in some protonemal and embryogenic multicellular organisms, as well as representatives of "complex" organisms – lichens (Lecanoromycetes) — can be considered the appearance in them of a special polycytic budding (= blas-

togenesis), as a result of which specialized outgrowths are regularly formed from groups of somatic cells, over time, separating and growing into independent individuals. In many groups of organisms, such polycytic budding occupies a strictly defined place in the life cycle or even represents the main way of reproduction and distribution in space. So, in many highly organized representatives of lichens, polycytic budding is the only way of reproduction (not counting accidental fragmentation of the body). This process is carried out through the formation of the so-called soredia and isidia (Fig. 17) – microscopic multicellular outgrowths of the thallus that combine symbiotic fungal hyphae and algae cells (Golubkova 1977: 419; Belyakova et al. 2006a: 224–226).

Polycytic budding is highly developed in Charophyceae s.s. and is provided by special nodules on rhizoids or by special "stellate cell clusters" (Belyakova et al. 2006b: 270).

A significant diversity of polycytic "brood bodies" is observed in gametophytes of various liverworts (Marchantiophyta) and mosses (Bryophyta) (Abramov and Abramova 1978: 65–66; 81–82; Potemkin and Sofronova 2009: 30). Brood nodules and buds are known in sporophytes of some Lycopodiophyta (Filin 1978: 106, 114) and Psilotopsida (Timonin and Filin 2009: 298). The sporophytes of many horsetails (Equisetopsida) are characterized by the formation of numerous underground nodules (Filin 1978: 140–141). Brood nodules and buds are known in sporophytes of some species of ferns (Pteridiophytina) (Timonin and Filin 2009: 255). However, polycytic budding is most common among flowering plants (Reproductive Systems 2000: 315).

In animals, polycytic budding is widespread among sponges (Porifera), trichoplax (Fig. 18), cnidarias (Cnidaria), flatworms (Plathelmintes), camptozoa (Kamptozoa), annelids (Annelida), many tentaculata (Tentaculata) and hemichordates (Hemichordata), a number of lower chordates – tunicates (Tunicata), some species of echinoderms (Echinodermata) and, in rare examples, are known from representatives of some other groups (Ivanova-Kazas 1977, 1995).

Polyembryony can be considered a special type of polycytic budding, apomorphic for some embryogenic multicellular organisms. This term, like many others used in reproductive biology, has a rather vague meaning. In most cases (and in this article),
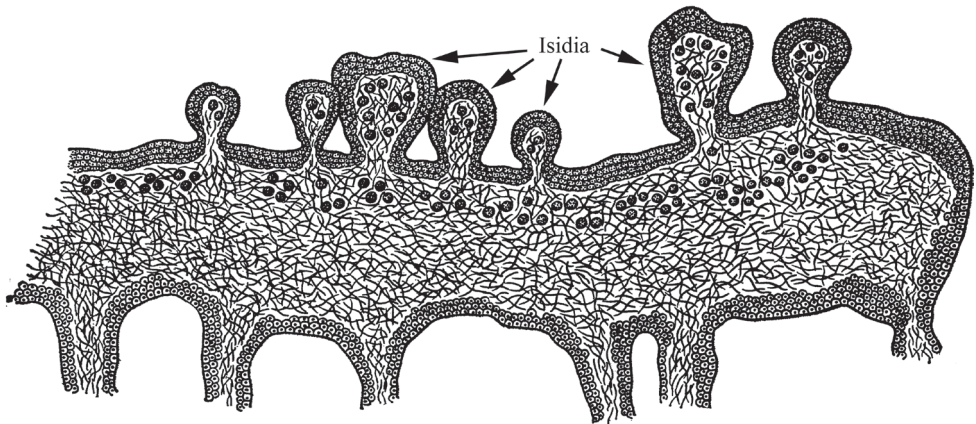
Figure 17. Polycytic budding in lichens: reproduction by isidia (after Golubkova 1977, with changes).
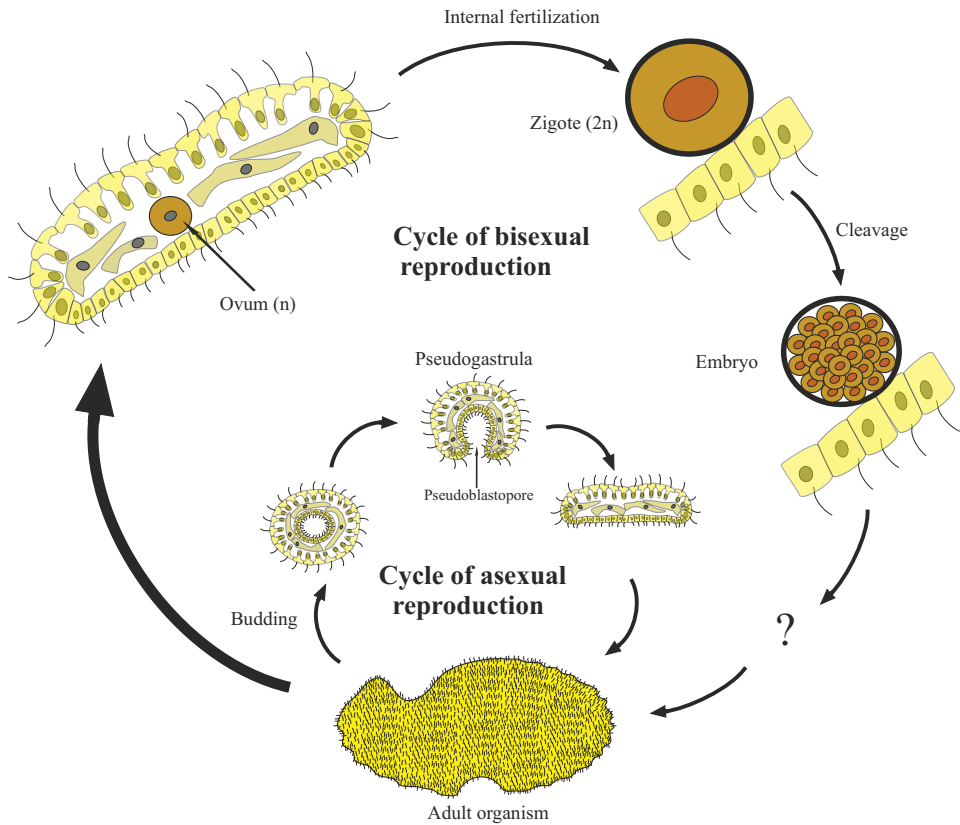
Figure 18. Scheme of the life cycle of *Trichoplax adhaerens* Schulze, 1883; asexual reproduction is provided by polycytic budding.

polyembryony means the regular division of a developing zygotic embryo into several secondary embryos (see, for example, Ivanova-Kazas 1995: 205). At the same time, in many embryogenic multicellular organisms, the division of the developing embryo into separate blastomeres is possible by chance or experimentally induced. So, for example, in some cases facultative "primary polyembryony" is noted, which manifests itself in experiments and is the cultivation of independent organisms from individual blastomeres, for example, in some hydromedusae (Zakhvatkin 1949: 217; Ivanova-Kazas 1977: 200–201). In my opinion, in this case we are not talking about polyembryony in the sense mentioned above, but only about the forced "reconstruction" of that stage of the life cycle that took place in the unicellular ancestors of the organisms under consideration, namely, the division of the zygote into separate zoospores.

Polyembryony is extremely widely understood in the literature on flowering plants (Reproductive Systems 2000: 401), where it is proposed to use this term not only for fairly rare cases of regular division of the initial zygotic embryo (as, for example, in peonies (*Paeonia* spp.), but also various cases the emergence of germ-like structures from

vegetative parts of the body. I am not ready to agree with such an expansive approach, since it generates terminological confusion.

Rare cases of polyembryony among animals are known in some genera of cyclostomes (Cyclostomatida), monogenetic flukes (Monogenea), endoparasitic hymenopterans (Hymenoptera) and Strepsiptera, as well as in mammals – armadillos of the genus *Dasypus* Linnaeus, 1758 (Ivanova-Kazas 1995: 205, 208, 257, 271, 275, 475, 480).

An extremely peculiar analogue of polyembryony can be seen in the development of the so-called "carposporophyte generation" of red floridian algae (Rhodophyta: Florideophyceae) (Fig. 8). In most floridias, the "zygote" (the fertilized "carpogon"), one way or another, merges with the "auxiliary" (nourishing) cells of the maternal thallus (gametophyte). After fusion, carpogon forms numerous multicellular processes ("gonimoblast threads"). Then all or only part of the gonimoblast cells become "carpospores". After separation from the mother plant, carpospores give rise to the next diploid generation (tetrasporophytes) (Vinogradova 1977a: 192–250; Searles 1980). If we accept the idea of an analogy of polyembryony with respect to floridias, then the need for such a complicated theoretical construction as an additional generation of sporophytes ("carposporophyte generation") disappears, and the life cycle of red algae then turns out to be quite comparable with the usual gametophyto-sporophyte cycle of other plants.

A number of groups of multicellular organisms completely lose the ability for polycytic reproduction (with the exception of the rarest cases of polyembryony mentioned above). Such, for example, are various taxa within the polyphyletic group of Nemathelmintes, echiurids (Echiurida), brachiopods (Brachiopoda), arthropods (Arthropoda), mollusks (Mollusca), vertebrates (Vertebrata). Obviously, such a loss is associated with a high degree of specialization of the tissues and organs of these organisms and the corresponding loss of totipotency in most of the somatic cells that make up their body. At the same time, the almost total absence of polycytic reproduction in gymnosperms and even in such simply organized multicellular plants as *Volvox* spp. is not entirely clear.

Conclusion

The multiple origin of multicellularity in different groups of organisms allows at the present time to give only a very approximate minimum estimate of the total number of such evolutionary events. Apparently, there were at least 50 cases of independent origin of multicellularity among eukaryotes and at least several dozens among prokaryotes. Examples of protonemal multicellularity among bacteria and algae are of particular difficulty for calculation, since the modern systems of these organisms abound in genera that simultaneously include species with simple unicellular, colonial-unicellular and obligate-multicellular bodies (see, for example, AlgaeBase: https://www.algaebase.org/). It is equally difficult to count the numerous cases of transition from siphon-unicellular to siphonoseptal multicellularity among fungi and algae, developing through the initial stage of a multinuclear "siphon". A much clearer picture emerges with regard to embryogenic multicellular organisms. Thus, there is no doubt about the sin-

gle independent appearance of animals and separately *Volvox* spp. on the basis of the corresponding ancestral spherical colonies with an internal cavity (Zakhvatkin 1949; Malakhov et al. 2019, etc.). The single occurrence of higher plants (Embryophyta) and charophyceae algae (Charophyceae s.s.) based on the preceding protonemal multicellularity of their ancestral forms is also generally accepted in the botanical literature (see, e.g., Umen 2014). It is believed that sporophytes of higher plants in all cases develop embryogenically, while gametophytes in many cases retain protonemal development. Embryogenic multicellularity among brown and red algae, apparently, arose repeatedly, but on the basis of the already achieved protonemal multicellularity of more archaic representatives of these groups (see above).

It is noteworthy that all complex multicellular organisms that have tissues and organs develop according to the type of embryogenic multicellularity based on obligate accumulative oogamy or accumulative aplanosporia. This is probably due to the well-known fact that a large volume of cytoplasm in the egg and its complex structure are very important for the initial differentiation, which then ensures the predetermination of cleavage and the formation of specific tissues and organs from certain blastomeres. For animals, in addition to the initial predetermination of cleavage, the formation of internal body cavities, in particular, the primary cavity (the blastocoel), is also important, and this, probably, cannot be achieved on the basis of protonemal or siphonoseptal development.

Summing up all of the above, I can highlight the following final suggestions:

1.   The proposed first reproductive criterion of multicellularity postulates that a unitary multicellular organism, in contrast to a colonial-unicellular organism, obligately develops as a multicellular organism and reproduces itself only after it reaches the multicellular "vegetative" stage of ontogenesis.

2.   The second reproductive criterion of multicellularity determines exactly how a multicellular body reproduces itself in course of the monocytic method and allows us to divide all known ways of implementing obligate multicellularity into three fundamentally different options: protonemal, siphonoseptal and embryogenic.

3.   The most complex, embryogenic multicellularity arises exclusively on the basis of obligate accumulative oogamy or accumulative aplanosporia, in which the gamete / spore exceeds in size (sometimes hundreds and thousands of times) the original mother cells. As a result of subsequent palintomic or syntomic divisions, an embryo or embryoid is formed from an oogamete/spore — stages of ontogenesis that are absent in other multicellular and unicellular organisms.

4.   The emergence of multicellularity, especially on the basis of oogamy, creates significant technical problems for the synchronization of copulatory processes. The simplest way out of this situation is to keep immobile female gametes in/on the body of the maternal organism until they are found by spermatozoa. This method is implemented in the vast majority of multicellular plants and fungi, as well as in the most archaic animals. In this regard, viviparity is considered as the original, plesiomorphic way of offspring in Metazoa.

Acknowledgements

References

Abramov II, Abramova AL (1978) Class liverworts (Marchantiopsida or Hepaticopsida). Class mosses (Bryopsida or Musci). In: Grushvitskiy IV, Zhilin SG (Eds) Plant life in 6 volumes. V. 4. Mosses. Clubmosses. Horsetails. Ferns. Gymnosperms. Moscow, 60–96. [In Russian]

Angel P, Herranz M, Leander BS (2021) Insights into the morphology of haplozoan parasites (Dinoflagellata) using confocal laser scanning microscopy. The Journal of Eukaryotic Microbiology 68: e12855. https://doi.org/10.1111/jeu.12855

Barr FA, Gruneberg U (2007) Cytokinesis: placing and making the final cut. Cell 131(5): 847–60. https://doi.org/10.1016/j.cell.2007.11.011.

Beklemishev VN (1964) Fundamentals of comparative anatomy of invertebrates. V. 2. Moscow, 446 pp. [In Russian]

Bell G (1978) The evolution of anysogamy. Journal of Theoretical Biology 73: 247–270. https://doi.org/10.1016/0022-5193(78)90189-3

Belyakova GA, Dyakov YuT, Tarasov KL (2006a) Botany in 4 volumes. V. 1. Algae and fungi. Moscow, 320 pp. [In Russian]

Belyakova GA, Dyakov YuT, Tarasov KL (2006b) Botany in 4 volumes. V. 2. Algae and fungi. Moscow, 320 pp. [In Russian]

Beyer TV (2007) Class Coccidea Leuckart, 1879 – Coccidia. In: Krylov MV, Frolov AO (Eds) Protista. Handbook on Zoology. Part 2. St. Petersburg, 149–248. [In Russian]

Blute M (2013) The evolution of anisogamy: more questions than answers. Biological Theory 7: 3–9. https://doi.org/10.1007/s13752-012-0060-4

Bogaert K, Beeckman T, De Clerck O (2017) Two-step cell polarization in algal zygotes. Nature Plants 3: 16221. https://doi.org/10.1038/nplants.2016.221

Bogaert KA, Delva S, De Clerck O (2020) Concise review of the genus *Dictyota* J.V. Lamouroux. Journal of Applied Phycology 32: 1521–1543. https://doi.org/10.1007/s10811-020-02121-4

Bonner JT (1998) The origins of multicellularity. Integrative Biology 1: 27–36. https://doi.org/10.1002/(SICI)1520-6602(1998)1:1<27::AID-INBI4>3.0.CO;2-6

Brawley SH (1992) Fertilization in natural populations of the dioecious brown alga *Fucus ceranoides* and the importance of the polyspermy block. Marine Biology 113: 145–157. https://doi.org/10.1007/BF00367648

Brawley SH, Johnson LE (1992) Gametogenesis, gametes and zygotes: an ecological perspective on sexual reproduction in the algae. British Phycological Journal 27(3): 233–252. https://doi.org/10.1080/00071619200650241

Brunet T, King N (2017) The origin of animal multicellularity and cell differentiation. Developmental Cell 43(2): 124–140. https://doi.org/10.1016/j.devcel.2017.09.016

Bulmer MG, Parker GA (2002) The evolution of anisogamy: a game-theoretic approach. Proceedings of the Royal Societe London B 269: 2381–2388. https://doi.org/10.1098/rspb.2002.2161

Chemin E (1937) Le développement des spores chez les Rhodophycées. Revue Générale de Botanique 49: 205–234, 300–327, 353–374, 424–448, 478–535.

Coates JC, Umm-E-Aiman, Charrier B (2015) Understanding "green" multicellularity: do seaweeds hold the key? Frontiers in Plant Science 5(737): 1–6. https://www.frontiersin.org/articles/10.3389/fpls.2014.00737

Cole K, Akintobi S (1963) The life cycle of *Prasiola meridionalis* Setchell and Gardner. Canadian Journal of Botany 41(5): 661–668. https://doi.org/10.1139/b63-053

Colizzi ES, Vroomans RMA, Merks RMH (2020) Evolution of multicellularity by collective integration of spatial information. eLife 9: e56349. https://doi.org/10.7554/eLife.56349

Davidovich NA (2019) Reproductive biology of diatoms. Doctor of Science Dissertation. Moscow, Russian Federation, Institute of Biology of the Southern Seas, 314 pp. [In Russian]

Davidovich NA, Kaczmarska I, Karpov SA, Davidovich OI, MacGillivary ML, Mather L (2012) Mechanism of male gamete motility in araphid pennate diatoms from the genus *Tabularia* (Bacillariophyta). Protistology 163(3): 480–494. https://doi.org/10.1016/j.protis.2011.09.002

Desnitsky AG (2018) Comparative analysis of embryonic inversion in the algae of the genus *Volvox* (Volovocales, Chlorophyta). Ontogenesis 49(3): 147–152. [In Russian]

Dobrovol'skaya TG (1974) Genus *Mycococcus*. In: Krasilnikov NA, Uranov AA (Eds) Plant life in 6 volumes. V. 1. Introduction. Bacteria and actinomycetes. Moscow, 296–300. [In Russian]

Dunn CW, Giribet G, Edgecombe GD, Hejnol A (2014) Animal phylogeny and its evolutionary implications. Annual Review of Ecology, Evolution and Systematics 45: 371–395. https://doi.org/10.1146/annurev-ecolsys-120213-091627

Edwards MS (2000) The role of alternate life-history stages of a marine macroalga: a seed bank analogue? Ecology 81: 2404–2415. http://dx.doi.org/10.1890/0012-9658(2000)081[2404:TROALH]2.0.CO;2

Egerod LE (1952) An analysis of the siphonous Chlorophycophyta with special reference to the Siphonocladales, Siphonales, and Dasycladales of Hawaii. University of California Publications in Botany 25: 327–367.

Enomoto S, Hirose H (1971) On the septum formation of *Microdictyon okamurai* Setchell. Bulletin of the Japanese Society of Phycology 19: 90–93.

Ereskovskiy AV (2005) Comparative embryology of sponges (Porifera). St. Petersburg, 2005. 304 pp. [In Russian]

Feis ME (2010) Reproduction in the genus *Fucus*. Haren, 18 pp.

Feist SW, Morris DJ, Alama-Bermejo G, Holzer AS (2015) Cellular processes in Myxozoans. In: Okamura B, Gruhl A, Bartholomew J (Eds) Myxozoan evolution, ecology and development. Cham, 139–154. https://doi.org/10.1007/978-3-319-14753-6_8

Felsenstein J (1974) The evolutionary advantage of recombination. Genetics 78(2): 737–756. https://doi.org/10.1093/genetics/78.2.737

Filin VR (1978) Phylum Lycopodiophyta. In: Grushvitskiy IV, Zhilin SG (Eds) Plant life in 6 volumes. V. 4. Mosses. Clubmosses. Horsetails. Ferns. Gymnosperms. Moscow, 99–121. [In Russian]

Foox J, Siddall ME (2015) The road to Cnidaria: history of the phylogeny of Myxozoa. The Journal of Parazitology 101(3): 269–274. https://doi.org/10.1645/14-671.1

Foster CA, Sarphief TG, Hawkinsf WE (1978) Fine structure of the peritrichous ectocommensal *Zoothamnium* sp. with emphasis on its mode of attachment to penaeid shrimp. Journal of Fish Diseases 1(4): 321–335. https://doi.org/10.1111/j.1365-2761.1978.tb00036.x

Franz V (1924) Geschichte der Organismen. Viena, 949 pp.

Friedmann I (1959) Structure, life-history, and sex determination of *Prasiola stipitata* Suhr. Annals of Botany 23(4): 571–594. https://doi.org/10.1093/oxfordjournals.aob.a083677

Fritsch FE (1929) The genus *Sphaeroplea*. Annals of Botany 43(1): 1–26. https://doi.org/10.1093/oxfordjournals.aob.a090153

Fursenko AV (1924) On the biology of *Zoothamnium arbuscula* Ehrenberg. Russian Archive of Protistology 3: 75–93. [In Russian]

Furuya H, Hochberg FG, Tsuneki K (2003) Reproductive traits of dicyemids. Marine Biology 142: 693–706. https://doi.org/10.1007/s00227-002-0991-6

Furuya H, Tsuneki K (2003) Biology of Dicyemid Mesozoans. Zoological Science 20: 519–532. https://doi.org/10.2108/zsj.20.519

Gavrilov-Zimin IA (2018) Ontogenesis, morphology and higher classification of archaeococcids (Homoptera: Coccinea: Orthezioidea). Zoosystematica Rossica. Supplementum 2, 260 pp. https://doi.org/10.31610/zsr/2018.supl.2.1

Gavrilov-Zimin IA (2022) Development of theoretical views on viviparity. Biological Bulletin Reviews 12: 570–595. https://doi.org/10.1134/S2079086422060032

Gavrilov-Zimin IA, Stekolshikov AV, Gautam DC (2015) General trends of chromosomal evolution in Aphidococca (Insecta, Homoptera, Aphidinea + Coccinea). Comparative Cytogenetics 9(3): 335–422. https://doi.org/10.3897/CompCytogen.v9i3.4930

Gladkova VN (1978) Family Aspleniaceae. In: Grushvitskiy IV, Zhilin SG (Eds) Plant life in 6 volumes. V. 4. Mosses. Clubmosses. Horsetails. Ferns. Gymnosperms. Moscow, 222–237. [In Russian]

Gollerbach MM (1977) Phylum Charophyta. In: Gollerbach MM (Ed.) Plant life in 6 volumes. V. 3. Algae. Lichens. Moscow, 338–350. [In Russian]

Golubkova NS (1977) External and internal structure of lichens. In: Gollerbach MM (Ed.) Plant life in 6 volumes. V. 3. Algae. Lichens. Moscow, 390–420. [In Russian]

Gonzalves EA, Mehra KR (1959) *Oocystaenium*, a new genus of the Chlorococcales. Hydrobiologia 13(1–2): 201–206. https://doi.org/10.1007/bf0004621810.1007/bf00046218

Grassé P-P (1953) Classe des Gregarinomorphes. In: Grassé P-P (Ed.) Traite de Zoologie 1(2): 550–690.

Grosberg RK, Stratchmann RR (2007) The evolution of multicellularity: a minor major transition? Annual Review of Ecology, Evolution, and Systematics 38: 621–654. https://doi.org/10.1146/annurev.ecolsys.36.102403.114735

Gruhl A, Okamura B (2015) Tissue characteristics and development in Myxozoa. In: Okamura B, Gruhl A, Bartholomew J (Eds) Myxozoan Evolution, Ecology and Development. Cham, 155–174. https://doi.org/10.1007/978-3-319-14753-6_9

Heesch S, Serrano-Serrano M, Barrera-Redondo J, Luthringer R, Peters AF, Destombe C, Mark Cock J, Valero M, Roze D, Salamin N, Coelho SM (2021) Evolution of life cycles and

reproductive traits: insights from the brown algae. Journal of Evolutionary Biology 34: 992–1009. https://doi.org/10.1111/jeb.13880

Herron MD, Rashidi A, Shelton DE, Driscoll WW (2013) Cellular differentiation and individuality in the "minor" multicellular taxa. Biological Reviews 88: 844–861. https://doi.org/10.1111/brv.12031

Hodgson AN (1997) Paraspermatogenesis in gastropod molluscs. Invertebrate Reproduction and Development 31: 31–38. https://doi.org/10.1080/07924259.1997.9672560

Hoyt WD (1927) The periodic fruiting of *Dictyota* and its relation to the environment. American Journal of Botany 14: 592–619. https://doi.org/10.1002/j.1537-2197.1927.tb04870.x

Ivanov AV (1968). Origin of multicellular animals. Phylogenetic essays. Leningrad, 287 pp. [In Russian]

Ivanova-Kazas OM (1975) Comparative embryology of invertebrates. Part 1. Protists and lower multicellular organisms. Novosibirsk, 372 pp. [In Russian]

Ivanova-Kazas OM (1977) Asexual reproduction of animals. Leningrad, 1975, 240 pp. [In Russian]

Ivanova-Kazas OM (1995) Evolutionary embryology of animals. St. Petersburg, 1995, 565 pp. [In Russian]

Jones WE, Moorjani Sh (1973) The attachment and and early development of tetraspores of some coralline red algae. MBAI Special Publication dedicated to Dr. N K Panikkar (1): 293–304. http://eprints.cmfri.org.in/id/eprint/2477

Kaczmarska I, Poulíčková A, Sato Sh, Edlund MB, Idei M, Watanabe T, Mann DG (2013) Proposals for a terminology for diatom sexual reproduction, auxospores and resting stages. Diatom Research 28(3): 263–294. https://doi.org/10.1080/0269249X.2013.791344

Kaplan-Levy RN, Hadas O, Summers ML, Rücker J, Sukenik A (2010) Akinetes: dormant cells of cyanobacteria. In: Lubzens E, Cerda J, Clark M (Eds) Dormancy and resistance in harsh environments. Topics in Current Genetics 21: 5–27. https://doi.org/10.1007/978-3-642-12422-8_2

Karpov SA (2011) Ichthyosporea. In: Karpov SA (Ed.) Protista. Handbook on Zoology. Part 3. St. Peterburg, 342–369. [In Russian]

Kawai H, Sasaki H, Maeda Y, Arai S (2001) Morphology, life history and molecular phylogeny of *Chorda rigida* sp. nov. (Laminariales, Phaeophyceae) from the Sea of Japan and the genetic diversity of *Chorda filum*. Journal of Phycology 37: 130–142. https://doi.org/10.1046/j.1529-8817.1999.014012130.x

Knoll AH (2011) The multiple origins of complex multicellularity. Annual Review of Earth and Planetary Sciences 39: 217–239. https://doi.org/10.1146/annurev.earth.031208.100209

Kobara T, Chihara M (1984) Spermatozoids of *Pseudodichotomosiphon constrictus* with special reference to the systematic position of the genus. Journal of Japanese Botany 59(1): 20–25.

Lamża Ł (2023) Diversity of 'simple' multicellular eukaryotes: 45 independent cases and six types of multicellularity. Biological Reviews 2023: 1–22. https://doi.org/10.1111/brv.13001

Leliaert F, De Clerck O, Verbruggen H, Boedeker C, Coppejans E (2007) Molecular phylogeny of the Siphonocladales (Chlorophyta: Cladophorophyceae). Molecular Phylogenetics and Evolution 44(3): 1237–1256. https://doi.org/10.1016/j.ympev.2007.04.016

Luthringer R, Cormier A, Ahmed S, Peters AF, Cock JM, Coelho SM (2014) Sexual dimorphism in the brown algae. Perspectives in Phycology 1(1): 11–25. https://doi.org/10.1127/2198-011X/2014/0002

Malakhov VV (1990) Mysterious groups of marine invertebrates. Moscow, 144 pp. [In Russian]

Malakhov VV, Bogomolova EV, Kuzmina TV, Temereva EV (2019) Evolution of Metazoan life cycles and the origin of pelagic larvae. Russian Journal of Developmental biology 50(6): 303–316. [In Russian] https://doi.org/10.1134/S1062360419060043

Matvienko AM (1977) Class Volvocophyceae. In: Gollerbach MM (Ed.) Plant life in 6 volumes. V. 3. Algae. Lichens. Moscow, 266–281. [In Russian]

McConnaughey BH (1951) The life cycle of the Dicyemid Mesozoa. University of California Publications in Zoology (55): 295–336.

McConnaughey B, McConnaughey E (1954) Strange life of Dicyemid Mesozoans. The Scientific Monthly 79(5): 277–284.

McDonald KL, Pickett-Heaps JD (1976) Ultrastructure and differentiation in *Cladophora glomerata*. I. Cell Division. American Journal of Botany 63: 592–601. https://doi.org/10.1002/j.1537-2197.1976.tb11847.x

Mikhailov KV, Konstantinova AV, Nikitin MA, Troshin PV, Rusin LYu, Lyubetsky VA, Panchin YV, Mylnikov AP, Moroz LL, Kumar S, Aleoshin VV (2009) The origin of Metazoa: a transition from temporal to spatial cell differentiation. BioEssays 31: 758–768. https://doi.org/10.1002/bies.200800214

Michetti KM, Martín LA, Leonardi PI (2013) Carpospore release and sporeling development in *Gracilaria gracilis* (Gracilariales, Rhodophyta) from the southwestern Atlantic coast (Chubut, Argentina). Journal of Applied Phycology 25: 1917–1924. https://doi.org/10.1007/s10811-013-0029-0

Nanda N (1993) Egg release and germling development of *Myagropsis myagroides* (Mertens ex Turner) Fensholt. Japanese Journal of Phycology (Sôrui) 41: 315–325.

Nayar BK, Kaur S (1971) Gametophytes of homosporous ferns. The Botanical Review 37(3): 295–396. https://doi.org/10.1007/BF02859157

Nehira K (1983) Spore germination, protonema development and sporeling development. In: Schuster RM (Ed.) New Manual of Bryology I. Nichinan, 343–385.

Niklas KJ, Newman SA (2013) The origins of multicellular organisms. Evolution and Development 15(1): 41–52. https://doi.org/10.1111/ede.12013

Noto T, Yazaki K, Endoh H (2003) Developmentally regulated extrachromosomal circular DNA formation in the mesozoan *Dicyema japonicum*. Chromosoma 111: 359–368. https://doi.org/10.1007/s00412-002-0216-2

Novozhilov YuK, Gudkov AV (2000) Class Eumycetozoea. In: Karpov SA (Ed.) Protista. Handbook on Zoology. Part 1. St. Petersburg, 417–443. [In Russian]

Okuda K, Sekida S, Hasebe A, Iwabuchi M, Kamiya M, Hishinuma T (2016) Segregative cell division and the cytoskeleton in two species of the genus *Struvea* (Cladophorales, Ulvophyceae, Chlorophyta). Phycological Research 64: 219–229. https://doi.org/10.1111/pre.12139

Ollerton J, Ren Z-X (2022) Did pollination exist before plants? Science 377: 471–472. https://doi.org/10.1126/science.add319

Parker GA, Baker RR, Smith VG (1972) The origin and evolution of gamete dimorphism and the male-female phenomenon. Journal of Theoretical Biology 36(3): 529–53. https://doi.org/10.1016/0022-5193(72)90007-0

Petrov YuE (1977) Phylum Phaeophyta. In: Gollerbach MM (Ed.) Plant life in 6 volumes. V. 3. Algae. Lichens. Moscow, 143–192. [In Russian]

Potemkin AD, Sofronova EV (2009) Liverworts and hornworts of Russia. V.1. St. Petersburg – Yakutsk, 368 pp. [In Russian]

Pugachev ON, Podlipaev SA (2007) Type Myxozoa Grassé, 1970. In: Krylov MV, Frolov AO (Eds) Protista. Handbook on Zoology. Part 2. St. Petersburg, 1045–1080. [In Russian]

Raikova EV (1994) Life cycle, cytology, and morphology of *Polypodium hydriforme*, a Coelenterate parasite of the eggs of Acipenseriform fishes. The Journal of Parasitology 80(1): 1–22. https://doi.org/10.2307/3283338

Reiswig HM (1970a) Population dynamics of three Jamaican Demospongiae. Bulletin of Marine Science 23(2): 191–226.

Reiswig HM (1970b) Porifera: sudden sperm release by tropical Demospongiae. Science 170(3957): 538–539. https://doi.org/10.1126/science.170.3957.538

Reiswig HM (1976) Natural gamete release and oviparity in Caribbean Demospongiae. In: Harrison FW, Cowden RR (Eds) Aspects of sponge biology. New York, 99–112. https://doi.org/10.1016/B978-0-12-327950-7.50013-0

Reproductive Systems (2000) In: Batygina (Ed.) Embryology of flowering plants. Terminology and concepts. V. 3. St. Petersburg, 639 pp. [In Russian]

Schulz HN, Brinkhoff T, Ferdelman TG, Mariné MH, Teske A, Jorgensen BB (1999) Dense populations of a giant sulfur bacterium in Namibian shelf sediments. Science 284(5413): 493–495. https://doi.org/10.1126/science.284.5413.493

Searles RB (1980) The strategy of the red algal life history. The American Naturalist 115(1): 113–120. https://doi.org/10.1086/283548

Seiler S, Heilig Y (2019) Septum formation and cytokinesis in Ascomycete fungi. In: Hoffmeister D, Gressler M (Eds) Biology of the Fungal Cell. The Mycota 8: 15–42. https://doi.org/10.1007/978-3-030-05448-9_2

Serrão EA, Pearson G, Kautsky L, Brawley SH (1996) Successful external fertilization in turbulent environments. Proceeding of the National Academy of Sciences of the USA 93: 5286–5290. https://doi.org/10.1073/pnas.93.11.5286

Siddall ME, Martin DS, Bridge D, Desser SS, Cone DK (1995) The demise of a phylum of protists: phylogeny of Myxozoa and other parasitic Cnidaria. The Journal of Parazitology 81(6): 961–967. https://doi.org/10.2307/3284049

Sizova TP (1976) Class of chytridiomycetes (Chytridiomycetes). In: Gorlenko MV (Ed.) Plant life in 6 volumes. V. 2. Fungi. Moscow, 23–34. [In Russian]

Simdyanov TG (2007) Class Gregarinea Dufour, 1828. In: Krylov MV, Frolov AO (Eds) Protista. Handbook on Zoology. Part 2. St. Petersburg, 20–149. [In Russian]

Slyusarev GS (2008) Type Orthonectida: structure, biology, position in the system of multicellular animals. Journal of General Biology 69(6): 403–427. [In Russian]

Smith GM (1947) On the reproduction of some Pacific Coast species of *Ulva*. American Journal of Botany 34(2): 80–87. https://doi.org/10.2307/2437232

Sparrow FK (1933) The Monoblepharidales. Annals of Botany 47(3): 517–542. https://doi.org/10.1093/oxfordjournals.aob.a090402

Štolc A (1899) Actinomyxidies, nouveau groupe de Mesozoaires parent des Myxosporidies. Bulletin International de l'Academie des Sciences de Boheme 22: 1–12.

Suga H, Ruiz-Trillo I (2013) Development of ichthyosporeans sheds light on the origin of metazoan multicellularity. Developmental Biology 377(1): 284–292. https://doi.org/10.1016/j.ydbio.2013.01.009

Swallow JG, Wilkinson GS (2002) The long and short of sperm polymorphisms in insects. Biological Reviews, Cambidge Philosophical Society 77(2): 153–82. https://doi.org/10.1017/s1464793101005851

Timonin AK (2007) Botany in 4 volumes. V. 3. Higher plants. Moscow, 352 pp. [In Russian]

Timonin AK, Filin VR (2009) Botany in 4 volumes. Vol. 4(1). Systematics of higher plants. Moscow, 320 pp. [In Russian]

Timonin AK, Sokolov DD, Shipunov AB (2009) Botany in 4 volumes. Vol. 4(2). Systematics of higher plants. Moscow, 352 pp. [In Russian]

Umen J, Coelho S (2019) Algal sex determination and the evolution of anisogamy. Annual Review of Microbiology 73(12): 267–291. https://doi.org/10.1146/annurev-micro-020518-120011

Umen JG (2014) Green algae and the origins of multicellularity in the plant kingdom. Cold Spring Harbor Perspectives in Biology 6(11): a016170. https://doi.org/10.1101/cshperspect.a016170

Vinogradova KL (1977a) Phylum Rhodophyta. In: Gollerbach MM (Ed.) Plant life in 6 volumes. V. 3. Algae. Lichens. Moscow, 192–250. [In Russian]

Vinogradova KL (1977b) Class Ulotrichophyceae. Class Siphonophyceae. In: Gollerbach MM (Ed.) Plant life in 6 volumes. V. 3. Algae. Lichens. Moscow, 281–305. [In Russian]

Wai MK (2018) Morphotaxonomy, culture studies and phytogeographical distribution of Amphiroa fragilissima (Linnaeus) Lamouroux (Corallinales, Rhodophyta) from Myanmar. Journal of Aquaculture & Marine Biology 7(3): 142–150. https://doi.org/10.15406/jamb.2018.07.00201

Westheide W, Rieger R (2004) Spezielle Zoologie. T. 1. Einzeller und Wirbellose Tiere. T. 2. Wirbel und Schädeltiere. Heidelberg – Berlin, [922 +] 712 pp.

Yakovlev MS (1981) Dictionary of basic terms. In: Yakovlev MS (Ed.) Comparative embryology of flowering plants. Winteraceae – Juglandaceae. Leningrad, 7–25. [In Russian]

Zakhvatkin AA (1949) Comparative embryology of lower invertebrates. Moscow, 395 pp. [In Russian]

Zakhvatkin AA (1956) Vergleichende Embryologie der niederen Wirbellosen. Berlin.

ORCID

Ilya A. Gavrilov-Zimin https://orcid.org/0000-0003-1993-5984

REVIEW ARTICLE

Chromosome study of the Hymenoptera (Insecta): from cytogenetics to cytogenomics

Vladimir E. Gokhman[1]

1 *Botanical Garden, Moscow State University, Moscow 119234, Russia*

Corresponding author: Vladimir E. Gokhman (vegokhman@hotmail.com)

Citation: Gokhman VE (2023) Chromosome study of the Hymenoptera (Insecta): from cytogenetics to cytogenomics. Comparative Cytogenetics 17: 239–250. https://doi.org/10.3897/compcytogen.17.112332

Abstract

A brief overview of the current stage of the chromosome study of the insect order Hymenoptera is given. It is demonstrated that, in addition to routine staining and other traditional techniques of chromosome research, karyotypes of an increasing number of hymenopterans are being studied using molecular methods, e.g., staining with base-specific fluorochromes and fluorescence *in situ* hybridization (FISH), including microdissection and chromosome painting. Due to the advent of whole genome sequencing and other molecular techniques, together with the "big data" approach to the chromosomal data, the current stage of the chromosome research on Hymenoptera represents a transition from Hymenoptera cytogenetics to cytogenomics.

Introduction

From the very introduction of the term "genome" (Winkler 1920), geneticists have been well aware that one of the most basic genomic features is how heritable matter of the nucleus is divided into separate cytological units, i.e., chromosomes. Moreover, the initial definition of this term, in fact, was indeed centered on the haploid chromosome set (Winkler 1920). Among organisms with sequenced genomes, insects play a crucial part due to their vast numbers and ecological significance (Li et al. 2019). Furthermore, they represent "an ideal group to examine the causes and consequences

of chromosomal evolution. Insects are diverse with over one million named species, and are highly variable in chromosome number and in many other traits, such as sex determination systems, population sizes, generation times, habitats, and natural history" (Alfieri et al. 2023). This is also undoubtedly true for the largest insect taxa, e.g., Hymenoptera, which is one of the most species-rich, taxonomically complicated and economically important orders of insects. The current number of described members of this group exceeds 150 thousand (Huber 2017), and the potential number of Hymenoptera may well exceed a million species, mostly due to a large number of still undescribed parasitoids (Forbes et al. 2018). Among these insects, karyotypic data are available for just about two thousand members, and for many of them little is known beyond the chromosome number (Gokhman 2023), not to mention a few hymenopteran superfamilies (e.g., Orussoidea, Megalyroidea and Stephanoidea), for which karyotypes are completely unknown. Nevertheless, certain taxa, e.g., some parasitoids, ants and wasps, are apparently better studied in this respect than the others. In addition, molecular data on this order, including results of the whole genome sequencing, are also rapidly accumulating now (see, e.g., Branstetter et al. 2018). This paper briefly overviews the present state of cytogenetic research on Hymenoptera and discusses its place in the context of the genomic study of this vast group.

Progress of the cytogenetic study of the order Hymenoptera

In a recently published review (Gokhman 2023), I have summarized the historical development of the karyotype research of the order Hymenoptera. According to this outline, three consecutive stages of this study took place in the 1890–1920s, 1930–1960s and 1970–1990s. Although chromosome research on this group was mostly done (and is still done today) involving traditional techniques, e.g., routine chromosome staining as well as C- and AgNOR-bandings, progressive accumulation of advanced methods did take place with time. This also applies to the current stage of karyotype research, which started in the 2000s (Gokhman 2023) with new techniques that involve both obtaining and analyzing primary karyotype data. Notably, a detailed description of the chromosome set of the honeybee, *Apis mellifera* Linnaeus, 1758 (Apidae), appeared in the paper containing the first report of the fully sequenced genome of a hymenopteran (The Honeybee Genome Sequencing Consortium 2006). Nowadays, the number of species with sequenced genomes in the current version of the Hymenoptera Genome Database (https://hymenoptera.elsiklab.missouri.edu) (Elsik et al. 2016) approaches 120 (Walsh et al. 2022 onwards), i.e., it is approximately six times larger than the number of these species at the time of the first publication on this database (Elsik et al. 2016). However, the real number of sequenced genomes is much higher (perhaps more than 300), since many studied hymenopterans are apparently still not included into the database (see, for example, Gokhman et al. 2017 for information on the sequenced genomes of the members of the parasitoid genus *Aphelinus* Dalman, 1820 from the chalcid family Aphelinidae).

Estimates of the genome sizes obtained using cytometry and/or whole genome sequencing (e.g., Moura et al. 2020, 2021; Cunha et al. 2021b) can also provide some insights on the genome evolution within the order Hymenoptera. Specifically, a simultaneous analysis of the karyotypes and genome sizes of *Aphelinus* species (Gokhman et al. 2017) demonstrated that chromosomal rearrangements in this group usually occurred independently of the changes in the genome size. In addition, comparative studies of these parameters conducted on different populations of two of the three known species of the ant genus *Mycetophylax* Emery, 1913 (Formicidae), *M. conformis* (Mayr, 1884) and *M. morschi* (Emery, 1888), showed that conspecific populations were significantly different in terms of the genome size and total karyotype length despite having the same chromosome number and karyotype morphology (Moura et al. 2020). The authors of this study suggest that these changes in the amount of genomic DNA could represent initial stages of karyotype evolution within certain ant species.

Molecular methods have played a crucial role in the recent progress of chromosome research on Hymenoptera. While initial attempts to employ base-specific fluorochromes and fluorescence *in situ* hybridization (FISH) for studying karyotypes of this order date back to the 1990s (Odierna et al. 1993; Lorite et al. 1997), use of these techniques has greatly increased since that time. Specifically, staining with 4',6-di-amidino-2-phenylindole (DAPI) proved that the DNA that constitutes hymenopteran chromosomes is predominantly AT-rich (as in most eukaryotes), with the exception of nucleolus organizing regions (NORs), which are usually GC-rich and are therefore stained with chromomycin A_3 (CMA_3) (see, e.g., Bolsheva et al. 2012). Nevertheless, most chromosomes of a few bee and parasitoid species carry GC-enriched segments (mostly terminal ones; see Gokhman 2023 for review), and at least some of them definitely do not represent NORs. Ultimately, FISH with probes derived from either full or partial large transcriptional units of ribosomal DNA, e.g., 45S or 18S rDNA, can reliably visualize NORs on hymenopteran chromosomes (Bolsheva et al. 2012; Gokhman et al. 2014; Piccoli et al. 2018; Micolino et al. 2019; Menezes et al. 2021; Pereira et al. 2021; Teixeira et al. 2021; Cunha et al. 2023). FISH also demonstrated that heterochromatin contains repetitive sequences which often differ between related genera and species of Hymenoptera (Lopes et al. 2014; Cunha et al. 2020). Moreover, in this order different microsatellites can be characteristic either of heterochromatin or euchromatin (dos Santos et al. 2018; Piccoli et al. 2018; Travenzoli et al. 2019; Elizeu et al. 2021; Cunha et al. 2023). In addition, FISH can detect the presence of certain transposable elements on the chromosomes of parasitoid and aculeate Hymenoptera (Lorite et al. 2012; Li et al. 2017). Finally, certain unique sequences were also localized on hymenopteran chromosomes using FISH (e.g., Matsumoto et al. 2002).

Nowadays, karyotype evolution of many insect taxa, including Hymenoptera, can be traced using a number of powerful cytogenetic methods, e.g., microdissection and chromosome painting, which is also based on the FISH technique. Using these methods, Fernandes et al. (2011) demonstrated that in the karyotype of the bee *Tetragonisca fiebrigi* (Schwarz, 1938) (Apidae), centromeres of different chromosome pairs are heterogeneous in terms of their DNA content. On the other hand, Martins et al. (2013)

explored B chromosomes of another bee species, *Partamona helleri* Friese, 1900 using the same approach. These authors showed that a probe derived from a certain type of B chromosomes hybridizes only with these elements. In addition, Rütten et al. (2004), who used both microdissection and whole chromosome painting (WCP), were able to identify every chromosome in the haploid karyotype of the parasitoid, *Nasonia vitripennis* (Walker, 1836) (Pteromalidae) containing five metacentrics of similar size (n = 5).

Supergenes, i.e., tightly linked sets of loci that are inherited together, control complex phenotypes and are usually characterized by reduced meiotic recombination due to certain features of the genome, now play an increasingly important role in studying many aspects of ecology and genetics of various organisms (see, e.g., Berdan et al. 2022). Since inversions apparently represent the most frequent case of rearrangements responsible for restricting recombination between homologous chromosomes, it is not surprising that the first detected case of the supergene in the order Hymenoptera, namely, in the ant *Solenopsis invicta* Buren, 1972, was explored, among other techniques, using cytogenetic analysis (Wang et al. 2013). In this species, a particular inversion was found to be responsible for the details of social organization of the colony, and similar rearrangements were later discovered in other members of the same family Formicidae (Brelsford et al. 2020; Lagunas-Robles et al. 2021; Kay et al. 2022; Chapuisat 2023) as well as in *Apis mellifera* (Wallberg et al. 2017). We have recently found another putative supergene in two cryptic species of parasitoids of the *Lariophagus distinguendus* (Förster, 1841) complex (Pteromalidae). These species have different chromosome numbers, n = 5 and 6, and a phylogenetic analysis based on molecular data indicates that chromosomal fusion occurred in this complex, with a certain acrocentric and a particular metacentric in the species with n = 6 corresponding to the shorter and longer arms of the largest metacentric chromosome in the species with n = 5 (König et al. 2019; Gokhman et al. 2019). This chromosomal fusion, together with a possible inversion in the longer arm of the above-mentioned metacentric in the species having n = 5, apparently prevents effective recombination between alternative variants of the supergene in these two morphologically indistinguishable species with strong biological differences (König et al. 2019). I therefore suggest that similar supergenes could also be responsible for the process of divergence of other groups of cryptic species of the order Hymenoptera.

A fascinating history of studying telomeric regions in the order Hymenoptera can serve as another example of applying a cytogenetic approach to the investigation of the genomic architecture of these insects. Specifically, these regions in most organisms have particular telomeric motifs; for example, the $(TTAGG)_n$ repeat is characteristic of many insects (see, e.g., Kuznetsova et al. 2020). Although initial cytogenetic analysis apparently confirmed presence of this motif in Hymenoptera (Frydrychová et al. 2004; Vítková et al. 2005), only several dozen ant species as well as *Apis mellifera* were studied at that time (Sahara et al. 1999; Lorite et al. 2002). However, the *Nasonia* Genome Working Group (2010) did not find this repeat in the genome of *Nasonia vitripennis*. Moreover, we also failed to reveal this motif on chromosomes of other studied parasitoids of the superfamilies Ichneumonoidea, Cynipoidea and Chalcidoidea (Gokhman et al. 2014).

In addition, Menezes et al. (2013, 2017) showed that the $(TTAGG)_n$ repeat is absent from the genomes of all studied aculeate Hymenoptera except for Apidae and Formicidae. Nevertheless, telomeric motifs in the suborder Symphyta remained unknown until the last five years, when we demonstrated presence of the canonical $(TTAGG)_n$ telomeric repeat in two members of the sawfly family Tenthredinidae, thus suggesting the ancestral nature of this motif in the order (Gokhman and Kuznetsova 2018). Two years later, Dalla Benetta et al. (2020) finally identified the $(TTATTGGG)_n$ repeat as the telomeric motif in *N. vitripennis* using both bioinformatic and cytogenetic approaches. Subsequent bioinformatic research has confirmed the two latter motifs, sometimes with a few variations, as characteristic features of the Symphyta and Chalcidoidea, respectively (Zhou et al. 2022). Furthermore, two recent studies (Fajkus et al. 2023; Lukhtanov and Pazhenkova 2023) have discovered an unprecedented diversity of telomeric repeats in the order Hymenoptera. Fajkus et al. (2023) demonstrated that short telomerase RNAs (TRs) in these insects are of the small nuclear RNA (snRNA) type, and are likely transcribed with RNA polymerase III. Surprisingly, this feature is characteristic of green plants and ciliates, apart from animals. Since TRs are used as templates for synthesizing telomeric motifs, the dramatic change in their structure and biogenesis have apparently led to an enormous increase in diversity of these repeats in the Hymenoptera. For example, TTAGGTCTGGG, TTGCGTCTGGG and TTAGGTTGGGG telomeric motifs were found in many aculeates, in the superfamily Vespoidea and in the genus *Bombus* Latreille, 1802 (Apidae) respectively (see also Lukhtanov and Pazhenkova 2023). On the other hand, Fajkus et al. (2023) did find the canonical insect repeat, $(TTAGG)_n$, in a few parasitoids, including the only studied member of the family Mymaridae, thus confirming its basal position among other Chalcidoidea. Analogously, Lukhtanov and Pazhenkova (2023) detected the same motif in a number of bees (Anthophila) and in a few other aculeates, and showed that telomeric sequences in most insects represent arrays of short repeats interspersed by non-LTR retrotransposons, with those of the SART family prevailing in the Hymenoptera. Lukhtanov and Pazhenkova (2023) also hypothesize that insect telomeres are usually maintained by both telomerase-dependent and independent mechanisms, and shifts in the balance between these processes can lead to an increased diversity in the telomere structure as well.

The information summarized above also indicates that use of molecular data and availability of computational analytical tools provide new opportunities for analyzing karyotype information. This process has twofold significance. First, an increased computer power allows handling enormous amounts of chromosomal data (the so-called "big data" approach). Second, it leads to new, much more reliable phylogenetic reconstructions resolving many aspects of karyotype evolution. In the framework of the "big data" approach, for example, the chromosome number can be considered as a proxy for the level of recombination, and therefore its variation both among and within specific clades can point to different features of the evolutionary chromosome change. Indeed, a particular study of that kind was implemented about a decade ago on more than 1,500 members of the order (Ross et al. 2015). By calculating variance in the chromosome number in solitary vs. eusocial Hymenoptera, we demonstrated that this

variance is about three times higher in the latter group, thus showing some specific features of the karyotype/genome evolution in the eusocial members of the order. Analogously, databases covering certain groups and/or particular chromosomal characters systematize our knowledge of the chromosome/genome features of the Hymenoptera and therefore help outlining pathways of the corresponding traits. These databases include the Bee Chromosome Database (https://bees.ufop.br) and the Ant Chromosome Database (https://ants.ufop.br) (Cardoso et al. 2018; Cunha et al. 2021a), as well as the databases on the number and position of ribosomal DNA (rDNA) clusters in animals (https://www.animalrdnadatabase.com) (Sochorová et al. 2021) and on the structure of telomere sequences, TeloBase (http://cfb.ceitec.muni.cz/telobase) (Lyčka et al. 2023). In addition, certain published reviews of chromosomal data of other large groups of Hymenoptera, e.g., Symphyta and Parasitica, are also available, although not in the form of online databases (Westendorff 2006; Gokhman 2009), but these publications are nevertheless substantially important.

The above-mentioned parallel accumulation of karyotypic and genomic data leads not only to general progress of cytogenetic studies of the Hymenoptera, but also to a qualitative transition toward a new level of cytogenetic knowledge, from studying separate DNA sequences to a network of interacting genes, and, ideally, to integral characteristics of whole genomes. On the other hand, this data accumulation allows independent checking of the results obtained by molecular and chromosomal techniques. For example, whole genome sequencing implies chromosome-level assemblies of different genomes, and counting chromosome numbers provides direct estimates of the numbers of linkage groups, which, in turn, can be compared to those of the obtained scaffolds.

Interestingly, all these features also characterize the newly introduced term "cytogenomics". Although this term apparently lacks a universally accepted clear-cut definition, most experts agree that it implies a modern synthesis of cytogenetic and molecular approaches aimed at comprehensive research of the structure and functions of eukaryotic chromosomes with a special emphasis on DNA that constitutes these chromosomes (see, e.g., Liehr 2021). In addition, cytogenomics, which is sometimes also called "chromosomics" (Deakin et al. 2019), rather focuses on features of the entire karyotypes and genomes, as opposed to those of particular chromosomal regions and certain DNA sequences. However, since a considerable amount of information on Hymenoptera chromosomes is still obtained using classical cytogenetic techniques (see, e.g., Gokhman 2009), I argue that we are currently experiencing a transition from cytogenetic to cytogenomic research on Hymenoptera.

Conclusions and future prospects

The present overview of cytogenetic research of the order Hymenoptera shows that, although many works still examine routinely stained chromosomes (see, e.g., König et al. 2019; Afonso Neto et al. 2022) and/or distribution and content of particular sequences and chromosomal segments, certain integral characteristics of the genomes

are also studied. All this information suggests that we currently are in the process of transitioning from cytogenetics to cytogenomics of the Hymenoptera. As far as further prospects in cytogenomic research of Hymenoptera are concerned, I believe that they imply a combination of cytogenetic and molecular approaches, which will be focused on large chromosomal regions and whole chromosomes. Specifically, microdissection and chromosome painting could become powerful instruments of studying syntenies among hymenopteran karyotypes, especially in the case of complex rearrangements between closely related species. For example, chromosome sets of the two morphologically similar parasitoids of the genus *Anisopteromalus*, *A. quinarius* Gokhman et Baur, 2014 and *A. calandrae* (Howard, 1881), with n = 5 and 7 respectively, differ to an extent that prevents any feasible reconstruction of chromosomal rearrangements that led to the origin of those karyotypes (Gokhman et al. 1998). Under these circumstances, sequencing of microdissection products as well as use of other combinations of the cytogenetic and molecular approaches seem very promising. Finally, I am aware of only one use of specific antibodies to visualize particular components of hymenopteran chromosomes. Specifically, fluorochrome-conjugated antibodies showed the distribution of 5-methylcytosine along chromosomes of a certain parasitoid (Bolsheva et al. 2012), and I believe that similar studies could reveal many details of fine chromosome structure in this order.

Acknowledgements

References

Afonso Neto PC, Micolino R, Cardoso DC, Cristiano MP (2022) Phylogenetic reconstruction of the ancestral chromosome number of the genera *Anochetus* Mayr, 1861 and *Odontomachus* Latreille, 1804 (Hymenoptera: Formicidae: Ponerinae). Frontiers in Ecology and Evolution 10: 829989. https://doi.org/10.3389/fevo.2022.829989

Alfieri JM, Jonika MM, Dulin JN, Blackmon H (2023) Tempo and mode of genome structure evolution in insects. Genes 14: 336. https://doi.org/10.3390/genes14020336

Berdan EL, Flatt T, Kozak GM, Lotterhos KE, Wielstra B (2022) Genomic architecture of supergenes: connecting form and function. Philosophical Transactions of the Royal Society B 377: 20210192. https://doi.org/10.1098/rstb.2021.0192

Bolsheva NL, Gokhman VE, Muravenko OV, Gumovsky AV, Zelenin AV (2012) Comparative cytogenetic study on two species of the genus *Entedon* Dalman, 1820 (Hymenoptera: Eulophidae) using DNA-binding fluorochromes and molecular and immunofluorescent

markers. Comparative Cytogenetics 6(1): 79–92. https://doi.org/10.3897/compcytogen.v6i1.2349

Branstetter MG, Childers AK, Cox-Foster D, Hopper KR, Kapheim KM, Toth AL, Worley KC (2018) Genomes of the Hymenoptera. Current Opinion in Insect Science 25: 65–75. https://doi.org/10.1016/j.cois.2017.11.008

Brelsford A, Purcell J, Avril A, Van PT, Zhang J, Brütsch T, Sundström L, Helanter H, Chapuisat M (2020) An ancient and eroded social supergene is widespread across *Formica* ants. Current Biology 30: 304–311. https://doi.org/10.1016/j.cub.2019.11.032

Cardoso DC, Santos HG, Cristiano MP (2018) The Ant Chromosome database – ACdb: an online resource for ant (Hymenoptera: Formicidae) chromosome researchers. Myrmecological News 27: 87–91.

Chapuisat M (2023) Supergenes as drivers of ant evolution. Myrmecological News 33: 1–18. https://doi.org/10.25849/myrmecol.news_033:001

Cunha MS, Campos LAO, Lopes DM (2020) Insights into the heterochromatin evolution in the genus *Melipona* (Apidae: Meliponini). Insectes Sociaux 67: 391–398. https://doi.org/10.1007/s00040-020-00773-6

Cunha MS, Cardoso DC, Cristiano MP, Campos LAO, Lopes DM (2021a) The Bee Chromosome database (Hymenoptera: Apidae). Apidologie 52(2): 493–502. https://doi.org/10.1007/s13592-020-00838-2

Cunha MS, Soares FAF, Clarindo WR, Campos LAO, Lopes DM (2021b) Robertsonian rearrangements in Neotropical Meliponini karyotype evolution (Hymenoptera: Apidae: Meliponini). Insect Molecular Biology 30(4): 379–389. https://doi.org/10.1111/imb.12702

Cunha MS, Garcia MVB, Campos LAO, Lopes DM (2023) Cytotaxonomy and karyotype evolution in Neotropical Meliponini (Hymenoptera: Apidae) inferred by chromosomal mapping of 18S rDNA and five microsatellites. Journal of Apicultural Research. https://doi.org/10.1080/00218839.2023.2179228

Dalla Benetta E, Antoshechkin I, Yang T, Nguyen HQM, Ferree PM, Akbari OS (2020) Genome elimination mediated by gene expression from a selfish chromosome. Science Advances 6: eaaz9808. https://doi.org/10.1126/sciadv.aaz9808

Deakin JE, Potter S, O'Neill R, Ruiz-Herrera A, Cioffi MB, Eldridge MDB, Fukui K, Marshall Graves JA, Griffin D, Grutzner F, Kratochvíl L, Miura I, Rovatsos M, Srikulnath K, Wapstra E, Ezaz T (2019) Chromosomics: Bridging the gap between genomes and chromosomes. Genes 10: 627. https://doi.org/10.3390/genes10080627

dos Santos JM, Diniz D, Rodrigues TAS, Cioffi MB, Waldschmidt AM (2018) Heterochromatin distribution and chromosomal mapping of microsatellite repeats in the genome of *Frieseomelitta* stingless bees (Hymenoptera: Apidae: Meliponini). Florida Entomologist 101: 33–39. https://doi.org/10.1653/024.101.0107

Elizeu AM, Travenzoli NM, Ferreira RP, Lopes DM, Tavares MG (2021) Comparative study on the physical mapping of ribosomal genes and repetitive sequences in *Friesella schrottkyi* (Friese 1900) (Hymenoptera: Apidae, Meliponini). Zoologischer Anzeiger 292: 225–230. https://doi.org/10.1016/j.jcz.2021.04.006

Elsik CG, Tayal A, Diesh CM, Unni DR, Emery ML, Nguyen HN, Hagen DE (2016) Hymenoptera Genome Database: integrating genome annotations in HymenopteraMine. Nucleic Acids Research 44(D1): D793–D800. https://doi.org/10.1093/nar/gkv1208

Fajkus P, Adámik M, Nelson ADL, Kilar AM, Franek M, Bubeník M, Čapková Frydrychová R, Votavová A, Sýkorová E, Fajkus J, Peška V (2023) Telomerase RNA in Hymenoptera (Insecta) switched to plant/ciliate-like biogenesis. Nucleic Acids Research 51: 420–433. https://doi.org/10.1093/nar/gkac1202

Fernandes A, Scudeler PES, Diniz D, Foresti F, Campos LAO, Lopes DM (2011) Microdissection: a tool for bee chromosome studies. Apidologie 42: 743–748. https://doi.org/10.1007/s13592-011-0082-0

Forbes AA, Bagley RK, Beer MA, Hippee AC, Widmayer HA (2018) Quantifying the unquantifiable: Why Hymenoptera, not Coleoptera, is the most speciose animal order. BMC Ecology 18: 1–21. https://doi.org/10.1186/s12898-018-0176-x

Frydrychová R, Grossmann P, Trubač P, Vítková M, Marec F (2004) Phylogenetic distribution of TTAGG telomeric repeats in insects. Genome 47: 163–178. https://doi.org/10.1139/g03-100

Gokhman VE (2009) Karyotypes of Parasitic Hymenoptera. Springer Science + Business Media B.V., Dordrecht, 183 pp. https://doi.org/10.1007/978-1-4020-9807-9

Gokhman VE (2023) Chromosome study of the Hymenoptera: History, current state, perspectives. Biology Bulletin Reviews 13(3): 247–257. https://doi.org/10.1134/S2079086423030040

Gokhman VE, Anokhin BA, Kuznetsova VG (2014) Distribution of 18S rDNA sites and absence of the canonical TTAGG insect telomeric repeat in parasitoid Hymenoptera. Genetica 142(4): 317–322. https://doi.org/10.1007/s10709-014-9776-3

Gokhman VE, Cioffi MB, König C, Pollmann M, Gantert C, Krogmann L, Steidle JLM, Kosyakova N, Liehr T, Al-Rikabi A (2019) Microdissection and whole chromosome painting confirm karyotype transformation in cryptic species of the *Lariophagus distinguendus* (Förster, 1841) complex (Hymenoptera: Pteromalidae). PLoS ONE 14(11): e0225257. https://doi.org/10.1371/journal.pone.0225257

Gokhman VE, Kuhn KL, Woolley JB, Hopper KR (2017) Variation in genome size and karyotype among closely related aphid parasitoids (Hymenoptera, Aphelinidae). Comparative Cytogenetics 11(1): 97–117. https://doi.org/10.3897/CompCytogen.v11i1.10872

Gokhman VE, Kuznetsova VG (2018) Presence of the canonical TTAGG insect telomeric repeat in the Tenthredinidae (Symphyta) suggests its ancestral nature in the order Hymenoptera. Genetica 146(3): 341–344. https://doi.org/10.1007/s10709-018-0019-x

Gokhman VE, Timokhov AV, Fedina TY (1998) First evidence for sibling species in *Anisopteromalus calandrae* (Hymenoptera: Pteromalidae). Russian Entomological Journal 7(3–4): 157–162.

Huber JT (2017) Biodiversity of Hymenoptera. In: Foottit RG, Adler PH (Eds) Insect Biodiversity: Science and Society (2nd edn). Wiley Blackwell, Oxford, 419–461. https://doi.org/10.1002/9781118945568.ch12

Kay T, Helleu Q, Keller L (2022) Iterative evolution of supergene-based social polymorphism in ants. Philosophical Transactions of the Royal Society B 377: 20210196. https://doi.org/10.1098/rstb.2021.0196

König C, Paschke S, Pollmann M, Reinisch R, Gantert C, Weber J, Krogmann L, Steidle JLM, Gokhman VE (2019) Molecular and cytogenetic differentiation within the *Lariophagus distinguendus* (Förster, 1841) species complex (Hymenoptera, Pteromalidae). Comparative Cytogenetics 13(2): 133–145. https://doi.org/10.3897/CompCytogen.v13i2.34492

Kuznetsova V, Grozeva S, Gokhman V (2020) Telomere structure in insects: A review. Journal of Zoological Systematics and Evolutionary Research 58(1): 127–158. https://doi.org/10.1111/jzs.12332

Lagunas-Robles G, Purcell J, Brelsford A (2021) Linked supergenes underlie split sex ratio and social organization in an ant. Proceedings of the National Academy of Sciences of the USA 118(46): e2101427118. https://doi.org/10.1073/pnas.2101427118

Li F, Zhao X, Li M, He K, Huang C, Zhou Y, Li Z, Walters JR (2019) Insect genomes: progress and challenges. Insect Molecular Biology 28(6): 739–758. https://doi.org/10.1111/imb.12599

Li Y, Jing XA, Aldrich JC, Clifford C, Chen J, Akbari OA, Ferree PM (2017) Unique sequence organization and small RNA expression of a "selfish" B chromosome. Chromosoma 126: 753–768. https://doi.org/10.1007/s00412-017-0641-x

Liehr T (2021) A definition for cytogenomics – Which also may be called chromosomics. In: Liehr T (Ed.) Cytogenomics. Academic Press, London etc., 1–7. https://doi.org/10.1016/B978-0-12-823579-9.00001-1

Lopes DM, Fernandes A, Diniz D, Scudeler PES, Foresti F, Campos LAO (2014) Similarity of heterochromatic regions in the stingless bees (Hymenoptera: Meliponini) revealed by chromosome painting. Caryologia 67: 222–226. https://doi.org/10.1080/0144235X.2014.974349

Lorite P, Aránega AE, Luque F, Palomeque T (1997) Analysis of the nucleolar organizing regions in the ant *Tapinoma nigerrimum* (Hymenoptera, Formicidae). Heredity 78: 578–582. https://doi.org/10.1038/hdy.1997.96

Lorite P, Carrillo JA, Palomeque T (2002) Conservation of $(TTAGG)_n$ telomeric sequences among ants (Hymenoptera, Formicidae). The Journal of Heredity 93(4): 282–285. https://doi.org/10.1093/jhered/93.4.282

Lorite P, Maside X, Sanllorente O, Torres MI, Periquet G, Palomeque T (2012) The ant genomes have been invaded by several types of *mariner* transposable elements. Naturwissenschaften 99: 1007–1020. https://doi.org/10.1007/s00114-012-0982-5

Lukhtanov VA, Pazhenkova EA (2023) Diversity and evolution of telomeric motifs and telomere DNA organization in insects. Biological Journal of the Linnean Society: blad068. https://doi.org/10.1093/biolinnean/blad068

Lyčka M, Bubeník M, Závodník M, Peska V, Fajkus P, Demko M, Fajkus J, Fojtová M (2023) TeloBase: a community-curated database of telomere sequences across the tree of life. Nucleic Acids Research: gkad672. https://doi.org/10.1093/nar/gkad672

Martins CCC, Diniz D, Sobrinho-Scudeler PE, Foresti F, Campos LAO, Costa MA (2013) Investigation of *Partamona helleri* (Apidae, Meliponini) B chromosome origin. An approach by microdissection and whole chromosome painting. Apidologie 44: 75–81. https://doi.org/10.1007/s13592-012-0157-6

Matsumoto K, Yamamoto DS, Sumitani M, Lee JM, Hatakeyama M, Oishi K (2002) Detection of a single copy on a mitotic metaphase chromosome by fluorescence in situ hybridization (FISH) in the sawfly, *Athalia rosae* (Hymenoptera). Archive of Insect Biochemistry and Physiology 49: 34–40. https://doi.org/10.1002/arch.10005

Menezes RST, Silva TM, Carvalho AF, Andrade-Souza V, Silva JG, Costa MA (2013) Numerical and structural chromosome variation in the swarm-founding wasp *Metapolybia decorata* Gribodo 1896 (Hymenoptera, Vespidae). Genetica 141: 273–280. https://doi.org/10.1007/s10709-013-9726-5

Menezes RST, Bardella VB, Cabral-de-Mello DC, Lucena DAA, Almeida EAB (2017) Are the TTAGG and TTAGGG telomeric repeats phylogenetically conserved in aculeate Hymenoptera? The Science of Nature 104: 85. https://doi.org/10.1007/s00114-017-1507-z

Menezes RST, Cabral-de-Mello DC, Milani D, Bardella VB, Almeida EAB (2021) The relevance of chromosome fissions for major ribosomal DNA dispersion in hymenopteran insects. Journal of Evolutionary Biology 34: 1466–1476. https://doi.org/10.1111/jeb.13909

Micolino R, Cristiano MP, Travenzoli NP, Lopes DM, Cardoso DC (2019) Chromosomal dynamics in space and time: Evolutionary history of *Mycetophylax* ants across past climatic changes in the Brazilian Atlantic coast. Scientific Reports 9: 18800. https://doi.org/10.1038/s41598-019-55135-5

Moura MN, Cardoso DC, Baldez BCL, Cristiano MP (2020) Intraspecific variation in the karyotype length and genome size of fungus-farming ants (genus *Mycetophylax*), with remarks on procedures for the estimation of genome size in the Formicidae by flow cytometry. PLoS ONE 15(8): e0237157. https://doi.org/10.1371/journal.pone.0237157

Moura MN, Cardoso DC, Cristiano MP (2021) The tight genome size of ants: diversity and evolution under ancestral state reconstruction and base composition. Zoological Journal of the Linnean Society 193: 124–144. https://doi.org/10.1093/zoolinnean/zlaa135

Odierna G, Baldanza F, Aprea G, Olmo E (1993) Occurrence of G-banding in metaphase chromosomes of *Encarsia berlesei* (Hymenoptera: Aphelinidae). Genome 36: 662–667. https://doi.org/10.1139/g93-088

Pereira JA, Travenzoli NM, de Oliveira MP, Werneck HA, Salomão TMF, Lopes DM (2021) Molecular cytogenetics in the study of repetitive sequences helping to understand the evolution of heterochromatin in *Melipona* (Hymenoptera, Meliponini). Genetica 149: 55–62. https://doi.org/10.1007/s10709-020-00111-5

Piccoli MCA, Bardella VB, Cabral-de-Mello DC (2018) Repetitive DNAs in *Melipona scutellaris* (Hymenoptera: Apidae: Meliponidae): Chromosomal distribution and test of multiple heterochromatin amplification in the genus. Apidologie 49: 497–504. https://doi.org/10.1007/s13592-018-0577-z

Ross L, Blackmon H, Lorite P, Gokhman VE, Hardy NB (2015) Recombination, chromosome number and eusociality in the Hymenoptera. Journal of Evolutionary Biology 28(1): 105–116. https://doi.org/10.1111/jeb.12543

Rütten KB, Pietsch C, Olek K, Neusser M, Beukeboom LW, Gadau J (2004) Chromosomal anchoring of linkage groups and identification of wing size QTL using markers and FISH probes derived from microdissected chromosomes in *Nasonia* (Pteromalidae: Hymenoptera). Cytogenetic and Genome Research 105: 126–133. https://doi.org/10.1159/000078019

Sahara K, Marec F, Traut W (1999) TTAGG telomeric repeats in chromosomes of some insects and other arthropods. Chromosome Research 7: 449–460. https://doi.org/10.1023/A:1009297729547

Sochorová J, Gálvez F, Matyášek R, Garcia S, Kovařík A (2021) Analyses of the updated "Animal rDNA Loci Database" with an emphasis on its new features. International Journal of Molecular Sciences 22(21): 11403. https://doi.org/10.3390/ijms222111403

Teixeira GA, de Aguiar HJAC, Petitclerc F, Orivel J, Lopes DM, Barros LAC (2021) Evolutionary insights into the genomic organization of major ribosomal DNA in ant chromosomes. Insect Molecular Biology 30: 340–354. https://doi.org/10.1111/imb.12699

The Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. Nature 443: 931–949. https://doi.org/10.1038/nature05260

The *Nasonia* Genome Working Group (2010) Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. Science 327: 343–348. https://doi.org/10.1126/science.1178028

Travenzoli NM, Lima BA, Cardoso DC, Dergam JA, Fernandes-Salomão TM, Lopes DM (2019) Cytogenetic analysis and chromosomal mapping of repetitive DNA in *Melipona* species (Hymenoptera, Meliponini). Cytogenetic and Genome Research 158: 213–224. https://doi.org/10.1159/000501754

Vítková M, Král J, Traut W, Zrzavý J, Marec F (2005) The evolutionary origin of insect telomeric repeats, $(TTAGG)_n$. Chromosome Research 13: 145–156. https://doi.org/10.1007/s10577-005-7721-0

Wallberg A, Schöning C, Webster MT, Hasselmann M (2017) Two extended haplotype blocks are associated with adaptation to high altitude habitats in East African honey bees. PLoS Genetics 13(5): e1006792. https://doi.org/10.1371/journal.pgen.1006792

Walsh AT, Triant DA, Le Tourneau JJ, Shamimuzzaman M, Elsik CG (2022) Hymenoptera Genome Database: new genomes and annotation datasets for improved GO enrichment and orthologue analyses. Nucleic Acids Research 50(D1): D1032–D1039. https://doi.org/10.1093/nar/gkab1018

Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang Y-C, Shoemaker D, Keller L (2013) A Y-like social chromosome causes alternative colony organization in fire ants. Nature 493: 664–668. https://doi.org/10.1038/nature11832

Westendorff M (2006) Chromosomes of sawflies (Hymenoptera: Symphyta) – a survey including new data. In: Blank SM, Schmidt S, Taeger A (Eds) Recent Sawfly Research: Synthesis and Prospects. Goecke & Evers, Keltern, 39–60.

Winkler H (1920) Verbreitung und Ursache der Parthenogenesis im Pflanzen- und Tierreiche. Gustav Fischer, Jena, 232 pp. https://doi.org/10.5962/bhl.title.1460

Zhou Y, Wang Y, Xiong X, Appel AG, Zhang C, Wang X (2022) Profiles of telomeric repeats in Insecta reveal diverse forms of telomeric motifs in Hymenopterans. Life Science Alliance 5(7): e202101163. https://doi.org/10.26508/lsa.202101163

ORCID

Vladimir E. Gokhman https://orcid.org/0000-0001-9909-7559

Divergent karyotypes in five genera of the African endemic fish family Distichodontidae (Cithariniformes, Osteichthyes)

Sergey A. Simanovsky[1], Dmitry A. Medvedev[1], Fekadu Tefera[2], Alexander S. Golubtsov[1]

1 *Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninsky Prospect 33, Moscow, 119071 Russia* **2** *National Fishery and Aquatic Life Research Center, Ethiopian Institute of Agricultural Research, Sebeta, P.O. Box 64, Ethiopia*

Corresponding author: Sergey A. Simanovsky (sergey.a.simanovsky@gmail.com)

Citation: Simanovsky SA, Medvedev DA, Tefera F, Golubtsov AS (2023) Divergent karyotypes in five genera of the African endemic fish family Distichodontidae (Cithariniformes, Osteichthyes). Comparative Cytogenetics 17: 251–262. https://doi.org/10.3897/compcytogen.17.107744

Abstract

The African family Distichodontidae comprises 109 species in 16 genera. Up-to-date cytogenetic information was available for the only distichodontid species *Distichodus affinis* Günther, 1873. Here we report chromosome number and morphology in: *Distichodus engycephalus* Günther, 1864 (2n = 52, FN = 104), *Ichthyborus besse* (Joannis, 1835) (2n = 46, FN = 92), *Nannocharax niloticus* (Joannis, 1835) (2n = 54, FN = 106) and three taxa, *Nannaethiops bleheri* Géry et Zarske, 2003, *Nannaethiops* sp., and *Neolebias unifasciatus* Steindachner**,** 1894, that exhibit the same karyotypes (2n = 50, FN = 98). To confirm the *Nannaethiops* Günther, 1872 and *Neolebias* Steindachner, 1894 species identification, mt-DNA sequences of the two markers (*COI* and *16S rRNA*) were obtained from karyotyped specimens and compared with the relevant sequences accessible from GenBank. The great prevalence of biarmed chromosomes (the karyotypes of most species contain exclusively biarmed chromosomes) is a distinctive characteristic of Distichodontidae and Cithariniformes as a whole.

Keywords

Africa, chromosomes, *Distichodus*, *Ichthyborus*, karyotype evolution, *Nannaethiops*, *Nannocharax*, *Neolebias*

Introduction

Until recently the two Afrotropical families, Citharinidae and Distichodontidae, were considered as belonging to characins, the order Characiformes, classified into two sub-orders: Citharinoidei with 117 species in two Afrotropical families and Characoidei with more than 2000 species in two Afrotropical and 20 Neotropical families (Nelson et al. 2016; Froese and Pauly 2023). Recently, however, sister group relationships between Characoidei and catfishes, the order Siluriformes, has been inferred from the molecular data (Melo et al. 2022). Therefore, Cithariniformes along with Characiformes (containing former Characoidei only) and Siluriformes should be recognized as distinct orders (Dornburg and Near 2021).

While Citharinidae include eight species in three genera, Distichodontidae are more species rich including 109 species in 16 genera (Eschmeyer et al. 2023, Froese and Pauly 2023). The molecular phylogeny of Citharinoidei is well established: there is the distinct family Citharinidae and six clades within the family Distichodontidae (Arroyave et al. 2013; Lavoué et al. 2017). Two representatives of the former family – *Citharinus citharus* (Geoffroy St. Hilaire, 1809) and *C. latus* Müller et Troschel, 1844 – and the only representative of the latter family – *Distichodus affinis* Günther, 1873 – were studied cytogenetically (Rab et al. 1998; Simanovsky et al. 2022). All three studied species have exclusively biarmed karyotypes with 2n = 40, 44 and 48 (for *C. citharus, C. latus and D. affinis*, respectively). Six distichodontid species from the five genera – *Distichodus* Müller et Troschel, 1844; *Ichthyborus* Günther, 1864; *Nannocharax* Günther, 1867; and *Nannaethiops* Günther, 1872 and *Neolebias* Steindachner, 1894 – involved in this study represent the four out of six clades identified by molecular methods within the family (Arroyave et al. 2013; Lavoué et al. 2017).

The present study is aimed at an estimation of the divergence of the karyotype structure (the number and morphology of chromosomes) between and within the phylogenetically distant lineages of the family Distichodontidae. The concordance between differences in karyotype structure and the molecular phylogenies elaborated for the family Distichodontidae by the previous researchers is considered.

Material and methods

Sample acquisition and characteristics

Ethiopian material was obtained from tributaries of the Sobat River, a tributary of the White Nile, in southwestern Ethiopia (Table 1). Fish were collected by the Joint Ethio-Russian Biological Expedition (JERBE) with the permissions of the National Fisheries and Aquatic Life Research Center under the Ethiopian Institute of Agricultural Research (EIAR) and the Ethiopian Ministry of Science and Technology. Two individuals – male and female – of *Nannaethiops bleheri* Géry et Zarske, 2003, collected from the roadside ditch in the interfluve of the Alvero and Gilo rivers (between towns

of Abobo and Funido, 7°45.307'N, 34°15.639'E) were karyotyped. The rest of karyotyped Ethiopian material was obtained from the two localities: (1) Alvero River just downstream of the Abobo Dam (7°52.503'N, 34°29.960'E) and (2) Baro River at the City of Gambela (8°14.878'N, 34°34.044'E). Two males and a female of *Distichodus engycephalus* Günther, 1864, as well as a female of *Ichthyborus besse* (Joannis, 1835), were collected at locality 1. Two males of *I. besse* and a female of *Nannocharax niloticus* (Joannis, 1835), were collected at locality 2.

Four specimens (a female, two males and one unsexed) of an unidentified species representing the genus *Nannaethiops* and seven specimens (five females and two males) of *Neolebias unifasciatus* Steindachner, 1894 were purchased from the Nigerian aquarium fish dealers through the mediation of the company Aqua Logo Engineering (https://www.aqualogo-engineering.ru).

After colchicine treatment, fish were euthanized with an overdose of tricaine methanesulfonate (MS-222), identified, measured with an accuracy of 1 mm, dissected for gonad examination and tissue sampling, and preserved in 10% formaldehyde or 70% ethanol. Species identification was done based on morphological characters (Gosse and Coenen 1990; Golubtsov et al. 1995). The experiments were carried out in accordance with the rules of the Severtsov Institute of Ecology and Evolution (IEE) and approved by IEE's Ethics Committee. Vouchers are deposited at the Severtsov Institute of Ecology and Evolution (Moscow), under provisional labels of JERBE.

DNA extraction, PCR amplification, and sequencing

In order to clarify the phylogenetic position of *Nannaethiops* and *Neolebias* specimens, two genetic markers – *Cytochrome oxidase subunit I (COI)* and *16S ribosomal RNA (16S rRNA)* – were studied in 13 karyotyped fish and one additional specimen of *N. bleheri* from an another location in Ethiopia (Suppl. material 1: table S1). We extracted total genomic DNA from the ethanol-preserved tissues using the DiatomDNA Prep 100 (Izogen, Moscow) extraction kit. The PCR mixture contained 5 pmol of each primer and the precast PCR mixture from DIALAT Ltd (Russia). The primers used for *COI* amplification were designed by Ward et al. (2005): FishF1-5′TCAACCAACCACAAAGACATTGGCAC3′ and FishR1-5′TAGACTTCTGGGTGGCCAAAGAATCA3′. The PCR cycle profiles were as follows: 5 min initial denaturation at 94 °C, followed by 35 cycles of 1 min at 94 °C, annealing for 45 sec at 55 °C, extension for 1 min at 72 °C; final extension for 7 min at 72 °C. The primers 8f-5′AGAGTTTGATCCTGGCTCAG3′ (Edwards et al. 1989) and 1492r-5′GGTTACCTTGTTACGACTT3′ (Stackebrandt and Liesack 1993) were employed for the *16S rRNA* amplification. The PCR cycle profiles were as follow: 3 min initial denaturation at 94 °C, followed by 30 cycles of 30 sec at 94 °C, annealing for 30 s at 50 °C, extension for 30 sec at 72 °C; final extension for 7 min at 72 °C. PCR products were visualized by electrophoresis in 1.5% agarose gel in TBE buffer with addition of ethidiumbromide. DNA sequencing was performed using an Applied Biosystems 3500 genetic analyzer. All new DNA sequences were deposited in GeneBank (Suppl. material 1: table S1).
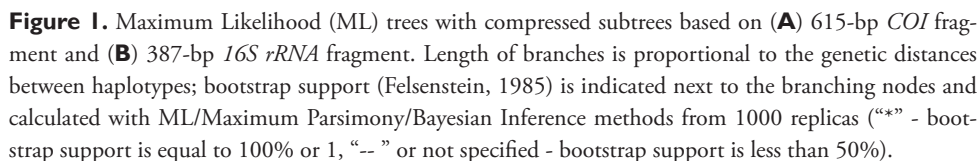
Sequence alignment and phylogenetic reconstruction

Preprocessing and alignment of the obtained sequences was carried out using Seq-Man Pro 7.1.0 and BioEdit 5.0.9. For phylogenetic reconstruction all sequences of the two markers (*COI* and *16S rRNA*) available in GenBank for *Nannoethiops* and *Neolebias* specimens were used. These sequences are listed below. The distichodontid species *Belonophago hutsebouti* Giltay, 1929, *Distichodus nefasch* (Bonnaterre, 1788) and *D. sexfasciatus* Boulenger, 1897, as well as citharinid *Citharinus citharus* (Geoffroy Saint-Hilaire, 1809), were selected as outgroups. The GenBank accession numbers for outgroups are given in Suppl. material 1: table S1.

Comparative material included the GenBank sequences of six species representing the genera *Nannoethiops* and *Neolebias* for *CO1* and seven such species for *16S rRNA* (Fig. 1, Suppl. material 1: table S1). For *CO1*, these were *Nannaethiops bleheri* from Ethiopia (the GenBank accession number KF541848, Arroyave et al. 2013), *Nannaethiops gracilis* (Matthes, 1964) (KF541851, KF541852, Arroyave et al. 2013), *Nannaethiops unitaeniatus* Günther, 1872 (KF541849, KF541850, Arroyave et al. 2013), *Neolebias ansorgii* Boulenger, 1912 (KF541858, KF541859, KF541860, Arroyave et al. 2013; HM418212, HM418213, Sonet et al. 2019), *Neolebias trewavasae* Poll et Gosse, 1963 (KF541853, KF541857, Arroyave et al. 2013) and *Neolebias trilineatus* Boulenger, 1899 (KF541854, KF541855, KF541856, Arroyave et al. 2013; KT193336, Decru et al. 2016; HM418214, HM418215, MK074510, MK074511, Sonet et al. 2019), all from West Africa. For *16S rRNA*, these were *Nannaethiops bleheri* from Ethiopia (JX985104, Lavoué et al. 2017), *Nannaethiops unitaeniatus* (JX985105, Lavoué et al. 2017), *Neolebias ansorgii* (AY788058, Calcagnotto et al., 2005; JX985107, Lavoué et al. 2017), *Neolebias powelli* Teugels et Roberts, 1990 (AY788061, Calcagnotto et al. 2005), *Neolebias trewavasae* (JX985132, Lavoue et al. 2017), *Neolebias trilineatus* (AY788063, Calcagnotto et al. 2005) and *Neolebias unifasciatus* Steindachner, 1894 (JX985103, Lavoué et al. 2017), all from West Africa.

For phylogenetic reconstruction, we used Maximum Likelihood (ML), Maximum Parsimony (MP) (Nei and Kumar 2000) and Bayesian Inference (BI) methods. For ML, the chosen models of molecular evolution were as follows: Hasegawa-Kishino-Yano (HKY +G+I; parameter +G = 1.77; +I = 0.6) (Hasegawa et al. 1985) for *COI* and Tamura-Nei (TN93+G; parameter +G = 0.13) (Tamura and Nei 1993) for *16S rRNA*. For ML and MP, the bootstrap support for branch nodes was calculated with 1,000 replicates (Felsenstein 1985). Genetic distances and other parameters for phylogenetic ML and MP analysis were calculated using the MEGA X software package (Kumar et al. 2018). The nucleotide substitution model for BI was selected by means of the Bayesian Information Criterion (BIC) as implemented in jModel-Test (Posada 2008). BI was carried out in MrBayes version 3.1.2 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003) and implemented using Markov Chain Monte Carlo algorithm for 10,000 generations with a sampling period of 1,000 generations.

Figure 1. Maximum Likelihood (ML) trees with compressed subtrees based on (**A**) 615-bp *COI* fragment and (**B**) 387-bp *16S rRNA* fragment. Length of branches is proportional to the genetic distances between haplotypes; bootstrap support (Felsenstein, 1985) is indicated next to the branching nodes and calculated with ML/Maximum Parsimony/Bayesian Inference methods from 1000 replicas ("*" - bootstrap support is equal to 100% or 1, "-- " or not specified - bootstrap support is less than 50%).

Cytogenetic analysis

Before preparation, fish were treated intraperitoneally with 0.1% colchicine (0.01 ml / 1 g of their weight; for Ethiopian material, under field conditions) or 0.025% colchicine (0.01 ml / 1 g of their weight; for Nigerian material, under laboratory conditions) for 3–5 hours. After euthanasia, chromosome preparations were obtained from kidney tissue following Kligerman and Bloom (1977) for Ethiopian and Nigerian material or from

Table 1. Species, fish standard length (SL), numbers of individuals (N) and metaphases (N_{mt}) studied, and collection site. UD – undetermined sex.

Species	SL, mm	N	N_{mt}	Collection site
Distichodus engycephalus	149–163	3 (1♀, 2♂)	30	Alvero River
Ichthyborus besse	110	1 (1♀)	25	
	103–118	2 (2♂)	20	Baro River
Nannocharax niloticus	51	1 (1♀)	10	
Nannaethiops bleheri	19–23	2 (1♀, 1♂)	20	Interfluve of the Alvero and Gilo rivers
Nannaethiops sp.	23–26	4 (1♀, 2♂, 1UD)	40	West Africa (fish store)
Neolebias unifasciatus	25–31	7 (5♀, 2♂)	81	

kidney, spleen, intestine and liver following Bertollo et al. (2015) for Nigerian material with some modifications for both protocols, as described in Simanovsky et al. (2022). The chromosome spreads were stained conventionally with 4% Giemsa solution in a phosphate buffer solution at pH 6.8 for 8 min and then analysed using an Axioplan 2 Imaging microscope (Carl Zeiss, Germany) equipped with a CV-M4+CL camera (JAI, Japan) and Ikaros software (MetaSystems, Germany). Final images were processed using Photoshop software (Adobe, USA). Karyotypes were arranged according to the centromere position following the nomenclature of Levan et al. (1964), but modified as metacentric (m), submetacentric (sm) and subtelocentric/acrocentric (st/a). Chromosome pairs were arranged according to their size in each chromosome category. To determine the chromosomal arm number per karyotype (fundamental number, FN), metacentrics and submetacentrics were considered as biarmed, and subtelocentrics/acrocentrics as monoarmed. The total numbers of complete metaphases studied for each species is presented in Table 1.

Results and discussion

Molecular phylogenetic analysis

An analysis of 615 bp of the mitochondrial *CO1* in 13 individuals representing the genera *Nannoethiops* and *Neolebias* and 387 bp of the mitochondrial *16S rRNA* in seven individuals representing the same genera included the Ethiopian samples of *Nannaethiops bleheri*, as well as the West African samples (from the Nigerian aquarium fish dealers) of the genera *Nannoethiops* and *Neolebias*. The alignment used for phylogenetic reconstructions included 47 *CO1* sequences and 18 *16S rRNA* sequences in total.

The thirteen newly obtained *COI* sequences were collapsed in six haplotypes deposited in GenBank with accession numbers OQ891056–OQ891061. Two of them made an independent cluster corresponding to *Neolebias unifasciatus* (Fig. 1). Genetic distance (p-distance) was 0.002 between haplotypes. Two more cluster together with a sequence of *Nannaethiops bleheri* deposited earlier by Arroyave et al. (2013) (p-d 0.002–0.003). The remaining two new haplotypes formed an independent cluster recognized by us as *Nannaethiops* sp. that is a sister to *Nannaethiops bleheri* (Fig. 1). In general,

haplotypes of the genera *Nannoethiops* and *Neolebias* comprise a monophyletic group without a clear division into two genera (Fig. 1). This is fully consistent with the conclusion of Géry and Zarske (2003) – supported by Arroyave et al. (2013) and Lavoué et al. (2017) – who considered *Neolebias* as a junior synonym of *Nannaethiops*.

The seven newly obtained *16S rRNA* sequences were collapsed in three haplotypes. One of them appeared to be identical to the sequence (JX985103) earlier deposited in GenBank for *Neolebias unifasciatus* by Lavoue et al. (2017). Two other haplotypes we deposited in GenBank with the accession numbers OQ911366 and OQ911367. The former cluster together with the haplotype deposited for *Nannaethiops bleheri* by Lavoué et al. (2017) (p-d 0.003); the latter belongs to the *Nannaethiops* sp. clade.

In summary, both the *COI* and *16S rRNA* analyses support: (1) our identification of *Nannaethiops bleheri*; (2) the distinctiviness of *Nannaethiops* sp.; and (3) the *16S rRNA* analysis supports our identification of *Neolebias unifasciatus*.

Cytogenetic analysis

The karyotype of *Distichodus engycephalus* has 2n = 52 and consists of 30 metacentrics and 22 submetacentrics, FN = 104 (Fig. 2). It differs substantially from the karyotype of *D. affinis* (2n = 48, 32m + 16sm, FN = 96) reported by Rab et al. (1998) (Table 2). No distinguishable sex chromosomes were observed in complements of *D. engycephalus*, similar to the finding by Rab et al. (1998) in *D. affinis*. This is true for all distichodontids studied by us.

The karyotype of *Ichthyborus besse* has 2n = 46 and consists of 40 metacentrics and 6 submetacentrics, FN = 92. The karyotype of *Nannocharax niloticus* has 2n = 54 and consists of 46 metacentrics, 6 submetacentrics, and 2 subtelocentrics/acrocentrics, FN = 106. The latter species exhibits the highest numbers of chromosomes and chromosome arms among all distichodontids studied (Table 2).

The karyotypes of *Nannaethiops bleheri*, *Nannaethiops* sp. and *Neolebias unifasciatus* appeared to be similar. These karyotypes have 2n = 50 and consists of 38 metacentric, 10 submetacentric, and 2 subtelocentrics/acrocentrics, FN = 96. These taxa, along with *Nannocharax niloticus*, possess the only pair of monoarmed chromosomes; the other distichodontids studied have exclusively biarmed chromosomes in their compliments.

The molecular phylogeny of the order Cithariniformes as it is reconstructed by Arroyave et al. (2013) and Lavoué et al. (2017) is as follows. The family Citharinidae is a sister group to the family Distichodontidae. *Xenocharax* Günther, 1867 comprises a sister group to all other distichodontids. *Nannaethiops* + *Neolebias* represent a sister group to other distichodontids excluding *Xenocharax*. *Monostichodus* Vaillant in Rivière, 1886 + *Ichthyborus* comprise a sister group to all remaining distichodontids. Branching of the remaining three clades (*Distichodus* + *Paradistichodus* Pellegrin, 1922, *Nannocharax*, *Belonophago* Giltay, 1929 + *Phago* Günther, 1865 with the related genera) is not well supported and different in Arroyave et al. (2013) and Lavoué et al. (2017). Nevertheless the monophyly of the each of three groups is well supported. Thus, we analysed the representatives of four clades out of six excluding *Xenocharax* and *Belonophago* + *Phago* with the related genera.

Figure 2. Karyotypes of six representatives of the family Distichodontidae. Scale bar: 10 μm.

There is an apparent correspondence between molecular phylogeneetic and cytogenetic data. There are differences in cytogenetic characteristics between *Distichodus* (2n = 48–52), *Ichthyborus* (2n = 46), *Nannocharax* (2n = 54) and *Nannaethiops* + *Neolebias* (2n = 50) representing the four different clades revealed by phylogenetic analyses. Moreover, there are differences in cytogenetic characteristics between all these distichodontids and the two species of *Citharinus* (2n = 40–44) (Table 2). These data clearly suggest a substantial role of chromosome fusions/fissions in the evolution of Cithariniformes karyotypes.

Regarding variation within the clades, we see two opposing trends. Two species of *Distichodus*, *D. affinis* and *D. engycephalus*, differ both in diploid chomosome numbers and karyotypic formulae. On the contrary, no differences were found between karyotypes of *Nannaethiops bleheri*, *Nannaethiops* sp. and *Neolebias unifasciatus* representing another clade. The latter point corroborates the position of authors who considered *Neolebias* as a junior synonym of *Nannaethiops* (Géry and Zarske 2003, Arroyave et al. 2013, Lavoué et al. 2017). Variability of karyotype structure in the genus *Distichodus* makes it possible to use the cytogenetic data in its taxonomy when a sufficient array of such data is accumulated. The same is true for the family Distichodontidae as a whole.

Due to the lack of data on the diversity of karyotypes in both the families Citharinidae and Distichodontidae it might be premature to make assumptions about the trend of karyotype evolution in the order Cithariniformes. The great prevalence of biarmed chromosomes (the karyotypes of most species contain exclusively biarmed chromosomes) is a distinctive characteristic of Cithariniformes compared to Characiformes and Siluriformes,

Table 2. Cytogenetically studied taxa of the order Characiniformes. Diploid chromosome number (2n), karyotypic formula, fundamental number (FN) and geographic origin.

Taxon	2n	Karyotypic formula	FN	Origin	References
Family Citharinidae					
Citharinus citharus (Geoffroy St. Hilaire, 1809)	40	26m + 14sm	80	West Africa (fish store)	Simanovsky et al. 2022
Citharinus latus Muller et Troschel, 1844	44	30m + 14sm	88	White Nile Basin, southwest Ethiopia	Simanovsky et al. 2022
Family Distichodontidae					
Distichodus affinis Günther, 1873	48	32m + 16sm	96	Unknown (aquarium stock)	Rab et al. 1998
Distichodus engycephalus Günther, 1864	52	30m + 22sm	104	White Nile Basin, southwest Ethiopia	This study
Ichthyborus besse (Joannis, 1835)	46	40m + 6sm	92	White Nile Basin, southwest Ethiopia	This study
Nannocharax niloticus (Joannis, 1835)	54	46m + 6sm + 2st/a	106	White Nile Basin, southwest Ethiopia	This study
Nannaethiops bleheri Géry et Zarske, 2003	50	38m + 10sm + 2st/a	98	White Nile Basin, southwest Ethiopia	This study
Nannaethiops sp.	50	38m + 10sm + 2st/a	98	West Africa (fish store)	This study
Neolebias unifasciatus Steindachner, 1894	50	38m + 10sm + 2st/a	98	West Africa (fish store)	This study

sister groups to Characiniformes. Characiformes and Siluriformes are characterized by karyotypes with various proportions of biarmed and monoarmed chromosomes (Arai 2011; Simanovsky et al. 2022). There is reason to suggest that the ancestral karyotype of Characiniformes consisted exclusively/predominantly of biarmed chromosomes. However, the karyotypes of representatives of the basal group of Distichodontidae – genus *Xenocharax* – have yet to be determined. Thus, the cytogenetic information about this genus and other unexamined taxa of Characiniformes would be of great interest.

Acknowledgements

References

Arai R (2011) Fish karyotypes – a Check List. Springer, 340 pp. https://doi.org/10.1007/978-4-431-53877-6

Arroyave J, Denton JSS, Stiassny MLJ (2013) Are characiform fishes Gondwanan in origin? Insights from a time-scaled molecular phylogeny of the Citharinoidei (Ostariophysi: Characiformes). PLoS ONE 8(10): e77269. https://doi.org/10.1371/journal.pone.0077269

Bertollo LAC, Cioffi MB, Moreira-Filho O (2015) Direct chromosome preparation from freshwater teleost fishes. In: Ozouf-Costaz C, Pisano E, Foresti F, de Almeida Toledo LF (Eds) Fish Cytogenetic Techniques: Ray-Fin Fishes and Chondrichthyans, CRC Press, Boca Raton, 21–26. https://doi.org/10.1201/b18534

Calcagnotto D, Schaefer SA, DeSalle R (2005) Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. Molecular Phylogenetics and Evolution 36(1): 135–153. https://doi.org/10.1016/j.ympev.2005.01.004

Decru E, Moelants T, De Gelas K, Vreven E, Verheyen E, Snoeks J (2016) Taxonomic challenges in freshwater fishes: a mismatch between morphology and DNA barcoding in fish of the north-eastern part of the Congo basin. Molecular Ecology Resources 16: 342–352. https://doi.org/10.1111/1755-0998.12445

Dornburg A, Near TJ (2021) The emerging phylogenetic perspective on the evolution of Actinopterygian fishes. Annual Review of Ecology, Evolution, and Systematics 52(1): 427–452. https://doi.org/10.1146/annurev-ecolsys-122120-122554

Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC (1989) Isolation and direct complete nucleotide determination of entire genes: characterization of a gene coding for 16S ribosomal RNA. Nucleic Acids Research 17: 7843–7853. https://doi.org/10.1093/nar/17.19.7843

Eschmeyer WN, Fricke R, van der Laan R (2023) Catalog of Fishes: Genera, Species, References. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [accesed 09.06.2023]

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Froese R, Pauly D (2023) FishBase. http://www.fishbase.org [accessed 01.08.2023]

Géry J, Zarske A (2003) *Nannaethiops bleheri* sp. n. – ein neuer, afrikanischer Salmler (Teleostei, Characiformes, Distichodidae) vom oberen Weißen Nil in Südwestäthiopien. Zoologische Abhandlungen (Dresden) 53: 37–45.

Golubtsov AS, Darkov AA, Dgebuadze YY, Mina MV (1995) An artificial key to fish species of the Gambela region (the White Nile basin in the limits of Ethiopia). Joint Ethio-Russian Biological Expedition, Addis Abeba, 84 pp.

Gosse J-P, Coenen EJ (1990) Distichodontidae. In: Lévêque C, Paugy D, Teugels GG (Eds) Faune des poissons d'eaux douces et saumâtres de l'Afrique de l'Ouest, T. 1, Paris, ORSTOM; Tervuren, MRAC, 237–260.

Hasegawa M, Kishino H, Yano T (1985) Dating the human-ape split by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22: 160–174. https://doi.org/10.1007/BF02101694

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754–755. https://doi.org/10.1093/bioinformatics/17.8.754

Kligerman AD, Bloom SE (1977) Rapid chromosome preparations from solid tissues of fishes. Journal of the Fisheries Research Board of Canada 34(2): 266–269. https://doi.org/10.1139/f77-039

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35: 1547–1549. https://doi.org/10.1093/molbev/msy096

Melo BF, Sidlauskas BL, Near TJ, Roxo FF, Ghezelayagh A, Ochoa LE, Stiassny MLJ, Arroyave J, Chang J, Faircloth BC, Macguigan DJ, Harrington RC, Benine RC, Burns MD, Hoekzema K, Sanches NC, Maldonado-Ocampo JA, Castro RMC, Foresti F, Alfaro ME, Oliveira C (2022) Accelerated diversification explains the exceptional species richness of tropical characoid fishes. Systematic Biology 71(1): 78–92. https://doi.org/10.1093/sysbio/syab040

Lavoué S, Arnegard ME, Rabosky DL, McIntyre PB, Arcila D, Vari RP, Nishida M (2017) Trophic evolution in African citharinoid fishes (Teleostei: Characiformes) and the origin of intraordinal pterygophagy. Molecular Phylogenetics and Evolution 113: 23–32. https://doi.org/10.1016/j.ympev.2017.05.001

Levan A, Fredga K, Sandberg A (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201–220. https://doi.org/10.1111/j.1601-5223.1964.tb01953.x

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, 348 pp.

Nelson JS, Grande T, Wilson MVH (2016) Fishes of the World, Fifth Edition, John Wiley & Sons, Inc., Hoboken, New Jersey, 707 pp. https://doi.org/10.1002/9781119174844

Posada D (2008) jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25: 1253–1256. https://doi.org/10.1093/molbev/msn083

Rab P, Rabova M, Ozouf-Costaz C (1998) Karyotype analysis of the African citharinid fish *Distichodus affinis* (Osteichthyes, Characiformes) by different staining techniques. Journal of African Zooljgy 112: 185–191.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Simanovsky SA, Medvedev DA, Tefera F, Golubtsov AS (2022) First cytogenetic data on Afrotropical lutefishes (Citharinidae) in the light of karyotype evolution in Characiformes. Comparative Cytogenetics 16(2): 143–150. https://doi.org/10.3897/compcytogen.v16.i2.79133

Sonet G, Snoeks J, Nagy ZT, Vreven E, Boden G, Breman FC, Decru E, Hanssens M, Zamba AI, Jordaens K, Mamonekene V, Musschoot T, Van Houdt J, Van Steenberge M, Wamuini SL, Verheyen E (2019) DNA barcoding fishes from the Congo and the Lower Guinean provinces: Assembling a reference library for poorly inventoried fauna. Molecular Ecology Resources 19: 728–743. https://doi.org/10.1111/1755-0998.12983

Stackebrandt E, Liesack W (1993) Nucleic acids and classification. In: Goodfellow M, O'Donnell AG (Eds) Handbook of new bacterial systematics. London, Academic Press, 152–189.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10: 512–526. https://doi.org/10.1093/oxfordjournals.molbev.a040023

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005) DNA barcoding of Australia's fish species. Philosophical Transactions of the Royal Society B: Biological sciences 360: 1847–1857. https://doi.org/10.1098/rstb.2005.1716

ORCID

Sergey A. Simanovsky https://orcid.org/0000-0002-0830-7977
Dmitry A. Medvedev https://orcid.org/0000-0001-8560-8186
Alexander S. Golubtsov https://orcid.org/0000-0002-8317-7527

Supplementary material 1

Supporting information
Authors: Sergey A. Simanovsky, Dmitry A. Medvedev, Fekadu Tefera, Alexander S. Golubtsov
Data type: pdf
Explanation note: supplementary text, tables S1–S3, figures S1–S4, supplementary references.
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/10.3897/compcytogen.17.107744.suppl1

SHORT COMMUNICATION

Cell culture and karyotypic description of *Pseudophryne coriacea* (Keferstein, 1868) (Amphibia, Anura) from the New South Wales Central Coast

Richard Mollard[1,2], Michael Mahony[3]

1 *Melbourne Veterinary School, Faculty of Science, The University of Melbourne, Parkville, 3052, Australia*
2 *Amphicell Pty Ltd, Cairns, Queensland, Australia* **3** *School of Environmental and Life Sciences, University of Newcastle, Callaghan, New South Wales, 2308, Australia*

Corresponding author: Richard Mollard (rmollard@unimelb.edu.au)

Citation: Mollard R, Mahony M (2023) Cell culture and karyotypic description of *Pseudophryne coriacea* (Keferstein, 1868) (Amphibia, Anura) from the New South Wales Central Coast. Comparative Cytogenetics 17: 263–272. https://doi.org/10.3897/compcytogen.17.113526

Abstract

The karyotype of the IUCN least concern red-backed toadlet *Pseudophryne* (*P.*) *coriacea* (Keferstein, 1868) from the New South Wales Central Coast is described following tissue culture of toe clipping macerates and conventional DAPI staining. The diploid number is 2n = 24. The karyotype is represented by six large and five small chromosomal pairs and one very small chromosomal pair. The very small chromosome 12 is 12% the size of chromosome 1. One of the large chromosomes is subtelocentric, two of the large chromosomes are submetacentric and the remaining chromosomes are metacentric. The putative nucleolus organiser region (NOR) is observed on chromosome 4. The diploid number and location of the putative NOR correlates to that of the previously published IUCN critically endangered *P. corroboree* (Moore 1953) and unpublished descriptions of the *P. coriacea* karyotype. This is the first described cell culture of a species from the genus *Pseudophryne* Fitzinger, 1843, first published analysis of the *P. coriacea* karyotype and the first published analysis of centromeric allocation of this genus. Globally there exists a large inventory of tissue samples in cryobanks that are not associated with known recovery mechanisms such as basic cell culture techniques. Detailed cytogenetic analyses of these cryobanked samples are therefore not possible. This work therefore enables: (i) a comparison of the *P. coriacea* karyotype with that of the critically endangered *P. corroboree* and (ii) a benchmark for repeat and future cytogenetic and genomic analyses of cryostored samples of this genus.

Introduction

Recently documented amphibian declines resulting from disease and habitat destruction have placed nearly one third of all amphibian species at risk of extinction (Silla and Byrne 2019). Examples of animals at the forefront of this decline are *P. corroboree* (the southern corroboree frog) and *P. pengilleyi* (Wells and Wellington 1985), both critically endangered species for which restorative husbandry programs are required and cryobanking proposed (Morgan et al. 2008; Kouba et al. 2013; McFadden et al. 2013; Clulow and Clulow 2016; Skerratt et al. 2016; Rojahn et al. 2018). To date, captive breeding programs have demonstrated some success towards the long-term reintroduction of these animals into the wild (McFadden et al. 2013; Silla et al. 2018). No examples of successful cell culture with or without cryobanking to provide a non-invasive technique for long term auxiliary and repeat genomic monitoring or assisted reproduction programs, however, have been reported for any representative of this genus.

The red-backed toadlet *P. coriacea*, an IUCN least concern listed species, is endemic to the east coast and ranges of Australia, north of Sydney to southern Queensland (White 1993; Donnellan et al. 2012). The genus *Pseudophryne* Fitzinger, 1843 comprises 14 known species (O'Brien et al. 2018) for which the karyotype of only *P. corroboree* has been published (Mahony and Robinson 1986). Reports from unpublished data (Morescalchi and Ingram 1974; Mahony and Robinson 1986) suggest a highly conserved 2n = 24 karyotype across nine species in this genus, including *P. coriacea*, with an NOR also highly conserved on chromosome set 4. Centromeric positions remain to be described for any species in this genus. Detailed karyotypic information available for interspecies comparisons within this genus and associated information for assistance in conservation management programs are therefore wanting (Mahony and Robinson 1986; Potter and Deakin 2018).

This report serves four aims: (1) to demonstrate successful cultivation, passaging and cryopreservation of cells from *P. coriacea*, (2) to formally describe their karyotype including centromeric positions and NOR locations, (3) to facilitate future genetic comparisons for conservation management programs of species within this genus, and (4) provide a tissue resource for future cytogenetic and genomic work that would not require harming living animals.

Material and methods

Ethics

Relevant Australian State governmental and institutional ethics, licenses and permissions were obtained and the described research was conducted in accordance with The Code of Ethics of the World Medical Association of The Declaration of Helsinki and in compliance with the EU Directive 2010/63/EU for animal experiments. The animal specimen was collected by Michael Mahony under New South Wales National Parks Scientific Licence SL00190.

Tissue culture and cryopreservation

Toe clippings obtained from an unsexed and deceased *P. coriacea* toadlet, euthanised for alternative research purposes, were prepared for culture and karyotyping according to previously described and detailed methods (Mollard 2018; Mollard et al. 2018; Bui-Marinos et al. 2022). Tissue was first rinsed in 70% v/v ethanol (Sigma Aldrich) and then washed consecutively three times in 0.22 µM (Merck Millipore) syringe (Terumo) filtered Amphibian Ringer's Solution (AR; Cold Spring Harbor Protocols) at 4 °C. Tissues were macerated with fine scissors (Solingen) and transferred to 24 well plates (Falcon Multiwell™; GIBCO) containing preequilibrated Dulbecco's Modified Eagle Medium (DMEM; Thermo Fisher Scientific) at room temperature (26 °C) in a 5% CO_2/ 95% air atmospheric incubator (Steri-Cycle CO_2 Incubator; FORMA) according to previously described methods (Speare and Smith 1992; Fukui et al. 2003; Ferris et al. 2010; Strauß et al. 2011) with gentamicin replaced with 1000 units/mL penicillin (Sigma Aldrich) and 1000 µg/mL streptomycin (Sigma Aldrich). After four days, when individual cells could be observed to emerge from the tissue pieces, one half of the media was changed daily for three days and thereafter the entire media was changed every one to two days. For passaging, cultures were rinsed with AR and adherent cells were trypsinised with 0.25% trypsin/0.02% ethylenediaminetetraacetic acid solution (Sigma Aldrich) and replated at a 1:3 dilution. For cryopreservation, following trypsinisation, cells were resuspended in 100 µl of culture DMEM containing 10% dimethyl sulfoxide (DMSO; Sigma Aldrich) in 1 ml cryotubes (Nunc®), placed at minus 80 °C (CSK Group) overnight and then transferred to liquid nitrogen for storage (Taylor Wharton). For culture following cryopreservation, cryovials were placed on ice until the medium was visibly thawed. DMSO in the cryovial was diluted to 0.5% with DMEM culture media and the media was then transferred to one well of a 24 well plate for reseeding. The medium was changed with fresh culture medium after 48 hours and then daily until karyotyping. Cells were photographed periodically using an Olympus IX70 – S8F2 inverted microscope, a ProgRes®C3 (Jenoptik, Germany) camera and ProRes® CapturePro Software Version 2.8.8.

Karyotyping

Karyotyping was performed according to modifications of previously described techniques (Howe et al. 2014). When culture wells had reached approximately 70% confluency, cells were treated for six hours with 0.1 µg/ml KaryoMAX® colcemid (GIBCO), removed from the culture dish with a two minute trypsin treatment, incubated in hypotonic 0.027 M Na_3Citrate (Sigma Aldrich) for five minutes, and fixed in Carnoy's fixative (Cold Spring Harbor) overnight at 4 °C. Microdrops were released from a 20 µl Gilson pipette onto ethanol cleaned glass microscope slides (Thermo Fisher Scientific) held at an approximately 45 degree angle from a height of approximately 20 cm, and above a water bath (Sigma Aldrich) preheated to 80 °C. Karyotype preparations were airdried overnight in a dust free environment. Spread cells were then stained with

4',6-diamino-2-phenylindole (DAPI; 500 ng/ml; Sigma Aldrich) and coverslipped (Menzel-Gläser) under Gelvatol mounting medium (Cold Spring Harbor Protocols). For numbering chromosomes, the largest chromosome was designated chromosome 1 and the remainder were designated in descending size order. Image J software with the Levan plugin (Levan et al. 1964) was used to measure chromosomal arm lengths. Metacentric, submetacentric and subtelocentric chromosomal designation were defined as a long arm to short arm ratios of 1–1.69, 1.7–2.99 and 3–6.99, respectively (Levan et al. 1964). Images were captured at 1000 × with an Olympus BX60 microscope, colour CCD Leica DFC425C camera, EL-6000 Leica light source and Leica LAS-AF and QCapture Pro7 Version 7.0.5 Build 4325 software (QImaging Inc, USA).

Results and discussion

Toe macerates from an unsexed *P. coriacea* were placed in culture and individual cells were observed as attached single cells or within small expanding cell masses during the following two weeks (Figs 1, 2). Under high power inverted phase contrast microscopy, and at day 18 (D18), a mixed cell population comprising spindle-shaped and ovoid/polygonal morphology was observed. Rounded/semi-detached cells were presumed to be mitotic cells. At D18 and approximately 40% confluency, cells were trypsinised and cryopreserved in liquid nitrogen.

Figure 1. *Pseudophryne coriacea*. Photographed by Michael Mahony at Wallingat State Forrest, New South Wales, Australia, 1982.

Pre-cryopreservation   Post-cryopreservation
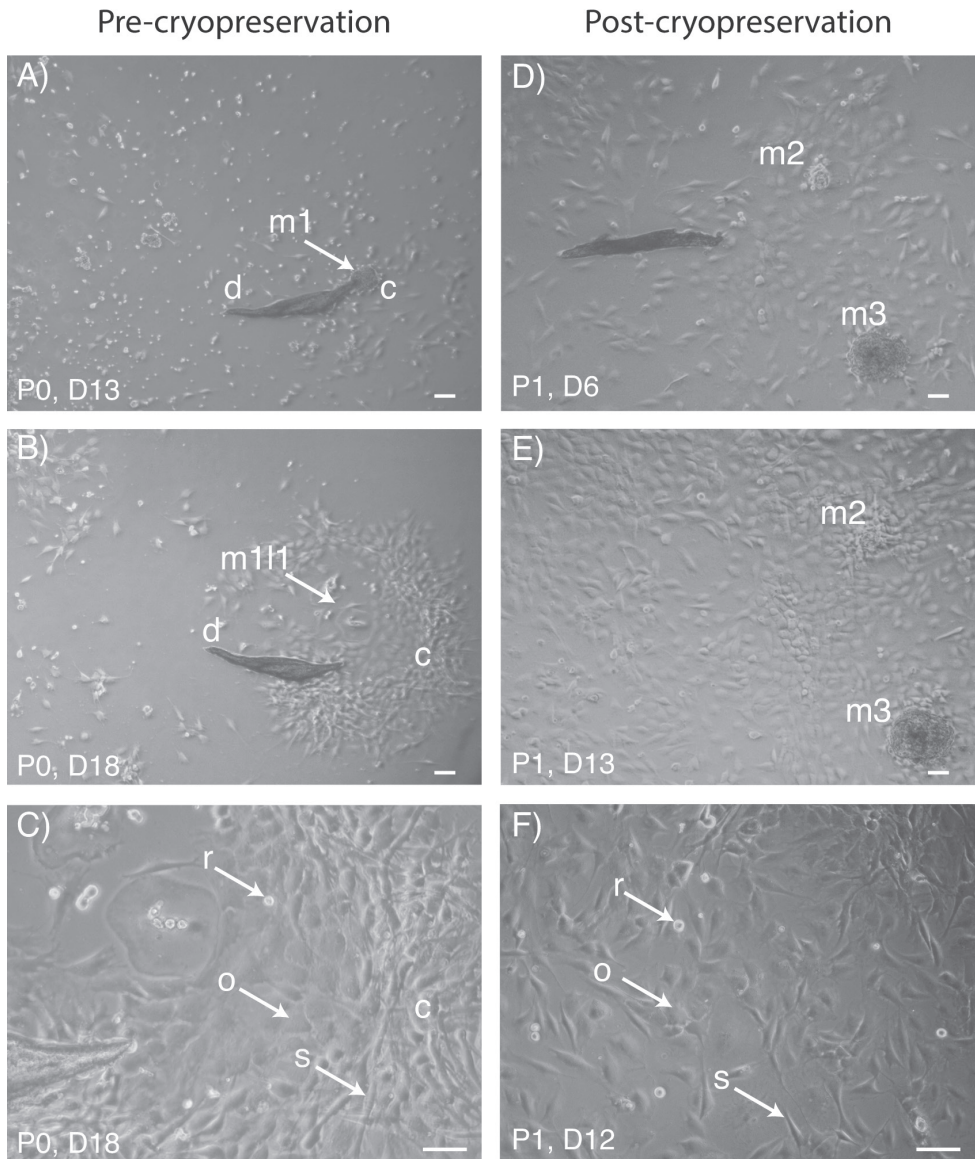
Figure 2. *Pseudophryne coriacea* macerated and cultured toe clippings **A–C** primary culture prior to cryopreservation (P0) **D–F** passage 1 cells, post cryopreservation (P1) **A, B, D, E** low power **C, F** high power **A, B** P0 cells form an expanding cluster (c) adjacent to the reference debris (d) at days 13 (D13) and D18 **C** P0 cells are either ovoid/polygonal (o) or spindle-shaped (s); rounded cells (r) are also observed **D, E** post-cryopreservation, P1 cells form two expanding mass reference points (m2 and m3) to reach approximately 70% confluency by D13 **F** post-cryopreservation P1 cells are both ovoid/polygonal (o) and spindle-shaped (s); rounded cells (r) are also observed. Scale bars: 10 µM.

Following a 12 month period of cryopreservation, cells were thawed into two separate wells of a 24 well plate. Passage 1 (P1) cells post-thawing attached within 48 hours as both cell clumps and single cells, and formed colony outgrowths resulting

in approximate 70% confluency by D13 (Fig. 1). Mixed spindle-shaped and ovoid/ polygonal cell populations were observed, as well as more condensed rounded cells characterizing a presumed mitotic phenotype. Cells from one dish were processed for karyotyping, while cells from the second dish were passaged and subsequently processed for cryopreservation. A total of 200 000 passage two *P. coriacea* cells were cryopreserved in a seven week total culture period.

Of the first 27 metaphase *P. coriacea* chromosome spreads identified and counted, 26 displayed a 2N = 24 chromosomal count and one displayed a chromosomal count of 15, with the latter a probable artefact of the cell spreading technique ($24_{incidence}$ = 96%; Fig. 3A). Six metaphase spreads were arranged in descending order of size to identify six large and five small chromosomal pairs and one very small chromosomal pair (Fig. 3B). A DAPI negative region was observed on the short arms of each chromosome pair number four (representing a presumptive NOR). Chromosomes 1, 2, 6, 7, 8, 9, 10, 11 and 12 are apparently metacentric, chromosomes 4 and 5 are apparently submetacentric and chromosome 3 is apparently subtelocentric (Table 1). Chromosome lengths were calculated relative to chromosome 1, not including measurements of the secondary restrictions on chromosome 4 (Table 1). The notably smaller chromosome 12 is 12% the size of chromosome 1.
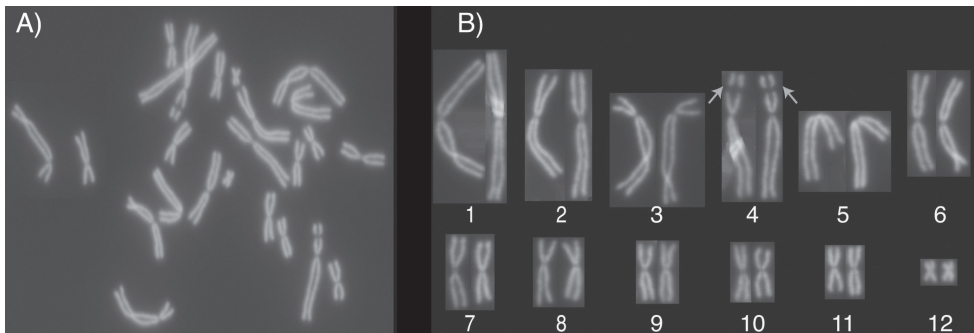
Figure 3. *Pseudophryne coriacea* karyotype **A** metaphase spread and **B** chromosomal pairs arranged in descending order relative to size and aligned by centromeric position. A 2N = 24 diploid chromosome number and the presence of a DAPI negative region on each of the short arms of chromosome 4 (arrows) are evident. Chromosomes 1 to 6 are larger, whereas chromosomes 7 to 11 are smaller in size, and chromosome 12 is smaller still.

While sperm cryobanking techniques have made significant advancements, methods for the cryopreservation of oocytes, embryos or amphibians suitable for conservation programs have not (Browne et al. 2019; Burger et al. 2022; Lampert et al. 2022). The cryostorage of karyotypically stable diploid nuclei amenable to recovery therefore represents a near term and important process for genomic and cytogenetic work and an additional resource for future conservation related assisted reproductive technologies (Kouba et al. 2013; Clulow and Clulow 2016; Zimkus et al. 2018).

Table 1. Chromosome log arm to short arm ratios with centromeric designations and overall relative lengths. The long arm to short arm ratios are provided from the average of six prepared and measured karyotypes +/- standard deviation. Relative lengths are provided from the percentage sum of each allocated and corresponding chromosomal set from the six individual karyotypes. The relative chromosome 5 length is smaller than relative chromosome 4 length at only four decimal places. Of note, inclusion of the secondary restriction measurement places chromosome 4 as chromosome 3, with a relative length of 0.7262.

	Chromosome number					
	1	**2**	**3**	**4**	**5**	**6**
Arm ratios	1.23 ± 0.1	1.62 ± 0.3	3.9 ± 0.6	2.15 ± 0.4	1.95 ± 0.2	1.27 ± 0.2
Designation	Metacentric	Metacentric	Subtelocentric	Submetacentric	Submetacentric	Metacentric
Relative length	1	0.7793	0.7104	0.7046	0.7043	0.6311
	Chromosome number					
	7	**8**	**9**	**10**	**11**	**12**
Arm ratios	1.57 ± 0.4	1.35 ± 0.3	1.30 ± 0.2	1.37 ± 0.2	1.3 ± 0.2	1.28 ± 0.3
Designation	Metacentric	Metacentric	Metacentric	Metacentric	Metacentric	Metacentric
Relative length	0.3941	0.3762	0.3365	0.3101	0.2823	0.1175

Conclusion

The overall *P. coriacea* karyotype with 2n = 24 and the location of the presumptive NOR on chromosome 4 agrees with previously unpublished reports for this species (Morescalchi and Ingram 1978; Mahony and Robinson 1980). It is also consistent with that described for *P. corroboree* (Mahony and Robinson 1980) and the unpublished data (Morescalchi and Ingram 1978; Mahony and Robinson 1980) on a further seven species from this genus including *P. pengilleyi*. The description of the centromeric positions and relative lengths for any species of this genus is novel. This study serves as a prototype for future comparisons of centromeric descriptions and karyotypes of species from the genus *Pseudophryne*, thus aiding conservation management programs.

Competing Interests

Acknowledgments

References

Browne RK, Silla AJ, Upton R, Della-Togna G, Marcec-Greaves R, Shishova NV, Uteshev VK, Proaño B, Pérez OD, Mansour N, Kaurova SA, Gakhova EN, Cosson J, Dyzuba B, Kramarova LI, McGinnity D, Gonzalez M, Clulow J, Clulow S (2019) Sperm collection and storage for the sustainable management of amphibian biodiversity. Theriogenology 133: 187–200. https://doi.org/10.1016/j.theriogenology.2019.03.035

Bui-Marinos MP, Todd LA, Douglas AJ, Katzenback BA (2022) So, you want to create a frog cell line? A guide to establishing frog skin cell lines from tissue explants. MethodsX 9: 101693. https://doi.org/10.1016/j.mex.2022.101693

Burger IJ, Lampert SS, Kouba CK, Morin DJ, Kouba AJ (2022) Development of an amphibian sperm biobanking protocol for genetic management and population sustainability. Conservation Physiology 10(1): coac032. https://doi.org/10.1093/conphys/coac032

Clulow J, Clulow S (2016) Cryopreservation and other assisted reproductive technologies for the conservation of threatened amphibians and reptiles: bringing the ARTs up to speed. Reproduction, Fertility and Development 28(8): 1116–1132. https://doi.org/10.1071/RD15466

Donnellan SC, Mahony M, Bertozzi T (2012) A new species of *Pseudophryne* (Anura: Myobatrachidae) from the central Australian ranges. Zootaxa 3476: 69–85. https://doi.org/10.11646/zootaxa.3476.1.4

Ferris DR, Satoh A, Mandefro B, Cummings GM, Gardiner DM, Rugg EL (2010) *Ex vivo* generation of a functional and regenerative wound epithelium from axolotl (*Ambystoma mexicanum*) skin. Development, Growth & Differentiation 52: 715–724. https://doi.org/10.1111/j.1440-169X.2010.01208.x

Fukui Y, Furue M, Myoishi Y, Sato JD, Okamoto T, Asashima M (2003) Long-term culture of *Xenopus* presumptive ectoderm in a nutrient-supplemented culture medium. Development, Growth & Differentiation 45: 499–506. https://doi.org/10.1111/j.1440-169X.2003.00717.x

Howe B, Umrigar A, Tsien F (2014) Chromosome preparation from cultured cells. JoVE: e50203. https://doi.org/10.3791/50203

Keferstein WM (1868) Ueber die Batrachier Australiens. Archiv für Naturgeschichte Berlin 34: 251–290. https://doi.org/10.5962/bhl.part.20476

Kouba AJ, Lloyd RE, Houck ML, Silla AJ, Calatayud N, Trudeau VL, Clulow J, Molinia F, Langhorne C, Vance C, Arregui L, Germano J, Lermen D, Della Togna G (2013) Emerging trends for biobanking amphibian genetic resources: the hope, reality and challenges for the next decade. Biological Conservation 164: 12. https://doi.org/10.1016/j.biocon.2013.03.010

Lampert SS, Burger IJ, Julien AR, Gillis AB, Kouba AJ, Barber D, Kouba CK (2022) Sperm Cryopreservation as a Tool for Amphibian Conservation: Production of F2 Generation Offspring from Cryo-Produced F1 Progeny. Animals (Basel) 13(1): 53. https://doi.org/10.3390/ani13010053

Levan A, Fredga K, Sandberg AA (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201–220. https://doi.org/10.1111/j.1601-5223.1964.tb01953.x

Mahony MJ, Robinson ES (1980) Polyploidy in the australian leptodactylid frog genus *Neobatrachus*. Chromosoma 81: 199–212. https://doi.org/10.1007/BF00285949

Mahony MJ, Robinson ES (1986) Nucleolar organiser region (NOR) location in karyotypes of Australian ground frogs (family Myobatrachidae). Genetica 68: 119–127. https://doi.org/10.1007/BF02424409

McFadden M, Hobbs R, Marantelli G, Harlow P, Banks C, Hunter D (2013) Captive management and breeding of the critically endangered Southern Corroboree frog (*Pseudophryne corroboree*) (Moore, 1953) at Taronga and Melbourne Zoos. Amphibian & Reptile Conservation 5: 70–87. https://oa.mg/work/2184318371

Mollard R (2018) Karyomaps of cultured and cryobanked *Litoria infrafrenata* frog and tadpole cells. Data in Brief 18: 1372–1377. https://doi.org/10.1016/j.dib.2018.04.025

Mollard R, Mahony M, Marantelli G, West M (2018) The critically endangered species *Litoria spenceri* demonstrates subpopulation karyotype diversity. Amphibian & Reptile Conservation 12(2): 28–36 [(e166)]. http://amphibian-reptile-conservation.org/pdfs/Volume/Vol_12_no_2/ARC_12_2_[Special_Section]_28-36_e166_high_res.pdf

Moore JA (1953) A new species of *Pseudophryne* from Victoria. Proceedings of the Linnean Society of New South Wales 78: 179–180. https://www.biodiversitylibrary.org/page/34949150

Morescalchi A, Ingram GJ (1974) New chromosome numbers in Australian Leptodactylidae (Amphibia, Salientia). Experientia 30: 1134–1136. https://doi.org/10.1007/BF01923651

Morescalchi A, Ingram GJ (1978) Cytotaxonomy of the myobatraehid frogs of the genus *Limnodynastes*. Experientia 34: 584. https://doi.org/10.1007/BF01923651

Morgan MJ, Hunter D, Pietsch R, Osborne W, Keogh JS (2008) Assessment of genetic diversity in the critically endangered Australian corroboree frogs, *Pseudophryne corroboree* and *Pseudophryne pengilleyi*, identifies four evolutionarily significant units for conservation. Molecular Ecology 17: 3448–3463. https://doi.org/10.1111/j.1365-294X.2008.03841.x

O'Brien DM, Keogh JS, Silla AJ, Byrne PG (2018) The unexpected genetic mating system of the red-backed toadlet (*Pseudophryne coriacea*): A species with prolonged terrestrial breeding and cryptic reproductive behaviour. Molecular Ecology 27: 3001–3015. https://doi.org/10.1111/mec.14737

Potter S, Deakin JE (2018) Cytogenetics: an important inclusion in the conservation genetics toolbox. Pacific Conservation Biology 24: 280–288. https://doi.org/10.1071/PC18016

Rojahn J, Gleeson D, Furlan EM (2018) Monitoring post-release survival of the northern corroboree frog, *Pseudophryne pengilleyi*, using environmental DNA. Wildlife Research 45: 620–626. https://doi.org/10.1071/WR17179

Silla AJ, Byrne PG (2019) The Role of Reproductive Technologies in Amphibian Conservation Breeding Programs. Annual Review of Animal Biosciences 7: 499–519. https://doi.org/10.1146/annurev-animal-020518-115056

Silla AJ, McFadden M, Byrne PG (2018) Hormone-induced spawning of the critically endangered northern corroboree frog *Pseudophryne pengilleyi*. Reproduction, Fertility and Development 30: 1352–1358. https://doi.org/10.1071/RD18011

Skerratt LF, Berger L, Clemann N, Hunter DA, Marantelli G, Newell DA, Philips A, McFadden M, Hines HB, Scheele BC, Brannelly LA, Speare R, Versteegen S, Cashins SD, West M (2016) Priorities for management of chytridiomycosis in Australia: saving frogs from extinction. Wildlife Research 43(2): 105–120. https://doi.org/10.1071/WR15071

Speare R, Smith JR (1992) An iridovirus-like agent isolated from the ornate burrowing frog *Limnodynastes ornatus* in northern Australia. Diseases of Aquatic Organisms 14: 51–57. https://doi.org/10.3354/dao014051

Strauß S, Ziegler T, Allmeling C, Reimers K, Frank-Klein N, Seuntjens R, Vogt PM (2011) In vitro culture of skin cells from biopsies from the Critically Endangered Chinese giant salamander, *Andrias davidianus* (Blanchard, 1871) (Amphibia, Caudata, Cryptobranchidae). Amphibian & Reptile Conservation 5: 51–63. https://biostor.org/reference/170463

Wells RW, Wellington CR (1985) A classification of the Amphibia and Reptilia of Australia. Australian Journal of Herpetology, Supplemental Series 1: 1–61.

White AW (1993) Ecological and behavioural observations on populations of the toadlets *Pseudophryne coriacea* and *Pseudophryne bibronii* on the Central Coast of New South Wales. In: Lunney D, Ayers D (Eds) Herpetology in Australia: A Diverse Discipline, 139–150. https://doi.org/10.7882/RZSNSW.1993.021

Zimkus BM, Hassapakis CL, Houck ML (2018) Integrating current methods for the preservation of amphibian genetic resources and viable tissues to achieve best practices for species conservation. Amphibian and Reptile Conservation 12: 1–27 [(e165)]. http://hdl.handle.net/20.500.12634/150

ORCID

Richard Mollard https://orcid.org/0000-0001-5820-7491

SHORT COMMUNICATION

Metaphase chromosomes of five Neotropical species of the genus *Drosophila* (Diptera, Drosophilidae)

Doris Vela[1], Erika Villavicencio[1]

I *Pontificia Universidad Católica del Ecuador, Facultad de Ciencias Exactas y Naturales, Laboratorio de Genética Evolutiva. Av. 12 de Octubre 1076 y Roca, Quito, Ecuador*

Corresponding author: Doris Vela (dvela508@puce.edu.ec)

Citation: Vela D, Villavicencio E (2023) Metaphase chromosomes of five Neotropical species of the genus *Drosophila* (Diptera, Drosophilidae). Comparative Cytogenetics 17: 273–281. https://doi.org/10.3897/compcytogen.17.108265

Abstract

The mitotic metaphases of five Andean species of genus *Drosophila* are described for the first time. The evolutionary and interspecific genetic relationships within three Neotropical *Drosophila* species groups are analyzed. The diploid chromosome number for each species is as follows: *D. cashapamba* Céspedes et Rafael, 2012 2n = 6 (2V, 1J) (X = J, Y = R), *D. ecuatoriana* Vela et Rafael, 2004 2n = 10 (3R, 2V) (X = V, Y = R), *D. ninarumi* Vela et Rafael, 2005 2n = 10 (3R, 1V, 1D) (X = V, Y = R), *D. urcu* Vela et Rafael, 2005 2n = 12 (4R, 2V) (X = V, Y = R), *D. valenteae* Llangarí-Arizo et Rafael, 2018 2n = 8 (3R, 1J) (X = J, Y = R).

Keywords

Andean, *Drosophila* chromosomes, *guarani*, *mesophragmatica*, metaphase, *tripunctata*

Introduction

The ancestral karyotype for the genus *Drosophila* Fallén, 1823 (Diptera, Drosophilidae) consists of five pairs of large chromosomes (V shape or J shape) and one pair of dots (Sturtevant and Novitski 1941). This *Drosophila* metaphase chromosome configuration has been commonly observed, for instance, in some species of Neotropical groups of the type subgenus *Drosophila*: *D. guarani* group (King 1947), *D. mesophragmatica* group (Brncic and Koref 1957; Hunter and Hunter 1964), *D. repleta* group (Wasserman 1960) and *D. tripunctata* group (Pipkin and Heed 1964). The species

of the type subgenus present a chromosome configuration ranging from three to six pairs of chromosomes. Cytogenetics studies demonstrated that in the genus *Drosophila* karyotypes of species may differ from the ancestral karyotype by the number of chromosomes and the chromosomal configuration, but chromosomal rearrangements do not break the integrity of Muller elements (chromosome arms and associated linkage groups) (Schaeffer 2018).

By means of the karyotypes, it is possible to observe the chromosomal rearrangements (inversions, translocations, duplications etc.) in species, and how they can limit the genetic exchange and potentially drive speciation (Noor et al. 2001). In addition, it is possible to detect interspecific and intraspecific polymorphism in species of *Drosophila* (Deng et al. 2007). Therefore, karyotypes are an important tool for understanding the evolutionary history of the *Drosophila* species, to conduct comparative genomics studies and to allow genome assembly at the chromosome level (Schaeffer 2018).

Most of the available cytological data about Neotropical species of *Drosophila* were reported in the past century (Metz and Moses 1923; Patterson and Wheeler 1942; Wharton 1943; Burla et al. 1949; Clayton and Wasserman 1957; Clayton and Wheeler 1975). In the most recent cytological studies of Neotropical species of *Drosophila* karyotypes of ten species from four sibling species groups have been described: *D. chorlavi* Céspedes et Rafael, 2012, *D. mesophragmatica* Duda, 1927 and *D. rucux* Céspedes et Rafael, 2012 from the *D. mesophragmatica* group (Mafla 2012), *D. butantan* Ratcov, Vilela et Goñi, 2017, *D. sachapuyu* Peñafiel-Vinueza et Rafael, 2018, and *D. zamorana* Peñafiel-Vinueza et Rafael, 2018 from the *D. guarani* group (Ratcov et al. 2017; Vela and Villavicencio 2021), *D. huancavilcae* Rafael et Arcos, 1989, *D. inca* Dobzhansky et Pavan, 1943, and *D. yangana* Rafael et Vela,;2003 from the *D. repleta* group (Mafla 2005, 2008), and *D. montevidensis* Goñi et Vilela, 2016 from the *D. tripunctata* group (Goñi and Vilela 2016).

In this study, the karyotypes of five Andean species of *Drosophila* from three sibling species groups are described for the first time: *D. ecuatoriana* Vela et Rafael, 2004 and *D. valenteae* Llangarí-Arizo et Rafael, 2018 from the *D. guarani* group, *D. cashapamba* Céspedes et Rafael, 2012 from the *D. mesophragmatica* group, *D. ninarumi* Vela et Rafael, 2005 and *D. urcu* Vela et Rafael, 2005 from the *D. tripunctata* group.

Methods

Species stock

The species analysed correspond to natural populations of: *D. cashapamba* (QCAZ-I 2349), Sangolquí Canton (location 0°19'59.3"S, 78°25'51"W DMS); *D. ecuatoriana* (QCAZ-I 1609), Yanacocha Forest (location 0°7'3.8"S, 78°35'9.4"W DMS); *D. ninarumi* (QCAZ-I 1765), Cruz Loma Forest (location 0°11'22"S, 78°31'17.2"W DMS); *D. urcu* (QCAZ-I 1755), Cruz Loma Forest (location 0°11'22"S, 78°31'17.2"W DMS) and *D. valenteae* (QCAZ-I 3142), Sangolquí Canton (location 0°19'59.3"S, 78°25'51"W DMS).

All species were provided by the Evolutionary Genetics Laboratory of Pontificia Universidad Católica del Ecuador. The flies were maintained in banana culture medium supplemented with fresh fruit, in a temperate room at 17 °C, with a 12 h light/dark cycle.

Chromosome plates

The metaphase nuclei of cerebral ganglia were obtained from third-instar larvae (ten males, ten females) of each species. Chromosomal plates were prepared by the cell suspension method (Cardoso and Dutra 1979) and thermic shock (Holmquist 1975) and stained with Giemsa. Ten metaphase nuclei were observed for each sex and species. A Ziess Axioskop 2 plus – HAL 100 microscope and a Cannon PowerShot A640 camera (100× objectives lens and optovar 2×) were used to observe and take the pictures of the mitotic chromosome cells. The modal number was considered the chromosome number of each species.

Mitotic chromosome analysis

For each species, the total length (TL), relative length (RL) and centromeric index (CI) of the chromosomes were estimated using the Axio Vision 4.4. Standard deviation of relative length was analysed using the SPSS statistical package 26.0v (Table 1).

Results

The description of new karyotypes of *Drosophila* species is presented below:

The *Drosophila guarani* group

The karyotype of *D. ecuatoriana* is 2n = 10 (3R, 2V), comprising of four autosomes – a large V-shaped metacentric (pair 2) and three pairs of rod-shaped telocentric chromosomes (pairs 3, 4 and 5) – and the sexual pair (X = V, Y = R). The X chromosome is V-shaped metacentric and the Y chromosome is rod-shaped telocentric (Fig. 1A, B, Table 1).

The karyotype of *D. valenteae* is 2n = 8 (3R, 1J), comprising of three rod-shaped telocentric autosomes (pairs 2, 3 and 4), and the sexual pair (X = J, Y = R). The X chromosome is J-shaped submetacentric, and the Y chromosome is rod-shaped telocentric (Fig. 1C, D, Table 1).

The *Drosophila mesophragmatica* group

The karyotype of *D. cashapamba* is 2n = 6 (2V, 1J) comprising of two V-shaped metacentric autosomes (pairs 2 and 3) and the sexual pair (X = J, Y = R). The X chromosome is J-shaped submetacentric and the Y chromosome is rod-shaped telocentric (Fig. 1E, F, Table 1).

Table 1. Measurement of metaphase chromosomes of five Andean *Drosophila* species.

Species	Chromosome	TL (µm)	RL (%)	CI	SD (n = 10)	Morphology
D. ecuatoriana	X	2,49	24,22	0,47	0,27	metacentric
2n = 10	Y	1,85	17,99	0,05	0,03	telocentric
	2	1,65	16,05	0,49	0,12	metacentric
	3	1,54	14,98	0,06	0,19	telocentric
	4	1,42	13,81	0,07	0,21	telocentric
	5	1,33	12,93	0,08	0,16	telocentric
D. valenteae	X	2,09	27,42	0,37	0,23	submetacentric
2n = 8	Y	1,73	22,7	0,06	0,31	telocentric
	2	1,4	18,37	0,07	0,21	telocentric
	3	1,26	16,53	0,08	0,23	telocentric
	4	1,14	14,96	0,09	0,14	telocentric
D. cashapamba	X	2,88	26,2	0,38	0,12	submetacentric
2n = 6	Y	1,94	17,65	0,05	0,04	telocentric
	2	3,21	29,2	0,47	0,11	metacentric
	3	2,96	26,93	0,49	0,12	metacentric
D. ninarumi	X	1,71	27,49	0,46	0,25	metacentric
2n = 10	Y	1,59	25,56	0,06	0,04	telocentric
	2	1,12	18	0,09	0,26	telocentric
	3	0,95	15,27	0,11	0,18	telocentric
	4	0,83	13,34	0,12	0,2	telocentric
	5	0,02	0,32	0,05	0,01	dot
D. urcu	X	3,09	24,75	0,48	0,23	metacentric
2n = 12	Y	2,65	21,23	0,04	0,07	telocentric
	2	1,62	12,98	0,49	0,17	metacentric
	3	1,58	12,66	0,06	0,27	telocentric
	4	1,45	11,61	0,07	0,21	telocentric
	5	1,21	9,69	0,08	0,14	telocentric
	6	0,88	7,05	0,11	0,29	telocentric

TL: Total Length, RL: Relative Length, CI: Centromeric Index, SD: Standard deviation.

The *Drosophila tripunctata* group

The karyotype of *D. ninarumi* is 2n = 10 (3R, 1V, 1D), comprising of four autosomes – three rod-shaped telocentric (pairs 2, 3 and 4) and one pair of dot-shaped chromosomes (pair 5), and the sexual pair (X = V, Y = R). The X chromosome is V-shaped metacentric and the Y chromosome is rod-shaped telocentric (Fig. 1G, H, Table 1).

The karyotype of *D. urcu* is 2n = 12 (4R, 2V) comprising of five autosomes – a pair of V-shaped metacentric (pair 2) and four pairs of rod-shaped telocentric chromosomes (pairs 3, 4, 5 and 6) – and the sexual pair (X = V, Y = R). The X chromosome is V-shaped metacentric and the Y chromosome is rod-shaped telocentric (Fig. 1I, J, Table 1).
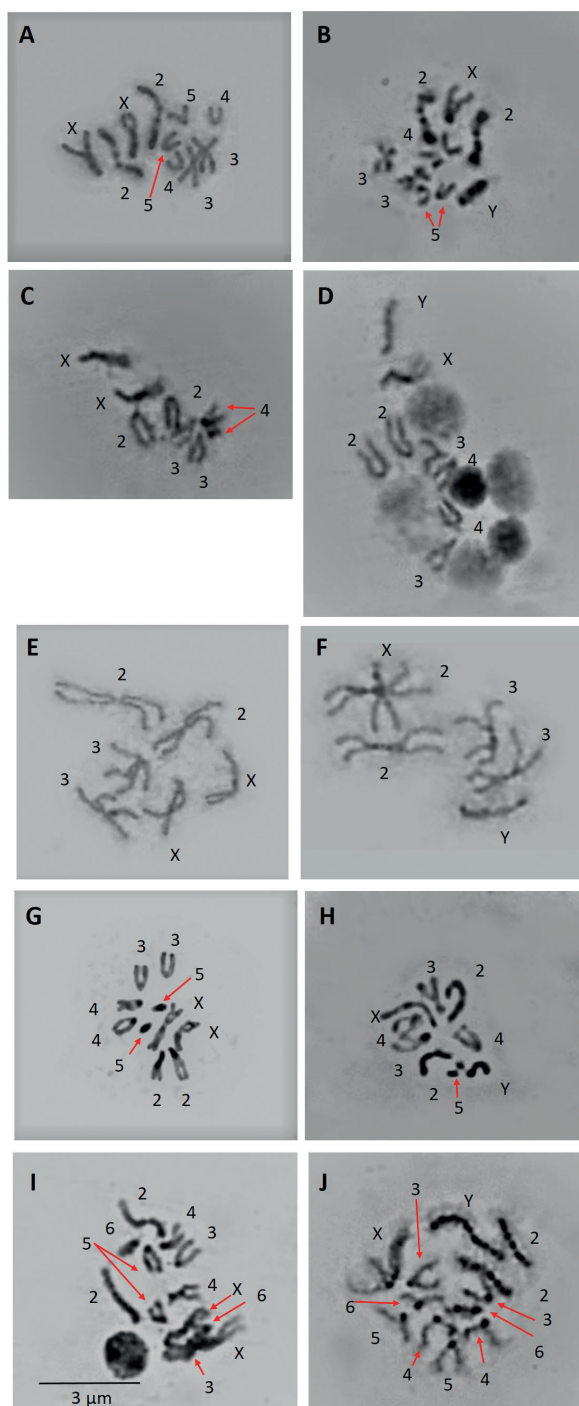
Figure 1. Metaphase karyotype of **A** *D. ecuatoriana* female **B** *D. ecuatoriana* male **C** *D. valenteae* female **D** *D. valenteae* male **E** *D. cashapamba* female **F** *D. cashapamba* male **G** *D. ninarumi* female **H** *D. ninarumi* male **I** *D. urcu* female **J** *D. urcu* male. Scale bar: 3 μm (**A–J**).

Discussion

Considering the high diversity of *Drosophila* species in the Neotropical region little is known about diploid chromosome numbers of these species.

In the *Drosophila guarani* group, the most common karyotype is 2n = 12. In the present study, the karyotype of *D. ecuatoriana* is 2n = 10 (Fig. 1A, B). A similar 2n = 10 karyotype was reported in other species of this group: *D. guaraja* King, 1947 (King 1947), *D. butantan* (Ratcov et al. 2017) and *D. sachapuyu* (Vela and Villavicencio 2021). The karyotype of *D. valenteae* is 2n = 8 (Fig. 1C, D) and is similar to *D. alexandrei* Cordeiro, 1951 (Cordeiro 1951), both species present the lowest diploid chromosome reported for the *Drosophila guarani* species group.

Several reports have shown that the karyotype of *Drosophila* species of the *D. mesophragmatica* group is highly conserved, 2n = 10, including a pair of rod-shaped or a dot-like fifth chromosomes (Brncic 1957). Additionally, paracentric inversions are the principal chromosomal rearrangements attributed to this species group (Brncic and Koref 1957). In our study, the chromosome number of *D. cashapamba* is 2n = 6, the chromosomes are large and present a small pericentromeric heterochromatin (Fig. 1E, F). It has been suggested that *D. cashapamba* is a junior synonym of *D. dreyfusi* Dobzhansky et Pavan, 1943 (Dr Carlos Vilela, pers. communication) due to the similarity of the male genitalia and the same chromosome number, 2n = 6 (Dobzhansky and Pavan 1943). However, in this study we maintain the current taxonomical classification until new taxonomic studies confirm the junior synonym status of *D. cashapamba*.

According to the information available in the *Drosophila* karyotype databases (Morelli et al. 2022), the chromosome number 2n = 6 is rarely reported in *Drosophila* subgenus. Only thirteen species of *Drosophila* subgenus present three pairs of chromosomes: *D. canalinea* Patterson et Mainland, 1944 from *D. canalinea* group, *D. dreyfusi* and *D. wingei* Cordeiro, 1964 from *D. dreyfusi* group, *D. albomicans* Duda, 1923, *D. annulipes* Duda, 1924, *D. neohypocausta* Lin et Wheeler, 1973 from *D. immigrans* group, *D. atalaia* Vilela et Sene, 1982 from *D. peruensis* group, *D. pinicola* Sturtevant, 1942 from *D. pinicola* group, *D. quinaria* Loew, 1866 from *D. quinaria* group; *D. neoguaramunu* Frydenberg, 1956 from *D. tripunctata* group, *D. montana* Patterson et Wheeler, 1942 from *D. virilis* group, *D. aracea* Heed et Wheeler, 1957 and *D. tranquilla* Spencer, 1942 (not grouped).

Most species of the *D. tripunctata* group have a karyotype 2n = 12, the sixth pair is a dot chromosome; some members of *D. tripunctata* group have a karyotype 2n = 10 (Morelli et al. 2022). In the karyotype of *D. ninarumi*, 2n = 10, it is present a dot-like fifth pair of chromosome (Fig. 1G, H) which is reported in the most species of *Drosophila tripunctata* group. This karyotype is similar to *D. fairchaldi* Pipkin et Heed, 1964 and *D. unipunctata* Patterson, 1943 (Wharton 1943; Pipkin and Heed 1964; Clayton and Wheeler 1975) but in these species the dot-like chromosome is absent. In the case of *D. urcu*, the karyotype is 2n = 12, all the chromosomes are large metacentric or telocentric (Fig. 1I, J). Our data show that the karyotype of *D. ninarumi* and *D. urcu* have a relevant similitud, the sexual chromosomes are the largest of the chromosome set, with a Y chromosome heteropycnotic (Fig. 1G, J).

Traditional studies like genetic crosses, in situ hybridization, polytene chromosomes maps or karyotype description are not commonly performed. However, for the genus *Drosophila*, the information provided by cytological studies is the initial tool in understanding the evolutionary history and the high radiation of the *Drosophila* species in the Neotropical region and also important in the beginning of genomic studies on these species.

Conclusions

This study reveals the first karyotype description of five Neotropical species of *Drosophila*. Only the karyotype of *D. urcu*, 2n = 12, is similar to the ancestral karyotype of *Drosophila*, but the sixth pair are large chromosomes. The karyotypes of *D. ecuatoriana* and *D. ninarumi* are 2n = 10, but only the last one has a dot-like chromosome. The karyotype of *D. valenteae* is 2n = 8; this is the second species of *D. guarani* group that have this chromosome number. The karyotype of *D. cashapamba* presents a low chromosome number, 2n = 6, which is only reported in other thirteen species of subgenus *Drosophila*.

Acknowledgements

References

Brncic D (1957) A comparative study of chromosomal variation in species of the *mesophragmatica* group of *Drosophila*. Genetics 42(6): 798–805. https://doi.org/10.1093/genetics/42.6.798

Brncic D, Koref S (1957) The *mesophragmatica* group of species of *Drosophila*. Evolution 11(3): 300–310. https://doi.org/10.2307/2405794

Burla H, Da Cunha AB, Cordeiro AR, Dobzhansky T, Malogolowkin C, Pavan C (1949) The *willistoni* group of sibling species of Drosophila. Evolution 3(4): 300–314. https://doi.org/10.2307/2405716

Cardoso H, Dutra A (1979) The Neo-X Neo-Y sex pair in Acrididae, its structure and association. Chromosoma 70(3): 323–336. https://doi.org/10.1007/BF00328770

Clayton FE, Wasserman M (1957) Chromosomal studies of several species of *Drosophila*. University of Texas Publications 5721: 125–131.

Clayton F, Wheeler M (1975) A catalog of *Drosophila* metaphase chromosome configurations. In: King R (Ed.) Handbook of Genetics (Vol. 3). Plenum Press. New York, 471–512. https://doi.org/10.1007/978-1-4615-7145-2_18

Cordeiro AR (1951) *Drosophila alexandrei*, una nova especie brasileira. Universidade do Rio Grande do Sul 3: 1–10.

Deng Q, Zeng Q, Qian Y, Li C, Yang Y (2007) Research on the karyotype and evolution of *Drosophila melanogaster* species group. Journal of Genetics and Genomics 34(3): 196–213. https://doi.org/10.1016/S1673-8527(07)60021-6

Dobzhansky T, Pavan C (1943) Studies on brazilian species of *Drosophila*. Boletim da Faculdade de Filosofia, Ciências e Letras da Universidade de São Paulo Biol Geral 36(4): 7–72. http://www.drosophila.jp/jdd/class/030703/03070372.pdf

Goñi B, Vilela CR (2016) Two new neotropical species of drosophilinae (Diptera: Drosophilidae) from Uruguay. Zoologia 33(6): 1–13. https://doi.org/10.1590/s1984-4689zool-20160142

Holmquist B (1975) A revision of the species *Archaeomysis grebnitzkii* Czernaivsky (sic) and *A. maculata* (Holmes) (Crustacea, Mysidacea). Ökologie und Geographie der Tiere 102: 51–71.

Hunter AS, Hunter RA (1964) The *mesophragmatica* species group of *Drosophila* in Colombia. Annals of the Entomological Society of America 57(6): 732–736. https://doi.org/10.1093/aesa/57.6.732

King JC (1947) A comparative analysis of the chromosomes of the *guarani* group of Drosophila. Evolution 1: 48–62. https://doi.org/10.2307/2405403

Mafla AB (2012) Cariología beta de tres especies pertenecientes al grupo de especies *Drosophila mesophragmatica*. Revista Ecuatoriana de Medicina y Ciencias Biológicas 33(1–2): 38–45. https://doi.org/10.26807/remcb.v33i1-2.221

Mafla AB (2005) Cariotipos metafásicos de *Drosophila inca* y *D. yangana*, subgrupo *inca*, grupo *repleta*. Revista Ecuatoriana de Medicina y Ciencias Biológicas 27: 21–25. https://doi.org/10.26807/remcb.v27i1-2.190

Mafla AB (2008) *Drosophila huancavilcae*: ciclo biológico y cariotipo metafásico. Revista Ecuatoriana de Medicina y Ciencias Biológicas 29(1–2): 7–10. https://doi.org/10.26807/remcb.v29i1-2.205

Metz CW, Moses MS (1923) Chromosome relationships and genetic behavior in the genus *Drosophila*: I. A comparison of the chromosomes of different species of *Drosophila*. Journal of Heredity 14(5): 195–205. https://doi.org/10.1093/oxfordjournals.jhered.a102315

Morelli MW, Blackmon H, Hjelmen CE (2022) Diptera and *Drosophila* karyotype databases: A useful dataset to guide evolutionary and genomic studies. Frontiers in Ecology and Evolution 10: 832378. https://doi.org/10.3389/fevo.2022.832378

Noor MAF, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. Proceedings of the National Academy of Sciences 98(21): 12084–12088. https://doi.org/10.1073/pnas.221274498

Patterson J, Wheeler MR (1942) Description of new species of the subgenera *Hirtodrosophila* and *Drosophila*. University of Texas Publications 4213: 67–109.

Pipkin SB, Heed WB (1964) Nine new members of the *Drosophila tripunctata* species group (Diptera: Drosophilidae). Pacific Insects 6(2): 256–273. http://www.drosophila.jp/jdd/class/030703/03070394.pdf

Ratcov V, Vilela CR, Goñi B (2017) A new species of neotropical *Drosophila* (Diptera, Drosophilidae) belonging to the *guarani* group. Revista Brasileira de Entomologia 61(3): 232–238. https://doi.org/10.1016/j.rbe.2017.06.002

Schaeffer SW (2018) Muller "elements" in *Drosophila*: How the search for the genetic basis for speciation led to the birth of comparative genomics. Genetics 210(1): 3–13. https://doi.org/10.1534/genetics.118.301084

Sturtevant AH, Novitski E (1941) The homologies of the chromosome elements in the genus *Drosophila*. Genetics 26(5): 517–541. https://doi.org/10.1093/genetics/26.5.517

Vela D, Villavicencio E (2021) Karyotype description of two andean species of the *guarani* group of *Drosophila* (Díptera: Drosophilidae) and Cytological Notes. Journal of Insect Science 21(3). https://doi.org/10.1093/jisesa/ieab032

Wasserman M (1960) Cytological and phylogenetic relationships in the *repleta* group of the genus *Drosophila*. Genetics 46: 842–859. https://doi.org/10.1073/pnas.46.6.842

Wharton LT (1943) An analysis of the metaphase and salivary chromosome morphology within the genus *Drosophila*. University of Texas Publications 4313: 282–319.

ORCID

Doris Vela https://orcid.org/0000-0001-7690-7758

SHORT COMMUNICATION

COMPARATIVE
Cytogenetics
A peer-reviewed open-access journal
International Journal of Plant & Animal Cytogenetics,
Karyosystematics, and Molecular Systematics

Karyotype and reproductive traits of the unique symbiotic mealybug *Orbuspedum machinator* G.-Z. (Homoptera, Coccinea)

Ilya A. Gavrilov-Zimin[1]

I *Zoological Institute, Russian Academy of Sciences, Universitetskaya nab. 1, St. Petersburg, 199034, Russia*

Corresponding author: Ilya A. Gavrilov-Zimin (coccids@gmail.com)

https://zoobank.org/D8C1606E-D00D-49E9-9641-0B86785DBC26

Citation: Gavrilov-Zimin IA (2023) Karyotype and reproductive traits of the unique symbiotic mealybug *Orbuspedum machinator* G.-Z. (Homoptera, Coccinea). Comparative Cytogenetics 17: 283–286. https://doi.org/10.3897/compcytogen.17.116550

Abstract

The karyotype and reproductive features of *Orbuspedum machinator* Gavrilov-Zimin, 2017 (Pseudococcidae) were studied for the first time. Diploid chromosome number is 18 in females. Reproduction is probably bisexual, as indicated by the presence of characteristic Lecanoid heterochromatinization of the paternal set of chromosomes in embryonic cells of about 50% of the embryos studied. The female reproductive system has a pair of lateral oviducts merged into enlarged common oviduct; the spermatheca and accessory glands are connected to the common oviduct in its proximal part. Complete ovoviviparity occurs in ontogenesis.

Keywords

Chromosomes, cytogenetics, Lecanoid genetic system, ovoviviparity, scale insects

Females of the peculiar legless mealybug *Orbuspedum machinator* Gavrilov-Zimin, 2017 from the monotypic genus *Orbuspedum* Gavrilov-Zimin, 2017 live inside conical domiciles constructed of densely packed fungal hyphae of the sooty mold *Capnodium* sp. mixed with wax secreted by the mealybug (Fig. 1a–c). The domicile grows together with the insect, which irrigates the hyphae with honeydew. This unique animal/fungus mutualistic symbiosis was described by me in details earlier (Gavrilov-Zimin 2017) from tropical rainforests of the Malay Peninsula (southern Thailand).
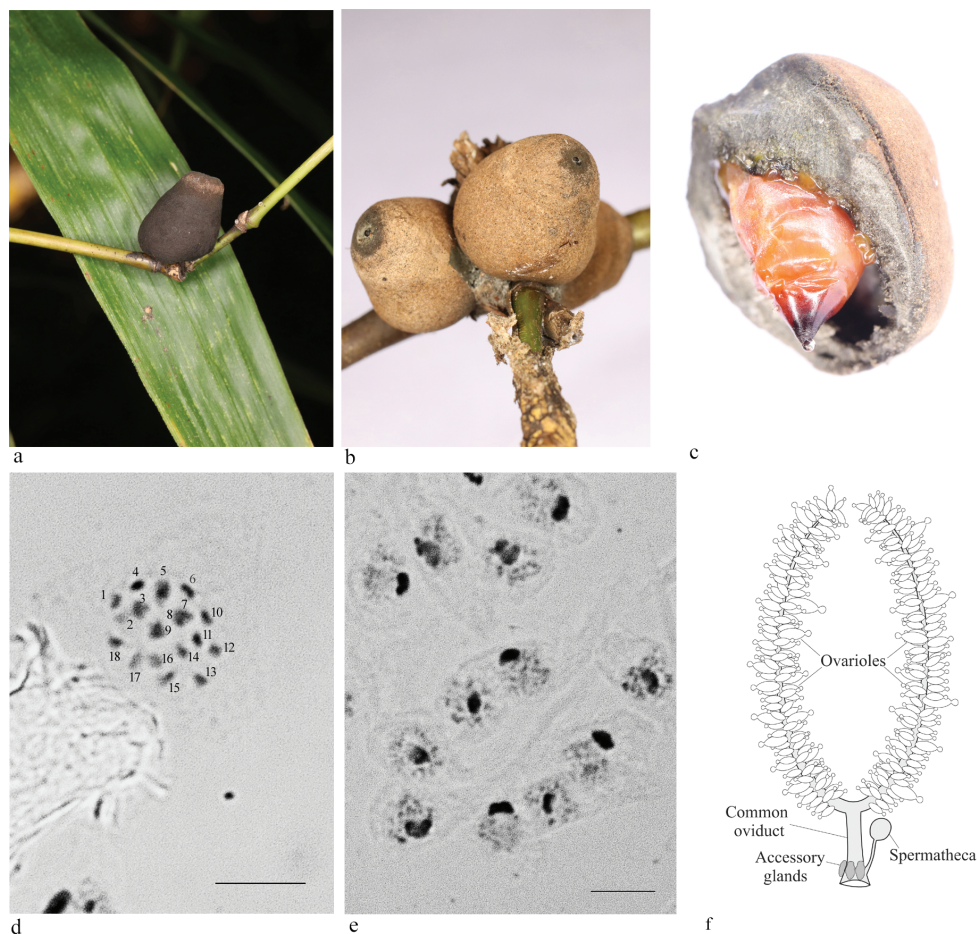
Figure 1. *Orbuspedum machinator*, Thailand, Khao-Sok **a** mature adult female inside a fungal domicile on twig of bamboo **b** younger females in three domiciles, **c** adult female inside a broken domicile (**a–c** photos by A.S. Kurochkin) **d** metaphase chromosomal plate in a cell of the female embryo, 2n = 18 **e** male embryo cells with heterochromatinized paternal chromosomes (deeply stained bodies) **f** scheme of the female reproductive system. Scale bar: 10 μm.

Such mutualistic symbiosis has never been reported for any other scale insect or for any other animals known to the author. In November 2023, I was able to revisit the type locality of *O. machinator* and collect gravid females for cytogenetic and reproductive studies. The karyotype of the species includes 18 chromosomes, quite similar in length (Fig. 1d). Such diploid number has not been previously reported for any member of the informal group "legless mealybugs", including at least 26 nominal genera in the world fauna (Gavrilov-Zimin 2017); the other studied species have 2n = 10, 12, 16, 20, 22+ Bs, 24, 24 + Bs, or 30 (Nur et al. 1987; Gavrilov-Zimin 2016, 2020).

About 50% of the embryos studied contained cells with characteristic Lecanoid heterochromatinization (Fig. 1e) of the paternal chromosomes set (see Nur 1980;

Gavrilov-Zimin et al. 2015 for more details). Usually in the Lecanoid system, the heterochromatic chromosome set exists in all stages of the male ontogenesis. In male meiosis, the chromosomes do not pair and separate equationally during the first division. Then, in the second division, two metaphase plates are formed, and the heterochromatic and euchromatic chromosomes segregate to the opposite poles. As a result of meiosis, quadrinucleate spermatids are formed, but only the nuclei of maternal origin produce sperm (Hughes-Schrader 1948; Nur 1980; Gavrilov-Zimin et al. 2015). Such heterochromatinization in *O. machinator* obviously indicates bisexual reproduction in the studied population. However, adult males or male larvae have not been found. This discrepancy can be explained by the probable separate life of minute males and larger females in different parts of the host plant (or even on different plants), which is a common feature of scale insects (Borchsenius 1963). Anatomical studies of the available females showed that their reproductive system is similar to that of other legless mealybugs studied (Gavrilov-Zimin 2020) and includes a pair of lateral oviducts merging into an enlarged common oviduct; the spermatheca and accessory glands are connected to the common oviduct in its proximal part (Fig. 1f). All embryonic development occurs within the ovarioles and oviducts (complete ovoviviparity). The hatched primolarvae leave the maternal fungal domicile through the apical orifice.

Acknowledgements

References

Borchsenius NS (1963) Practical guide to the determination of scale insects of cultivated plants and forest trees of the USSR. Leningrad, 311 pp. [In Russian]

Gavrilov-Zimin IA, Stekolshchikov AV, Gautam DC (2015) General trends of chromosomal evolution in Aphidococca (Insecta, Homoptera, Aphidinea + Coccinea). Comparative Cytogenetics 9(3): 335–422. https://doi.org/10.3897/CompCytogen.v9i3.4930

Gavrilov-Zimin IA (2016) Cytogenetic and taxonomic studies of some legless mealybugs (Homoptera: Coccinea: Pseudococcidae). Comparative Cytogenetics 10(4): 587–601. https://doi.org/10.3897/compcytogen.v10i4.10503

Gavrilov-Zimin IA (2017) A remarkable example of symbiosis between an animal and a fungus in a new species of legless mealybug (Insecta: Pseudococcidae). Journal of Natural History 51(37–38): 2211–2224. https://doi.org/10.1080/00222933.2017.1365180

Gavrilov-Zimin IA (2020) Chromosomal and reproductive features of some Oriental and Australasian scale insects (Homoptera, Coccinea). Comparative Cytogenetics 14(3): 339–352. https://doi.org/10.3897/CompCytogen.v14i3.53367

Hughes-Schrader S (1948) Cytology of coccids (Coccoidea-Homoptera). Advances in Genetics 2: 127–203. https://doi.org/10.1016/S0065-2660(08)60468-X

Nur U (1980) Evolution of unusual chromosome systems in scale insects (Coccoidea: Homoptera). In: Blackman RL, Hewitt GM & Ashburner M (Eds) Insect Cytogenetics. London, 97–117.

Nur U, Brown SW, Beardsley JW (1987) Evolution of chromosome number in mealybugs (Pseudococcidae: Homoptera). Genetica 74: 53–60. https://doi.org/10.1007/BF00055094

ORCID

Ilya A. Gavrilov-Zimin https://orcid.org/0000-0003-1993-5984

Karyotype diversity in the genus *Nysius* Dallas, 1852 (Hemiptera, Heteroptera, Lygaeidae) is much greater than you might think

Natalia V. Golub[1], Boris A. Anokhin[1], Valentina G. Kuznetsova[1]

1 *Department of Karyosystematics, Zoological Institute, Russian Academy of Sciences, Universitetskaya emb. 1, 199034 St. Petersburg, Russia*

Corresponding authors: Natalia V. Golub (nvgolub@mail.ru); Valentina G. Kuznetsova (valentina_kuznetsova@yahoo.com)

Citation: Golub NV, Anokhin BA, Kuznetsova VG (2023) Karyotype diversity in the genus *Nysius* Dallas, 1852 (Hemiptera, Heteroptera, Lygaeidae) is much greater than you might think. Comparative Cytogenetics 17: 287–293. https://doi.org/10.3897/compcytogen.17.116628

Abstract

We studied the karyotype and chromosomal distribution of 18S rDNA clustered in nucleolar organizer regions (NORs) in *Nysius graminicola* (Kolenati, 1845), belonging to the subfamily Orsillinae (Lygaeidae). It is shown that this species has a karyotype with 2n = 22(18+mm+XY), previously known in only one of 24 studied species of the genus *Nysius* Dallas, 1852, characterized by a similar karyotype, 2n = 14(12+mm+XY). In *N. graminicola*, 18S loci are located on sex chromosomes, which is a previously unknown trait for this genus. Our results in a compilation with previous data revealed dynamic evolution of rDNA distribution in *Nysius*. It is concluded that molecular chromosomal markers detected by FISH contribute to a better understanding of the structure and evolution of the taxonomically complex genus *Nysius*.

Keywords

18S rDNA, Ag-NOR, chromosome number, FISH, *Nysius graminicola*, Orsillinae, sex chromosomes, true bugs

Introduction

Nysius Dallas, 1852 is one of the most common and widely distributed genera within the family Lygaeidae (Heteroptera, Pentatomomorpha). Species of the genus are seed-predators; most species live in ruderal habitats and are often extremely abundant and

sometimes becoming agricultural pests (Ge and Li 2019). The genus currently includes more than 100 described species and subspecies, with many more species remaining unrecognized (Ashlock 1967; Schaefer and Panizzi 2000; Péricart 2001; Nakatani 2015; Dellapé and Henry 2023). *Nysius* is a taxonomically complex group, and its members are known as "difficult to identify" because of the striking similarity of morphological features (Nakatani 2015). Obviously, some new methods and approaches are needed to solve the problem of distinguishing between closely related *Nysius* species. It has been shown that DNA sequencing of a standard gene region or "DNA barcoding" might speed a solution (Matsuura et al. 2012; Nakatani 2015).

Quite a few species of *Nysius* have been studied cytogenetically. Data on the number of chromosomes, the mechanism of sex chromosomes and, in some cases, the peculiarities of meiosis are currently available for 24 species, i.e. about 25% of all known species of this genus (reviewed by Ueshima and Ashlock 1980; see also Golub et al. 2023). Routine cytogenetics of *Nysius* appears to be highly conserved: all species have 2n = 14(12+XY), with the only exception being *N. tennellus* Barber, 1947, which has 2n = 22(20+XY). Each species has a pair of very small, so-called m-chromosomes (microchromosomes).

Consistent advances in chromosomal analysis increased dramatically in recent decades, becoming more refined and accurate through molecular cytogenetics using fluorescence *in situ* hybridization (FISH) allowing physical location of DNA sequences in chromosomes. The chromosomes of true bugs are holokinetic (Ueshima 1979), that is, they lack centromeres; therefore, the search for chromosomal markers is of great importance for the comparative analysis of their karyotypes. *rRNA* genes are among the better-known multigene families in true bugs (Panzera et al. 2021; Kuznetsova et al. 2021). The first recent application of FISH to map *rRNA* genes on the chromosomes of two *Nysius* species with modal karyotypes of 2n = 14(12+XY), *N. cymoides* (Spinola, 1837) and *N. helveticus* (Herrich-Schäffer, 1850), showed that they both have rDNA sites on the largest pair of autosomes (Golub et al. 2023).

The present study is focused on karyotype description of *N. graminicola* (Kolenati, 1845) based on classical cytogenetics, including Ag-NOR staining, and FISH mapping of the 18S rDNA probe, which, we believe, opens up new perspectives for understanding the evolution of karyotypes in the genus *Nysius*.

Material and methods

Five males of *Nysius graminicola* were collected on August 15, 2023, 20 km NE of Voronezh (Russia) in a flood meadow on cereals. Males were freshly fixed in a mixture of alcohol and acetic acid (3:1) and stored in a refrigerator at 4 degrees until examination. Several slides were prepared from the testes of each male. Standard karyotypes were studied after staining by the Schiff–Giemsa method (Grozeva and Nokkala 1996). Nucleolus organizer regions (NORs) were localized by Ag-staining according to Howell and Black (1980) with minor modifications as described in Karagyan et al. (2020). To study the chromosomal distribution of major rDNA, FISH with an 18S

rDNA probe of the firebug *Pyrrhocoris apterus* (Linneus, 1758) was performed according to the protocol described by Grozeva et al. (2015). The entire procedure (labeling, hybridizing, and detecting) is described in Golub et al. (2019) and Gokhman and Kuznetsova (2022). All preparations were photographed under oil-immersion (X100 objective) using a Leica DM 6000 B microscope, Leica DFC 345 FX camera, and Leica Application Suite 3.7 software with Image Overlay module (Leica Microsystems, Wetzlar, Germany). Filter sets A and L5 (Leica Microsystems) were used. The specimens from which chromosome preparations were made and the preparations themselves are stored at the Zoological Institute RAS (St. Petersburg, Russia).

Results

Nysius graminicola (Kolenati, 1845) n = 11 (9AA+mm+XY), 2n = 22, XY

The karyotype of *N. graminicola* has been studied for the first time. We analyzed the stages of male meiosis from prophase and metaphase I (MI) to metaphase II (MII) after the classic routine staining (Fig. 1a–d), after FISH with an 18S rDNA probe (Fig. 1e, f), and after Ag-staining (Fig. 1g). At the early prophase stages (Fig. 1a, e, g), there are two heteropycnotic bodies corresponding to the X-chromosome (presumably larger) and Y- chromosome (smaller); both lie on the periphery of the nucleus, sometimes far apart (Fig. 1e, g), but sometimes quite close to one another (Fig. 1a). At MI (Fig. 1b, c) and diakinesis/MI transition (Fig. 1f), there are 10 bivalents of autosomes, including a small pair of m-chromosomes, and sex chromosomes X and Y placed separately from each other. Eleven elements, including ten autosomes split into chromatids and a pseudobivalent XY, were found in each of the sister MII nuclei (Fig. 1d). It is obvious that sex chromosomes, unlike autosomes and m-chromosomes, segregate equationally in the first round of meiosis and divide reductionally in the second round of meiosis (inverted or post-reductional meiosis), which is characteristic of all Pentatomomorpha and most Heteroptera in general (Ueshima 1979). The meioformula of the karyotype of *N. graminicola* can thus be denoted as n = 9AA+mm+X+Y (2n = 22, XY). The autosomes form a decreasing size series; sex chromosomes, as noted above, are a different size and behave like univalents, each splitting into chromatids. M-chromosomes exhibit negative heteropycnosis during meiotic divisions; they may be located separately or form a pseudobivalent at prophase (not shown) and at MI (Fig. 1b, c), a phenomenon known as "touch-and-go" pairing studied in depth by Nokkala (1986) on the example of *Coreus marginatus* (Linnaeus, 1758) (Coreidae). Both MI and MII plates are radial, with sex chromosomes and m- chromosomes lying in the center of a ring formed by bivalents (Fig. 1b, c, d). rDNA signals are visible on both sex chromosomes at all stages of meiosis, with larger and brighter signals on the Y-chromosome (Fig. 1e, f). Ag-staining revealed remnants of the nucleoli associated with both sex chromosomes in interphase/prophase cells, confirming the presence of *rRNA* genes in these chromosomes (Fig. 1g).
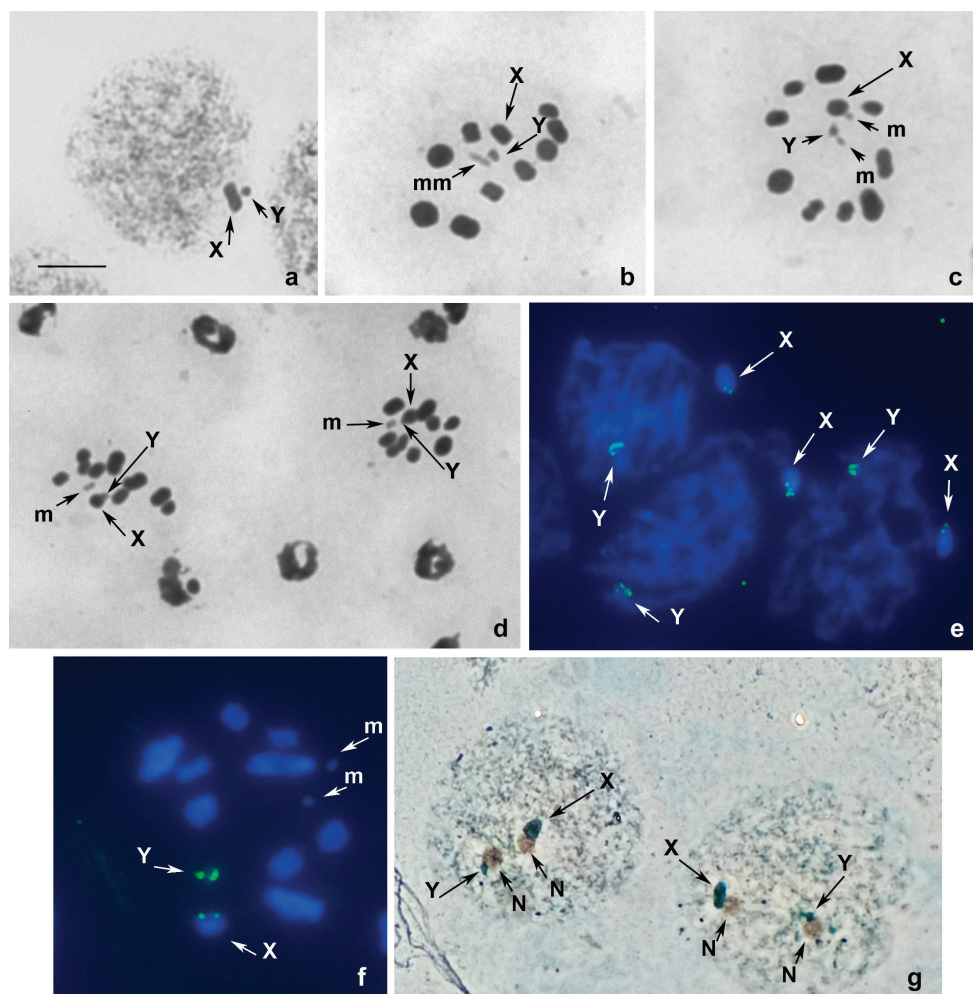
Figure 1. a–g Male meiotic karyotype of *N. graminicola* after standard staining (**a–d**), FISH with 18S rDNA probe (**e, f**), and Ag-staining (**g**) **a, e, g** interphase/prophase nuclei **b, c** metaphase I **f** diakinesis/ MI transition **d** metaphases II, daughter cells. N – nucleolus. Scale bar: 10 μm.

Discussion

Nysius graminicola is the second species in the genus *Nysius* to have 2n = 22(20+XY). This karyotype was previously known only in *N. tennellus*, and its origin was attributed to autosome fragmentations in the karyotype with 2n = 14(12+XY), representing a plesiomorphic state common to vast majority of *Nysius* species (Ueshima and Ashlock 1980). It should be noted that this karyotype is one of two (second 2n = 16, XY) modal karyotypes in the family Lygaeidae including the subfamily Orsillinae (Ueshima and Ashlock 1980; Papeschi and Bressa 2006). The above hypothesis is confirmed by the fact that in the karyotype with 2n = 14 there is a pair of very large chromosomes

(although for many species no karyotype illustration is given in the original publications), whereas in the karyotype with 2n = 22 (in both *N. graminicola* and *N. tennellus*) there is no such pair, and the chromosomes form a decreasing size series. The detection of a ribosomal cluster in autosomes in *N. cymoides* and *N. helveticus* sharing a modal karyotype (Golub et al. 2023) suggests an autosomal rDNA pattern to be the ancestral state for *Nysius*. Because the 18S ribosomal genes in these species are located on the largest pair of autosomes, we hypothesized that they would be found in one of the autosome pairs in *N. graminicola* with a derived karyotype. However, this hypothesis was not confirmed in our results, since the hybridization marks of the 18S rDNA probe were detected in the sex chromosomes of this species. Such a relocation of ribosomal sites from autosomes to the sex chromosomes is unlikely to be the result of chromosomal rearrangements alone. It is conceivable that transposable elements (also called "jumping genes" or mobile genetic elements) capable capturing entire genes and moving them from one genomic locus to another (Fambrini et al. 2020), could be involved in the dispersal of *rRNA* genes in the genus *Nysius*, as suggested for some other true bugs and some other insects (see examples and references in Panzera et al. 2021). The movement of rDNA clusters from autosomes to sex chromosomes is thought to be of evolutionary significance, causing genetic differentiation between divergent lineages and speciation events (see Pita et al. 2016; Panzera et al. 2021). We hypothesize that studies of other *Nysius* species will reveal a greater diversity of rDNA cluster distribution patterns, contributing to a better understanding of the structure and evolution of this taxonomically complex genus.

Conclusion

Our results show that the genus *Nysius* is characterized by a much more pronounced karyotype diversity than previously thought.

Acknowledgments

References

Ashlock PD (1967) A general classification of the Orsillinae of the world (Hemiptera-Heteroptera: Lygaeidae). University of the California Publications in Entomology 48: 1–82.

Dellapé PM, Thomas JH (2023) Lygaeoidea Species File. Version 5.0/5.0. http://lygaeoidea.speciesfile.org/ [accessed on 04 December 2023]

Fambrini M, Usai G, Vangelisti A, Mascagni F, Pugliesi C (2020) The plastic genome: The impact of transposable elements on gene functionality and genomic structural variations. Genesis 58: e23399. https://doi.org/10.1002/dvg.23399

Ge X, Li G (2019) Review of the genus *Nysius* Dallas from Mongolian Plateau (Hemiptera: Heteroptera: Orsillidae). Zootaxa 4560(1): 171–183. https://doi.org/10.11646/zootaxa.4560.1.10

Gokhman VE, Kuznetsova VG (2022) FISH – In Insect Chromosomes. In: Liehr T (Ed.) Cytogenetics and Molecular Cytogenetics. Boca Raton (USA), 319–338. https://doi.org/10.1201/9781003223658-27

Golub N, Anokhin B, Kuznetsova V (2019) Comparative FISH mapping of ribosomal DNA clusters and TTAGG telomeric sequences to holokinetic chromosomes of eight species of the insect order Psocoptera. Comparative Cytogenetics 13: 403–410. https://doi.org/10.3897/CompCytogen.v13i4.48891

Golub NV, Maryańska-Nadachowska A, Anokhin BA, Kuznetsova VG (2023) Expanding the chromosomal evolution understanding of lygaeoid true bugs (Lygaeoidea, Pentatomomorpha, Heteroptera) by classical and molecular cytogenetic analysis. Genes 14: 725–738. https://doi.org/10.3390/genes14030725

Grozeva S, Nokkala S (1996) Chromosomes and their meiotic behavior in two families of the primitive infraorder Dipsocoromorpha (Heteroptera). Hereditas 125: 189–192. https://doi.org/10.1111/j.1601-5223.1996.t01-1-00031.x

Grozeva S, Anokhin B, Kuznetsova VG (2015) Bed bugs (Hemiptera). In: Sharachov I (Ed.) Protocols for Cytogenetic Mapping of Arthropod Genomes. Boca Raton (USA), 285–326. https://doi.org/10.1201/b17450-9

Howell WM, Black DA (1980) Controlled silver-staining of nucleolus organizer regions with a protective colloidal developer: a 1-srep method. Experientia 36: 1014–1015. https://doi.org/10.1007/BF01953855

Karagyan G, Golub N, Sota T (2020) Cytogenetic characterization of periodical cicadas (Hemiptera: Cicadidae: *Magicicada*). European Journal of Entomology 117: 474–480. https://doi.org/10.14411/eje.2020.050

Kuznetsova VG, Gavrilov-Zimin IA, Grozeva SM, Golub NV (2021) Comparative analysis of chromosome numbers and sex chromosome systems in Paraneoptera (Insecta). Comparative Cytogenetics 15(3): 279–327. https://doi.org/10.3897/CompCytogen.v15.i3.71866

Matsuura Y, Kikuchi Y, Hosokawa T, Koga R, Meng X-Y, Kamagata Y, Nikoh N, Fukatsu T (2012) Evolution of symbiotic organs and endosymbionts in lygaeid stinkbugs. The ISME Journal 6: 397–409. https://doi.org/10.1038/ismej.2011.103

Nakatani Yu (2015) Revision of the lygaeid genus *Nysius* (Heteroptera: Lygaeidae: Orsillinae) of Japan, with description of a new species. Entomological Science 18(4): 435–441. https://doi.org/10.1111/ens.12141

Nokkala S (1986) The mechanisms behind the regular segregation of the m-chromosomes in *Coreus marginatus* L. (Coreidae, Hemiptera). Hereditas 105: 73–85. https://doi.org/10.1111/j.1601-5223.1986.tb00645.x

Panzera F, Pita S, Lorite P (2021) Chromosome structure and evolution of Triatominae: A review. In: Guarneri A, Lorenzo M (Eds) Triatominae – the biology of Chagas disease

vectors. Entomology in Focus. New York (USA), 65–99. https://doi.org/10.1007/978-3-030-64548-9_4

Papeschi AG, Bressa MJ (2006) Evolutionary cytogenetics in Heteroptera. Journal of Biological Research 5: 3–21.

Péricart J (2001) Family Lygaeidae. In: Aukema B, Rieger C (Eds) Catalogue of the Heteroptera of the Palaearctic region. 4. Amsterdam, 35–220.

Pita S, Lorite P, Nattero J, Galvão C, Alevi KCC, Teves SC, Azeredo-Oliveira MTV, Panzera F (2016) New arrangements on several species subcomplexes of *Triatoma* genus based on the chromosomal position of ribosomal genes (Hemiptera-Triatominae). Infection, Genetics and Evolution 43: 225–231. https://doi.org/10.1016/j.meegid.2016.05.028

Schaefer CW, Panizzi AR (2000) Economic importance of Heteroptera: A general view. In: Schaefer CW, Panizzi AR (Eds) Heteroptera of Economic Importance. Boca Raton (USA) 3–8. https://doi.org/10.1201/9781420041859.ch1

Ueshima N (1979) Animal Cytogenetics. Insecta 6. Hemiptera II: Heteroptera. Berlin, 118 pp.

Ueshima N, Ashlock PD (1980) Cytotaxonomy of the Lygaeidae (Hemiptera-Heteroptera). University of Kansas science bulletin 51(26): 717–801. https://doi.org/10.5962/bhl.part.3259

ORCID

Natalia V. Golub https://orcid.org/0000-0002-6048-9253
Boris A. Anokhin https://orcid.org/0000-0002-4110-6704
Valentina G. Kuznetsova https://orcid.org/0000-0001-8386-5453

RESEARCH ARTICLE

An updated Atlas of *Helophorus* chromosomes

Robert B. Angus[1]

1 *Department of Life Sciences (Insects), The Natural History Museum, London SW7 5 BD, UK*

Corresponding author: Robert B. Angus (r.angus@rhul.ac.uk)

Abstract

An account is given of my development of techniques to obtain well-spread Giemsa-stained banded chromosome preparations. Apparent G-banding could be obtained following very slight trypsin treatment of freshly prepared slides, but this banding was very fine (close-grained) and possibly not a reflection of chromosome structure. However, treatment of developing embryos *in vitro* with 5-fluorouridine produced a similar chromomere banding, which is therefore regarded as genuine. Steady accumulation of *Helophorus* Fabricius, 1775 karyotypes has resulted in the production of an Atlas covering 62 of the 170 species known to occur in the Palaearctic. Chromosome polymorphisms involving pericentric inversions and addition of extra C-banding regions have been found, as well as small B-chromosomes in a few species. In general, karyotypes have proved very useful in establishing the limits of individual species. Parthenogenesis involving triploidy has been found in two species. Karyotypes of experimentally produced hybrids have revealed irregularities in chromosome condensation.

Keywords

banding, chromosomes, experimental hybrids, *Helophorus*, karyotypes, parthenogenesis, triploidy

Introduction

My investigation of *Helophorus* chromosomes began in 1975 with my appointment as a Lecturer in the Zoology Department of Royal Holloway College, University of London. Earlier attempts at chromosome preparation had resulted in complete failure, but now the field was beginning to open up. The paper by Crozier (1968) describing an acetic acid dissociation, air-drying technique was a breakthrough. It

allowed preparations of well spread undistorted chromosomes. Initially Crozier had used aceto-lactic orcein staining, but application of Giemsa stains had already been described for similarly prepared mammalian chromosomes (Rothfels and Siminovitch 1958) and this gave excellent results. All my early preparations were from developing embryos.

In those early days insect chromosomes were known to display C-banding and to show active nucleolus organisers (NORs) by silver staining. G-banding was another matter.

C-banding is associated with highly repetitive DNA, with one base-pair to a short sequence of base-pairs repeated many times. Such bands are present in both dividing and interphase chromosomes. It is generally observed following treatment with alkali (for me saturated $Ba(OH)_2$ at room temperature), followed by incubation in salt-sodium citrate (2X SSC) at about 60°C. There have been attempts to differentiate "true C-bands" from other less distinctive types. With beetles a pretreatment with 1N HCl has been recommended–applied to my chromosome preparations it abolishes all traces of banding!

I have found silver staining tricky. I have not succeeded with acetic acid inflated material but can get it to work with centrifuge-spread material. The results are consistent.

G-banding is where the real rewards may lie, enabling chromosomes and even sections of chromosomes to be identified with great precision, demonstrating homologies between chromosomes of different species and their relatedness as with Man and the Great Apes (Pearson 1997) and the Giant Panda and the Brown Bear (O'Brien et al. 1985).

Published information on G-banding was not encouraging. Maudlin (1974) published information on G-banding in triatomine bugs (Heteroptera), but in most of the chromosomes there are only a few bands. Steiniger and Mukherjee (1975) obtained banding patterns in the mosquito *Aedes albopictus* (Skuse, 1895) by reducing normal fixation times. The results appear dramatic but ragged and certainly not finegrained. Webb (1976) obtained spectacular banding on B-chromosomes of the Australian plague locust *Chortoicetes terminifera* (Walker, 1870). These B-chromosomes are heterochromatic and the bands were in the same positions whether resulting from G- or C-banding protocols. Tambasco et al. (1974) reported G-banding in South American stingless bees, but again the bands were few in number, and hard to see in the photograph.

Bigger (1975), using the centrifugation method with cell suspensions, produced what he claimed to be G-bands on various Lepidoptera including the Large White butterfly, *Pieris brassicae* (Linnaeus, 1758). His published photographs are difficult to interpret, and his diagrams are interesting but may have to some extent been guided by the "eye of faith". However, Dutrillaux et al. (2022), using more refined microscopy, especially confocal microscopy, showed localized primary constrictions as well as some banding. In 2004, working with L. A. Dutton, then an undergraduate student doing a research project, I obtained some well-spread preparations from eggs, treated with

5-fluorouridine. Although not nearly as convincing as the *Helophorus* chromosomes shown in in Fig. 1 of this paper, they did hint at possible fine-grained banding.

After much experimentation I found that bands could be produced by a very slight trypsin treatment of freshly prepared slides (5 min. drying immediately after preparation)–0.01% Difco 1:250 trypsin in 0.75% NaCl buffered to pH 7.6 with Sörensen - for 5–15 sec. at 10°C, then quenched by rinsing in three changes of distilled water at pH 6.0. This is a very slight treatment but can give very good results, with numerous bands on all the chromosomes, inviting the hope that results comparable with those obtained from bear chromosomes might be possible. The problem was, the banding produced is not only very fine-grained but also very even, so did it reflect chromosome organisation or merely the last bits not destroyed by the trypsin? A solution came from studies by Rønne and others using various "antibiotic" reagents on *in vitro* cultures of human cells (Rønne 1977); (Rønne and Andersen 1978). Cycloheximide, chosen because it was relatively cheap, had very limited success, but 5-fluorouridine, which appeared to give the clearest results with human cells (Rønne and Andersen 1978) gave some very clear results (Fig. 1). Fig. 1 shows a comparison of the longest autosomes of *Helophorus aquaticus* (Linnaeus, 1758) and *H. aequalis* Thomson, 1868. The chromosomes of the two species show a similar sequence of bands (allowing for the different sizes of their centromeric C-bands), except for the distal region of the short arm, beyond a fairly distinct gap in about the middle of the arm. In *H. aquaticus* there are three very distinct bands in this distal section, but in *H. aequalis* the bands are less distinct, comprising a basal one with a hint of subdivision and apical less stained and more indistinct section. The conclusions are that the banding reflects the chromomeric organisation of the chromosomes, and that the apical sections of the short arms have been involved in translocations. The details are explored further by Angus (1982).

The extent to which this fine-grained banding, however useful, is the same as the G-banding obtained with mammalian chromosomes remains to some extent an open question. One interesting feature of mammalian G-banding is that the bands correspond with those observed on pachytene chromosomes during meiosis (Luciani et al. 1975). Dutrillaux et al. (2006) developed methods of using pachytene banding in beetles, and some of these bands appear very similar to the fine-grained banding in *Helophorus*. It therefore seems that these are G-bands.

Work on *H. aquaticus* and *H. aequalis* required chromosomally verified material to establish the extent of their morphological variation, especially of the aedeagus. To begin with, testes of freshly emerged adults were used as a source of mitotic chromosomes, and this solved the problem. Later the technique was extended to mid gut, where undifferentiated cells in the mid gut crypts undergo mitosis to replace epithelial cells lost in the course cells of food-digestion. I had been steadily accumulating karyotypes of various *Helophorus* species, and 1989 I produced a preliminary Atlas, covering 31 species, for the Balfour-Browne Club Newsletter (Angus 1989). This work has continued, often focusing on groups of similar-looking species requiring taxonomic clarification. So now there is a new version of the Atlas with 62 species. This is presented here, with Figs 2–12.
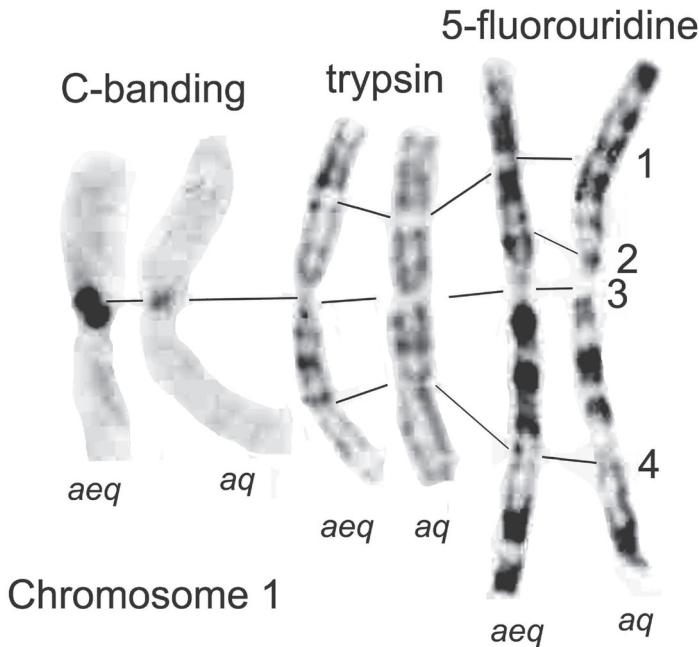
Figure 1. Detailed comparison of the banding patterns of Chromosome 1 of *Helophorus aequalis* (*aeq*) and *H. aquaticus* (*aq*). The lines joining the chromosomes indicate homologous points. Treatments are indicated above the illustrated chromosomes.

Atlas of *Helophorus* chromosomes

Helophorus species divide into two karyotype-groups, those with eight pairs of autosomes plus Xy_p sex chromosomes (the so-called "parachute-association" with the very small y chromosome attached to the X by a nucleolus or cytoplasmic vesicle, described by John & Lewis (1960) and with the possibility that the cytoplasmic vesicle was not always a true nucleolus (Juan et al. 1993) (subgenera *Helophorus* s. str., *Gephelophorus* Sharp, 1915 and *Eutrichelophorus* Sharp, 1915), and those with 10 pairs of autosomes plus Xy_p sex chromosomes (subgenera *Empleurus* Hope, 1838, *Trichohelophorus* Kuwert, 1886, *Lihelophorus* Zaitzev, 1908 and *Rhopalohelophorus* Kuwert, 1886).

Subgenus *Helophorus* s. str.
Figs 2a–j, 3a–j, 4a–g

Species of *Helophorus* s. str. divide morphologically into three groups, the *H. aquaticus* group with the last fixed abdominal segment bearing small but clearly square-ended teeth, the *H. grandis* Illiger, 1798 group, with much larger teeth and the *H. bergrothi* J. Sahlberg, 1880 group, in which the abdominal sternite is crinkled apically but with the shape of the teeth not really discernible except sometimes in cleared, slide-mounted preparations (Angus 1970a). One particularly distinctive feature of the karyotype,

originally discovered in *H. aequalis*, is the presence of a distinct secondary constriction, confirmed by silver staining as the site of a NOR (Angus 1982). In *H. aequalis* this chromosome goes as pair 6 in the row of chromosomes in the karyotype, and in other species the NOR-bearing chromosome is placed as pair 6 for ease of comparison.

H. aquaticus (Fig. 2a, b). The NOR-bearing chromosome 6 is about as long as pair 3, depending on the degree of opening of the secondary constriction. The centromeric C-bands are small (see Fig. 1) and the X chromosome is submetacentric.

H. thauma Angus et Toledo, 2010 (Fig. 2c, d). An Italian species very closely resembling *H. aequalis* but distinguished chromosomally by the NOR-bearing chromosome 6 being as long or longer than autosome 3.

H. aequalis (Fig. 2e, f). The centromeric C-bands are fairly strong, and autosome 7 and the X chromosome are subacrocentric. The NOR-bearing chromosome 6 is shorter than 5, but about the same length as pair 7. Other details of comparison with *H. aquaticus* are given in the section discussing banding, and by Angus (1982).

H. grandis (Fig. 2g–j). Although this is the first of the big-toothed group of species, it has been found by Martin Fikáček (*pers. comm.*16.VII.2023) in the course of his ongoing DNA analysis, to be the sister-species of *H. aequalis*, and chromosomally this is supported by the size and shape of the NOR-bearing autosome 6. The centromeric C-bands are clearly larger than in *H. aequalis*, and there may be an interstitial C-band in the middle of the long arm of the acrocentric X chromosome, which is thus polymorphic for long and short forms (Fig. 2h–j). Silver-staining (Fig. 2i) shows the interstitial C-band behaving rather differently from the centromeric one. The short form matches the *H. aequalis* X. Autosome 5 is polymorphic for a pericentric inversion, and may be either metacentric as in *H. aequalis*, or acrocentric. Smith (1960) correctly recorded Canadian "*H. aquaticus*" (actually *H. grandis*, a Palaearctic species introduced in Canada) as having 18 chromosomes including Xy_p. He also gave this number for *H. oblongus* LeConte, 1850, a Holarctic species of the subgenus *Rhopalohelophorus*, the group with 8-segmented antennae, and therefore expected to have 22 chromosomes including Xy_p. This result needs to be checked.

H. liguricus Angus, 1970 (Fig. 3a, b). The position of the NOR is not clear but it may be at the distal end of the short arm of autosome 7. Autosome 6 is acrocentric and the X chromosome is a smallish metacentric, similar in size to autosomes 4 and 5. The C-bands vary in size between the chromosomes, apparently absent from pair 1, very small in pairs 2 and 3, slightly larger in the others.

H. maritimus Rey, 1885 (Fig. 3c, d). The centromeric C-bands are small, the NOR may be located at the distal end of the long arm of autosome 6, and the X chromosome is a fairly long metacentric, about as long as autosomes 4 and 5.

H. occidentalis Angus, 1983 (Fig. 3e, f). Known from the southern parts of Spain and Portugal, and from Morocco. NOR-bearing autosome 6 is small, about as long as pairs 3–5, slightly longer than pairs 7 and 8. The metacentric X chromosome is slightly larger than autosomes 7 and 8, but smaller than 6. The centromeric C-
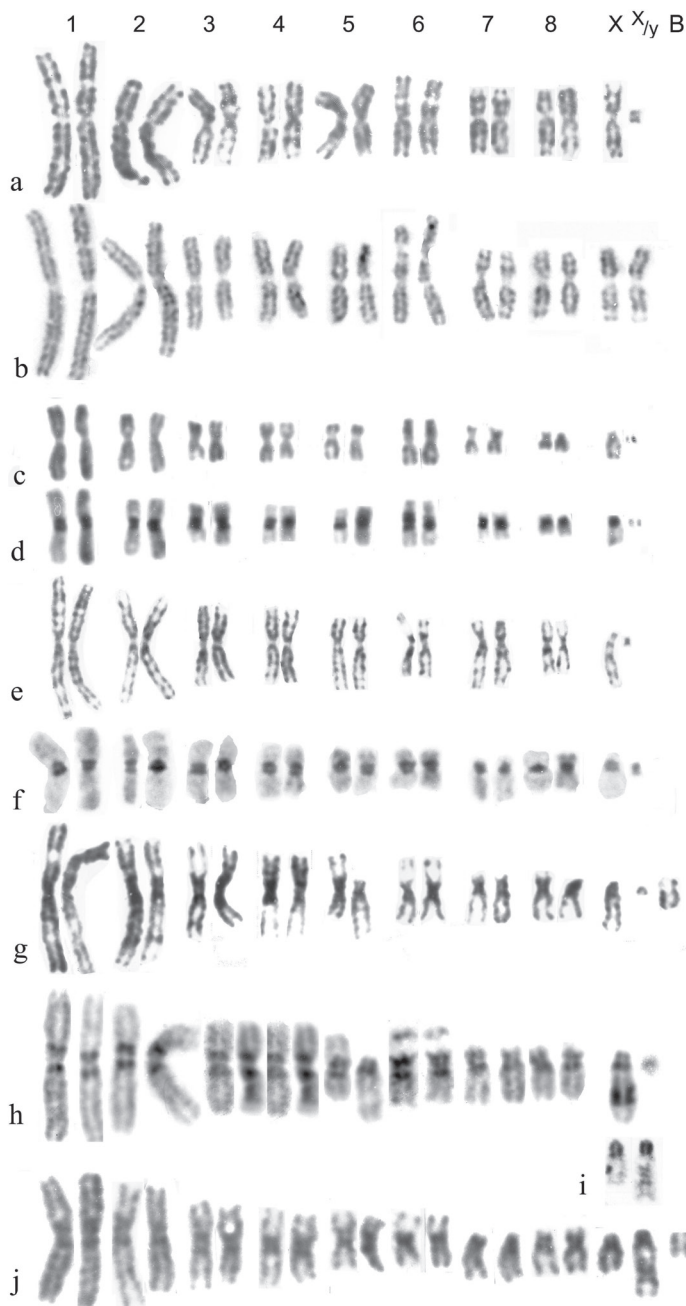
Figure 2. a–j *Helophorus* str. Mitotic chromosomes arranged as karyotypes **a, b** *H. aquaticus*, embryos, banded with trypsin **a** ♂, France, Fontanières **b** ♀, Russia, Strelna near St Petersburg **c, d** *H. thauma*, paratype ♂, mid gut **c** Giemsa-stained **d** the same nucleus, C-banded **e, f** *H. aequalis*, France, mid gut **e** Giemsa-stained **f** C-banded **g–j** *H. grandis*, embryos **g** ♂, France, Giemsa-stained **h** ♂, Russia, Pavlovsk near St Petersburg, C-banded **i** long and short X chromosomes C-banded by silver-staining **j** ♀, England, Surrey showing the long and short X chromosomes. Scale bar: 15 μm.
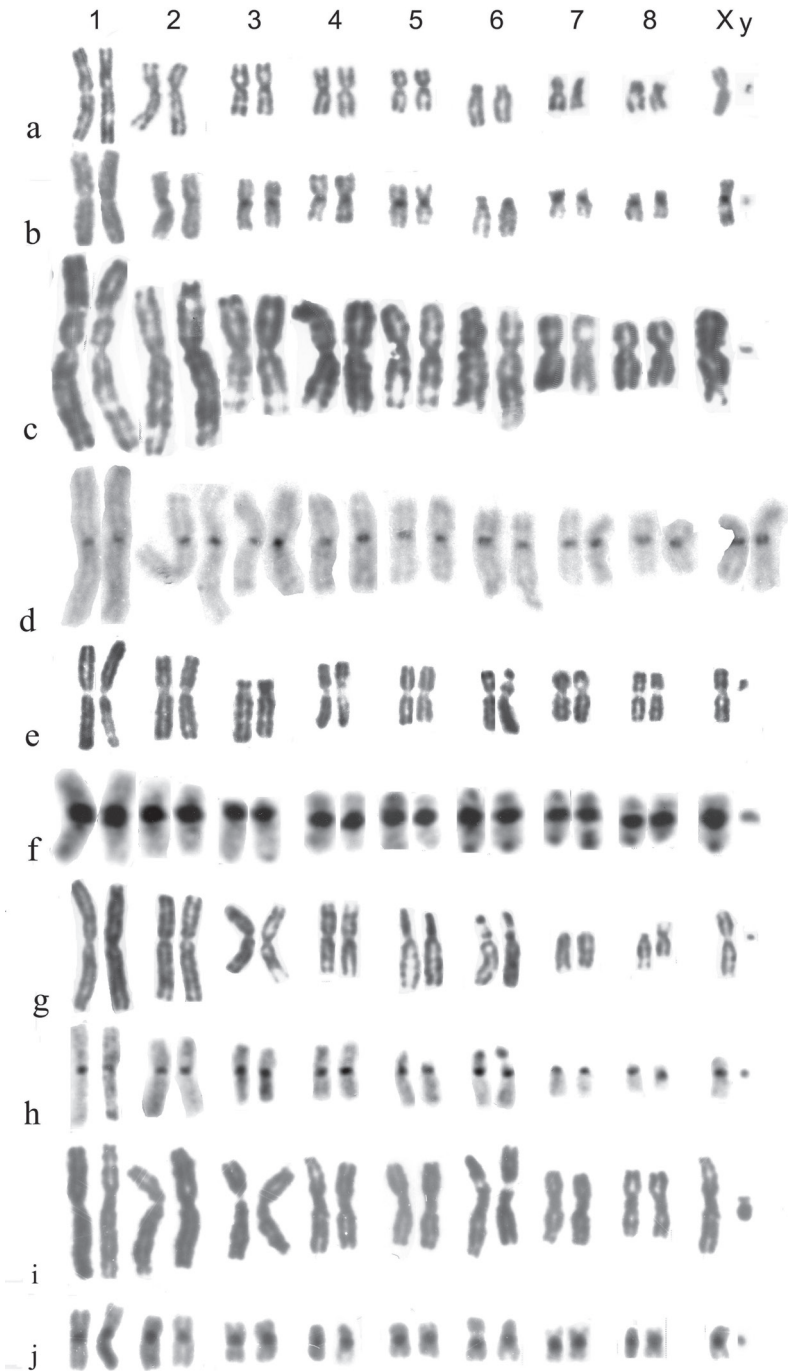
Figure 3. a–j *Helophorus* str. Mitotic chromosomes arranged as karyotypes **a, b** *H. liguricus*, ♂, Corfu, mid gut **a** Giemsa-stained **b** C-banded **c, d** *H. maritimus*, embryos, France, Camargue **c** ♂, Giemsa-stained **d** ♀, C-banded **e, f** *H. occidentalis*, mid gut, Spain, Province of Cáceres, Abadia **e** Giemsa-stained **f** C-banded **g, h** *H. milleri*, ♂, mid gut, Corfu **g** Giemsa-stained **h** C-banded **i, j** *H. syriacus*, ♂, mid gut, Israel **i** Giemsa-stained **j** C-banded. Scale bar: 15 μm.

bands are particularly heavy in all chromosomes except the dot-like y, pairs 6 and 7 have terminal C-bands at both ends and pairs 1 and 8 are polymorphic for the presence of a small C-band at the end of their long arms.

H. milleri Kuwert, 1886 (Fig. 3g, h). Characterised by small centromeric C-bands, autosome 5 being acrocentric, autosome 6 with its NOR located medially on the short arm and matching its position, being about the same size as autosome 5 but longer than 7. The distal part of the short arm, beyond the NOR, is heterochromatic, and there may be a C-band at the distal end of the long arm. Autosome 8 is polymorphic for a pericentric inversion, resulting in metacentric and acrocentric forms. The X chromosome is metacentric, longer than autosomes 7 and 8, but slightly shorter than 6. Described from Corfu, this species is widespread in the central Mediterranean area.

H. syriacus Kuwert, 1885 (Fig. 3i, j). The NOR-bearing autosome 6 is as long as pair 3 and the metacentric X chromosome is also long, as pair 4. The centromeric C-bands are fairly heavy, smaller and fainter on pairs 2 and 8. My material is from Israel, but this species is widely distributed from western Anatolia (and adjacent Greek islands) east to the mountains of Kazakhstan (Aksu-Dzhabagli).

H. oscillator Sharp, 1915 (Fig. 4a, b). Originally placed in *Trichohelophorus* Kuwert, 1886 by Sharp, this species was transferred to *Helophorus* s. str. by Angus et al. (2019), largely because of its karyotype. The chromosomes are all metacentric, with fairly large centromeric C-bands and the small y chromosome is also heavily C-banded. The intensity of the bands varies but this may be an experimental artefact.

H. hammondi Angus, 1970 (Fig. 4c, d). Autosomes 1–7 are more or less metacentric with moderate centromeric C-bands. The C-banding of pair 4, with a weak band in the middle of the short arm, suggests this may be the site of the NOR. Autosome pair 8 and the X chromosome are subacrocentric, with the X chromosome about the same size as pair 7.

H. jaechi Angus, 1995 (Fig. 4e–g). The general layout of the chromosomes is similar to that of *H. hammondi*, but autosome pair 4 is subacrocentric and the X chromosome is distinctly larger, about as long as pair 3. The centromeric C-bands are small but distinct.

Subgenus *Gephelophorus*
Fig. 4h–k

H. sibiricus Motschulsky, 1860 (Fig. 4h–j). Autosomes 1–6 and 8 are metacentric, 7 is acrocentric and the metacentric X chromosome is the longest in the nucleus. All the chromosomes, except the tiny y have heavy centromeric C-bands, and pair 7 has a size polymorphism with, in Fig. 4i, the longer replicate of pair 7 with an apparent C-band in the middle of its long arm. The somewhat fainter appearance of this band matches that of *H. grandis* when silver-stained (Fig. 2i). The nucleus shown in Fig. 4j has 2 y chromosomes.
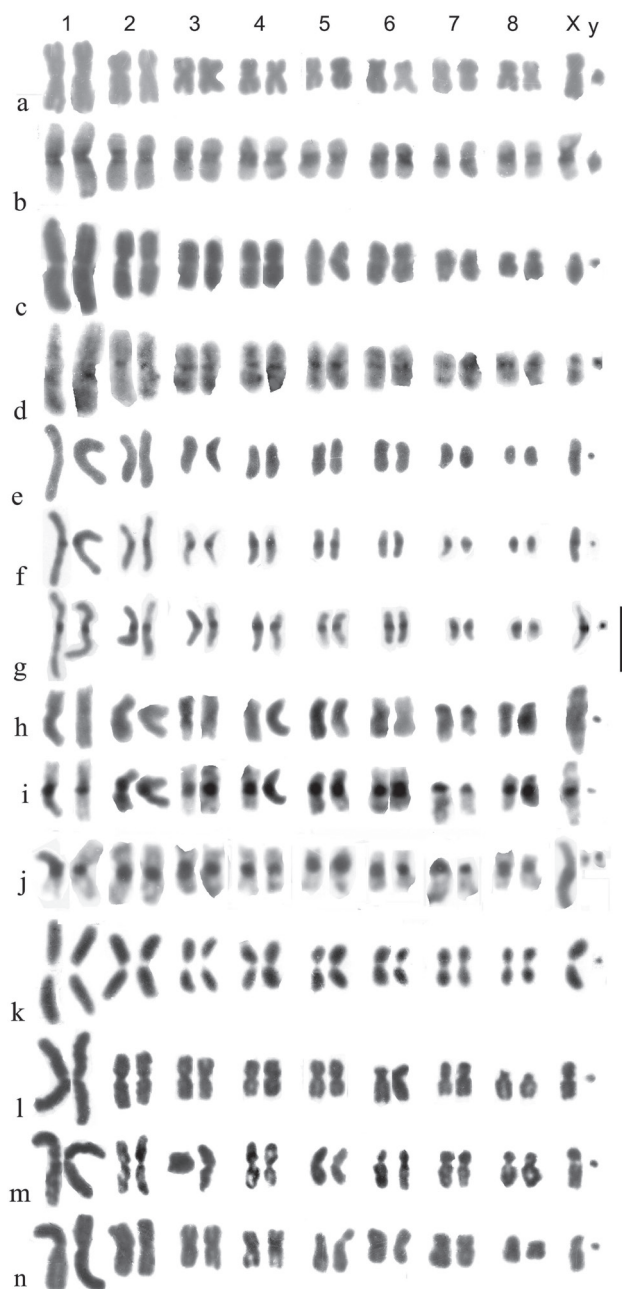
Figure 4. a–n Mitotic mid gut chromosomes of subgenera *Helophorus* s. str, *Gephelophorus* and *Eutrichelophorus*, arranged as karyotypes **a–h** *Helophorus* s. str **a, b** *H. oscillator*, ♂, Israel, Golan, Einot Summaga **a** Giemsa-stained **b** C-banded **c, d** *H. hammondi*, China, Qinghai, Gangca **c** Giemsa-stained **d** C-banded **e–g** *H. jaechi*, China, Sichuan, Xinduqiao **h–k** *Gephelophorus* **h–j** *H. sibiricus*, ♂, China, Heilongjiang, Mishan **h** Giemsa-stained **i** the same nucleus C-banded **j** a different nucleus from the same specimen, C-banded **k** *H. auriculatus*, ♂, Japan, Saitama prefecture near Tokyo, Giemsa-stained **l–n** *Eutrichelophorus*, ♂, Giemsa-stained **l, m** *H. micans* **l** Crete, Rethymnon **m** Hungary **n** *H. oxygonus*, Morocco, Ifrane. Scale bar: 15 μm.

H. auriculatus Sharp, 1884 (Fig. 4. k). All the autosomes, and the X chromosome, are metacentric, with the X chromosome about as long as pair 3. The y is very small, dot-like. No C-banding is available, but the distinct centromeric gaps suggest the presence of large C-bands.

Subgenus *Eutrichelophorus*
Fig. 4l–n

H. micans Faldermann, 1835 (Fig. 4l, m). Ongoing DNA investigation by Martin Fikáček (*pers. comm.*16.VII.2023) associates this species with *Helophorus* s. str., in agreement with its chromosome number. Autosome pairs 1–5 are metacentric, 6, 7 and the X chromosome are borderline submetacentric/subacrocentric, and pair 8 is subacrocentric. The y chromosome is very small, dot-like.

H. oxygonus Bedel, 1881 (Fig. 4n). A very similar karyotype to that of *H. micans* but with pair 5 borderline acrocentric/subacrocentric and possibly longer than pair 4, and pair 6 metacentric.

Subgenera with karyotypes of 20 +Xy$_p$

Subgenus *Empleurus*
Fig. 5a–c

H. nubilus Fabricius, 1777 (Fig. 5a). All the autosomes, and the X chromosome are metacentric, with the X chromosome about as long autosome 2. The y chromosome is a dot. No C-banding is available, but this Giemsa-stained karyotype suggests that at least some of the autosomes have large centromeric C-bands.

H. rufipes Bosc, 1791 (Fig. 5b, c). The general layout of the karyotype is similar to that of *H. nubilus*. This is especially clear in the Spanish specimen (Fig. 5b).

Subgenus *Trichohelophorus*

H. alternans Gené, 1836 (Fig. 5d). All the autosomes and the X chromosome are metacentric, with pairs 9 and 10 approaching the border with submetacentric. The y is a dot.

Subgenus *Lihelophorus*
Fig. 5e–l

The three species of this subgenus are endemic to the Tibetan Plateau. They are unique in *Helophorus* in having the outermost elytral interval (interval 10) completely flat, so that there is no trace of pseudepipleura outside the elytral epipleurs. The combination of elytral intercalary (scutellary) striae and asymmetrical apical segments of the maxillary palpi suggests association of *Lihelophorus* with *Helophorus* s. str. but the chromosomes show that this is not the case. The subgenus was reviewed by Angus et al. (2016).

Figure 5. a–l Mitotic mid gut chromosomes of subgenera *Empleurus*, *Trichohelophorus* and *Lihelophorus*, arranged as karyotypes **a–c** *Empleurus* **a** *H. nubilus*, ♂, Spain, Provincia de Salamanca, El Cubo, Giemsa-stained **b, c** *H. rufipes*, ♂, Giemsa-stained **b** Spain, Provincia de Segovia, Santa Maria la Real de Nieva **c** England, Worcestershire **d** *Trichohelophorus alternans*, ♂, Sardinia, Giemsa-stained **e–l** *Lihelophorus*, ♂, China, Qinghai, Zuimatan **e, f** *L. lamicola* **e** Giemsa-stained **f** C-banded **g, h** *L. ser* **g** Giemsa-stained, the y chromosome lost from this preparation **h** C-banded, with the y from a different preparation **i–l** *L. yangae* **i, k** Giemsa-stained **i, j** and **k, l** the same nuclei, Giemsa-stained and C-banded. Scale bar: 15 μm.

H. lamicola Zaitzev, 1908 (Fig. 5e, f). All the autosomes are more or less metacentric with distinct centromeric C-bands. The X chromosome, similar in size to autosome 10, is subacrocentric, again with a distinct centromeric C-band. The small, almost dot-like y chromosome also has a small C-band.

H. ser Zaitzev, 1908 (Fig. 5g, h). The general layout of the karyotype is very similar to that of *L. lamicola*. The X chromosome is slightly larger and with slightly longer short arms. The y chromosome is dot-like,

H. yangae Angus et al., 2016 (Fig. 5i–l). Autosomes 4–6 are clearly less metacentric than in the other two species, and the X chromosome is slightly larger, similar in size to autosome pair 7 rather than pair 8.

Subgenus *Rhopalohelophorus*
Figs 6a–m, 7a–i, 8a–p, 9a–p, 10a–o, 11a–g

Informal group *Atractohelophorus* (Fig. 6a–m). *Atractohelophorus* refers to the small species with symmetrically oval apical segments on their maxillary palpi. In most of Europe by far the commonest species is *H. brevipalpis* Bedel, 1881, and many of the other species tend to be associated with mountains.

H. brevipalpis, bisexual, diploid (Fig. 6a, b). Autosome pairs 1, 2, 4 and 7, and the X chromosome are metacentric, 3, 6, 8, 9 and 10 are borderline acrocentric/subacrocentric and 6 is acrocentric in the Spanish specimen (Fig. 6a) polymorphic for a pericentric inversion, either subacrocentric or metacentric in the Cretan one (Fig. 6b). The y chromosome is dot-like. For parthenogenetic triploids see Fig. 11a–d.

H. montenegrinus Kuwert, 1885 (Fig. 6c, d). The karyotype is very like that of *H. brevipalpis*, but autosome pair 3 is metacentric and 6 is submetacentric.

H. glacialis Villa et Villa, 1833 (Fig. 6e, f). Autosome pairs 1–4, 6 and 10 are metacentric and 5 and 7–9 are subacrocentric. The metacentric X chromosome is clearly the longest in the nucleus, a feature shared with *H. redtenbacheri* Kuwert, 1885 (Fig. 7d). The y chromosome is small, almost dot-like.

H. leontis Angus, 1985 (Fig. 6g, h). Autosomes 1–7 and 10 are metacentric, 8 and 9 and the X chromosome are submetacentric. The X chromosome is about the same size as autosomes 6 and 7.

H. dixoni Angus, 1987 (Fig. 6i). No male karyotype is available, so the X chromosome cannot be identified. Chromosomes 1–8, on the arrangement adopted here, match those of *H. leontis*. Of the smaller autosomes, one, placed as pair 10, is clearly smaller than anything in the *H. leontis* karyotype. In the current arrangement the X chromosome would be smaller than that of *H. leontis*, about the same size as pair 8.

H. biltoni Angus et al., 2005(Fig. 6j). Autosomes 1–6 match those of *H. leontis* and *H. dixoni*, but pairs 7–9 are smaller, and the small autosome 10 matches chromosome 10 of *H. dixoni*. The X chromosome is a small acrocentric, clearly smaller than the *H. leontis* X chromosome and not matching any of the *H. dixoni* chromosomes.

Figure 6. a–m Subgenus *Rhopalohelophorus*, informal grouping *Atractohelophorus*. Giemsa-stained mitotic mid gut chromosomes arranged as karyotypes **a, b** *H. brevipalpis*, diploid ♂♂ **a** Spain, Province of León, Algadefe **b** Crete, Rethymnon **c, d** *H. montenegrinus* **c** Bulgaria, Rila **d** Italy, Stirone **e, f** *H. glacialis* **e** ♂, Spain, Provincia de Madrid, Peña Labra **f** ♀ Corsica, Haute-Corse, Restonica **g, h** *H. leontis*, ♂, Spain, Province of Madrid, Peña Lara **i** *H. dixoni*, ♀, Israel, Golan **j** *H. biltoni*, Iran, Fars Province, Sishpir **k** *H. nevadensis*, ♂, Spain, Province of Madrid, Peña Lara **l** *H. korotyaevi*, ♂, Spain, Province of Cantabria, Puerto de Piedrasluengas **m** *H. lewisi*, ♂, Israel, Golan, Einot Summaga. The positions of missing chromosomes are indicated by small black discs. Scale bar: 15 μm.

H. leontis, *H. dixoni* and *H. biltoni* are a group of species which cannot be separated by their aedeagal morphology, though their body-forms differ. Their karyotypes leave no doubt that they are separate species.

H. nevadensis Sharp, 1916 (Fig. 6k). Autosomes 1, 2 and 4 are metacentric, with pair 1 about twice the length of pair 2. The remaining autosomes, and the X chromosome, are acrocentric to subacrocentric. The X chromosome is about the same length as autosome 8. One B-chromosome is present, about the same size as the diminutive y. I have seen this chromosome in both the males from which I have obtained karyotypes. I have no female preparations and cannot say which of the tiny chromosomes is the y and which is a B. They are both about a third of the length of the X.

H. korotyaevi Angus, 1985 (Fig. 6l). Autosomes 1–5, and the X chromosome are metacentric and autosomes 6–10 are submetacentric to subacrocentric. The X chromosome is about the same size as autosomes 6 and 7. The diminutive y chromosome is about a third of the length of the X.

H. lewisi Angus, 1985 (Fig. 6m). Autosomes 1–6 and the X chromosome are metacentric, 7 and 8 are acrocentric and 9 and 10 are subacrocentric. The X chromosome is about as long as autosome 6 and the diminutive y is about a quarter that length.

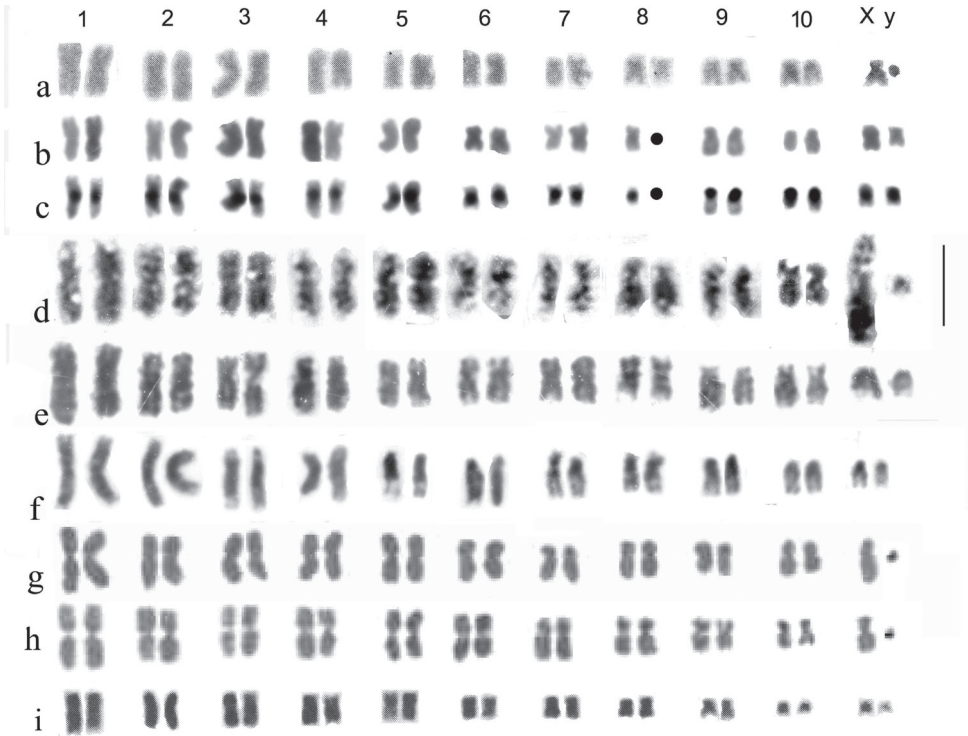## *Rhopalohelophorus*, species with 8-segmented antennae
Fig. 7a–i

Not a natural group, but convenient.

H. nanus Sturm, 1836 (Fig. 7a–c). Probably the most widely distributed species in the Palaearctic, from Britain, Ireland and France in the west to the Russian Far East (Primorye) in the east. Autosomes 1–8 are metacentric, 9 and 10, along with the X chromosome are subacrocentric. The diminutive, almost dot-like, y chromosome is about a third of the length of the X, itself one of the shortest in the nucleus. There appears to be no morphological difference between the French and Chinese specimens figured here.

H. redtenbacheri (Fig. 7d). This is one of the preparations made at Karasuk in 1982 and slide-mounted using polymerising UV setting resin. Unfortunately, it had deteriorated badly before it was photographed. Nevertheless, the main morphological features of the chromosomes can be discerned. As mentioned in the discussion of *H. glacialis* (Fig. 6f, g), this is a species whose X chromosome is clearly the longest in the nucleus. Autosomes 1–3, 5–7, and the X chromosome, are metacentric. Autosomes 4 and 8–10 are submetacentric.

H. pallidus Gebler, 1830 (Fig. 7e) As no male karyotype is available the X chromosome cannot be recognised. Chromosomes 1 and 2 and 5–8 are metacentric, 3 and 4 are submetacentric, and the rest are subacrocentric.

Figure 7. a–i Subgenus *Rhopalohelophorus*, species with 8-segmented antennae **a–c** *H. nanus* **a** ♂ embryo, France, Beaumont-sur-Sarthe, Giemsa-stained **b, c** ♀, mid gut, China, Heilongjiang, Mishan **b** Giemsa-stained **c** C-banded **d** *H. redtenbacheri*, ♂, embryo, Russia, West Siberia, Karasuk, Giemsa-stained **e** *H. pallidus*, ♀, embryo, Russia, West Siberia, Karasuk, Giemsa-stained **f** *H. villosus*, ♀, Germany, Bavaria, Deggendorf, embryo, Giemsa-stained **g, h** *H. pallidipennis*, ♂, embryo, Giemsa-stained **g** Cyprus **h** Crete **i** *H. kervillei*, ♀, embryo, Giemsa-stained, Corfu. The positions of missing chromosomes are indicated by small black discs. Scale bar: 15 µm.

H. villosus Duftschmid, 1805 (Fig. 7f). No male karyotype is available so the X chromosome cannot be identified. Chromosomes 1–4 are metacentric, 5, 7 and 9 are submetacentric, 8 is borderline submetacentric/subacrocentric and 6, 10 and 11 are acrocentric to subacrocentric.

H. pallidipennis Mulsant et Wachanru, 1852(Fig. 7g, h). Autosome pairs 1–6 are metacentric, while 7- 10 and the X chromosome are subacrocentric. The X chromosome is about the same length as autosomes 7 and 8 and the y is dot-like.

H. kervillei d'Orchymont, 1932 (Fig. 7i). No male karyotype is available so the X chromosome cannot be recognised. Chromosomes 1–8 are metacentric, 9 is submetacentric and 10 and 11 are apparently metacentric, but very small. This species, like *H. kirgisicus* Kniž, 1914, has only two larval instars. Angus (1992a) regarded this as a form of *H. pallidipennis*, which he therefore described as having only two

larval instars. Only when information on Cretan and Cypriot *H. pallidipennis* revealed not only a different karyotype but also a third larval instar, did the truth become apparent (Angus, 1998).

Rhopalohelophorus, the *H. minutus* Fabricius, 1775 group
Fig. 8a–p

For experimental hybrids see later, Fig. 12.

H. minutus (Fig. 8a–c). Autosomes 1–7 and 9 are metacentric, pairs 8 and 10, and the X chromosome are subacrocentric. The X chromosome is about the same size as pair 8 and the diminutive y is almost dot-like, perhaps metacentric. One autosome has a modified apical part of the short arm, probably the site of a NOR, very pale in one replicate of the Giemsa-stained pair, slightly C-banded in the English C-banded preparation (Fig. 8b) and more strongly so in the Spanish one (Fig. 8c). This autosome was originally placed as pair 2 (Angus, 1986) but, bearing in mind the large variation in the degree of condensation of this autosome and a more averaged interpretation of its length, placing it a pair 4 seems more appropriate. This also agrees with its position in the related *H. atlantis* and *H. calpensis*. Autosome 9, an even metacentric, often takes the form of a multiplication sign (X) and is one of the landmarks of the *H. minutus* karyotype. Angus (1986) reversed the positions of pairs 8 and 9, to place the metacentrics before the submetacentrics. This is in fact unhelpful and counter to the Relative Chromosome Length data presented in Table 1 of Angus (1986). The positions of autosomes 7 and 8 have also been reversed in the light of study of more material, including hybrids. *H. minutus* is widely distributed over much of Europe.

H. atlantis Angus et Aouad, 2009 (Fig. 8d, e). The karyotype is very similar to that of *H. minutus*, but the NOR appears to be at the distal end of the short arm of pair 4, pair 9 is less evenly metacentric and the X chromosome is as small as pair 10. The centromeric C-bands are noticeably heavy. This species is known from the Moyen Atlas of Morocco.

H. calpensis Angus, 1988 (Fig. 8f–j). The karyotype is very similar to that of *H. atlantis*, the most obvious difference being the size of the y chromosome, acrocentric and about half the length of the X. The position of the NOR-bearing chromosome is not easy to establish due to irregularities in condensation, but it appears to belong in position 4, as in *H. minutus* and *H. atlantis*. *H. calpensis* is so far known only from southernmost Spain, Tarifa and the Coto Doñana.

H. paraminutus Angus, 1986 (Fig. 8k, l). This species was initially recognised by Angus (1986) because it has a karyotype was apparently indistinguishable from that of *H. lapponicus* Thomson, 1858 but an egg cocoon like that of *H. minutus*, not *H. lapponicus*. Also, the beetles looked more like *H. minutus* than *H. lapponicus*, though they were often larger. Fig. 8k, l shows chromosomes from trypsin-treated

Figure 8. a–p Subgenus *Rhopalohelophorus*, *H. minutus*-group **a–c** *H. minutus* ♂, embryos **a, b** England, Surrey, Runnymede **a** Giemsa-stained **b** C-banded **c** Spain Province of Segovia, Villacastín, C-banded **d, e** *H. atlantis*, ♂, embryos, Morocco, Ifrane **d** Giemsa-stained **e** C-banded **f–j** *H. calpensis*, Spain **f–h** ♂ Provincia de Cádiz, Tarifa, embryos **f** Giemsa-stained **g, h** C-banded **i, j** ♀, Province of Huelva, Coto Doñana, mid-gut **i** Giemsa-stained **j** the same nucleus, C-banded **k, l** *H. paraminutus*, ♂, embryos, Giemsa stained **k** Russia, West Siberia, Karasuk **l** Austria, Neusiedler See area **m–p** *H. lapponicus*, ♂ **m–o** embryos **p** mid gut **m** ♀, Spain, Province of Cantabria X ♂, Sweden, Västerbotten **n, o** Russia, West Siberia, Karasuk **n** treated with cycloheximide then Giemsa-stained **o** Giemsa-stained **p** Israel, Golan, Einot Summaga, Giemsa-stained. Scale bar: 15 μm.

Giemsa-stained embryos. Autosome pairs 1–6 are metacentric, while the others, and the X chromosome, are acrocentric to subacrocentric. The X chromosome is about the same size as autosome 7 and the y chromosome is dot-like. None of the metacentric chromosomes shows any indication of a terminal NOR.

H. lapponicus (Fig. 8m–p). The karyotypes shown in m and o are from embryos, m from a Spanish female crossed with a Swedish male, and o from Karasuk. The arrangement and banding patterns of the chromosomes appear identical. Note that one replicate of autosome in m has been damaged in the course of preparation. As mentioned above, the arrangement appears to be the same as that of *H. paraminutus*. Fig. 8n shows a preparation from a Karasuk embryo which was treated *in vitro* with cycloheximide. There is no trace of banding but autosome 9 shows some extension of the short arm, suggesting that this may be the site of the NOR. Fig. 8p shows a karyotype from a mid-gut cell of an Israeli specimen. The sequence of sizes and shapes of the chromosomes appears the same as in the other material.

Rhopalohelophorus, various species

H. fulgidicollis Motschulsky, 1860 (Fig. 9a). A trypsin-treated Giemsa-stained preparation from an embryo. No banding has resulted. Autosomes 1–5 and 8, 9 and the X chromosome are metacentric. The X chromosome is slightly shorter than pair 5, and the y is dot-like. Pairs 6 and 7 are submetacentric and the rounded condensed appearance of the short arm of 7 suggests this may be the site of the NOR. Pair 10 is a short acrocentric.

H. asturiensis Kuwert, 1885 (Fig. 9b). The karyotype appears very similar to that of *H. fulgidicollis*, though the beetles and their aedeagi are quite distinctly different.

H. kirgisicus Kniž, 1914 (Fig. 9c, d). This is another of the preparations which had partially decomposed in the polymerising resin. Autosomes 1–3, 5, 6 8, 10 and the X chromosome are metacentric and 4, 7 and 9 are submetacentric. The X chromosome is almost as large as autosome 1 and the y is very small, about a sixth the length of the X. One replicate of autosome 2 has the shorter arm expanded and the short arms of autosome 8 look as though as though they have small secondary constrictions. The C-banded karyotype (Fig. 9e) shows moderate centromeric C-bands on the larger chromosomes (1–5) and the X chromosome, but the chromosomes are too condensed for the banding of the smaller ones to be established.

H. similis Kuwert, 1887 (Fig. 9e). Another decomposed preparation, this time viewed under phase contrast. The karyotype seems very like that of *H. kirgisicus*, but with a shorter X chromosome.

H. griseus Herbst, 1793 (Fig. 9f). A very distinctive karyotype with autosomes 1 and 2 metacentric and all the others, as well as the X chromosome, acrocentric, with the X about as long as autosome 6. The y chromosome, almost dot-like, is about a third of the length of the X, and there is a similarly small B-chromosome in this individual.

Figure 9. a–p Subgenus *Rhopalohelophorus*, various **a** *H. fulgidicollis*, ♂, England, Hampshire, Lymington, embryo, trypsin-treated, Giemsa-stained **b** *H. asturiensis*, ♂, France, Sarthe, Beaumont-sur-Sarthe, embryo, trypsin-treated, Giemsa-stained **c, d** *H. kirgisicus*, ♂, Russia, West Siberia, Karasuk, embryos **c** Giemsa-stained but partially decomposed in polymerising resin **d** C-banded **e** *H. similis*, ♂, Russia, West Siberia, Karasuk, embryo, Giemsa-stained but partially decomposed, phase-contrast **f** *H. griseus*, ♂, Sweden, Öland, embryo, Giemsa-stained **g** *H. granularis*, ♂, France, Sarthe, Beaumont-sur-Sarthe, embryo, Giemsa-stained **h** *H. discrepans*, ♀, Spain, Pyrenees, embryo, Giemsa-stained **i, j** *H. jocoteroi*, ♂, mid gut cells from the same paratype, Province of La Coruña, Esclavitud, Giemsa-stained **k** *H. strigifrons*, ♂, France, Indre, Scoury, embryo, Giemsa-stained **l** *H. asperatus*, ♂, France, Sarthe, Beaumont-sur-Sarthe, embryo, Giemsa-stained **m, n** *H. pumilio*, Netherlands, Druten, embryos, Giemsa-stained **m** ♂ **n** ♀ **o** *H. croaticus*, ♂, Netherlands, Druten, embryo, Giemsa-stained **p** *H. cincticollis*, ♂, Morocco, Fes, embryo, Giemsa-stained. Scale bar: 15 μm.

H. granularis (Linnaeus, 1760) (Fig. 9g). As published by Angus (1989) the X chromosome was one of the longer acrocentrics (as in *H. griseus*, Fig. 9f) and autosome 8 was regarded as polymorphic for a pericentric inversion. Here a different arrangement, suggested by an anonymous referee, is adopted. This places the single metacentric as the X chromosome and autosomes 3–10 as acrocentrics, as in *H. griseus*. This should be checked using fresh material, especially females, but it is adopted here, not least because it makes fewer assumptions.

H. discrepans Rey, 1885 (Fig. 9h). No male karyotype is available, so the X chromosome cannot be identified. Chromosomes 1–8 are more or less metacentric, 9–11 acrocentric to subacrocentric.

H. jocoteroi Angus et Diaz Pazos, 1991 (Fig. 9i, j). Mid gut preparations from a single male. A karyotype of 12 pairs of chromosomes, including two presumed B-chromosomes. The karyotype shown in Fig. 9j appears to be complete, while that in k is incomplete but shows the shapes of some of the chromosomes more clearly. Autosomes 1, 2 and 9 are metacentric, 4 and 5 are submetacentric and the others, as well as the X chromosome, are acrocentric to subacrocentric. The X chromosome is about the same size as autosome 7 and the y is a dot. The smallest chromosomes, presumed to be Bs are about the size of the X chromosome, though less substantial, and appear to be acrocentric.

H. strigifrons Thomson, 1868 (Fig. 9k). Autosomes 1–7 are metacentric, 5, 6, 8 and 9 are submetacentric, and 10 and the X chromosome are subacrocentric. The y is a dot.

H. asperatus Rey, 1885 (Fig. 9l). The configuration of the karyotype resembles that of *H. strigifrons*, many of the autosomes giving the impression of having very large C-bands.

H. pumilio Erichson, 1837 (Fig. 9m, n). Autosomes 1–4 are metacentric, 5–9 are submetacentric, and 10 and the X chromosome are subacrocentric. The X chromosome is about as long as autosome 7, and the almost dot-like y is about a quarter of the length of the X.

H. croaticus Kuwert, 1886 (Fig. 9o). Autosomes 1–4 and 6 are metacentric, 5–8 are submetacentric, and 9, 10 and the X chromosome are subacrocentric. The y chromosome is a dot and the X is slightly smaller that autosome 10.

H. cincticollis Guillebeau, 1893 (Fig. 9p). Autosomes 1, 3, 4 and 6 are metacentric, 5, 7, 8 and 10 and the X chromosome are submetacentric, and 9 is subacrocentric. The X chromosome is about the same size as autosome 10 and the y is a dot.

Rhopalohelophorus, the *H. flavipes* Fabricius, 1792 group, and *H. browni* McCorkle, 1970
Fig. 10a–o

The *H. flavipes* group are mainly dark coloured species, lacking yellow margins to the pronotum. *H. flavipes* and *H. obscurus* Mulsant, 1844 are two of the most widely distributed species in Europe.

Figure 10. a–o Subgenus *Rhopalohelophorus*, mainly *H. flavipes* group **a–c** *H. flavipes* **a** ♂, England, Hampshire, New Forest, embryo, Geimsa-stained **b** ♂, Spain, Province of Madrid, Peña Lara, embryo, Giemsa-stained **c** ♀, Sweden, mid gut, Giemsa-stained **d–g** *H. obscurus* **d** ♂, Öland, embryo, Giemsa-stained **e** ♀, England, Surrey, Chobham Common, embryo, Giemsa-stained **f** ♂, France, Corsica, Ajaccio, mid gut, Giemsa-stained **g** ♂, Crete, Rethymnon, embryo, Giemsa-stained **h, i** *H. algiricus* ♂, Morocco, Ifrane, mid gut, Giemsa-stained **j, k** *H. subarcuatus* ♂, Italy, Sardinia, Mandas, mid gut, Giemsa-stained **l, m** *H. seidlitzi*, mid gut, Spain **l** ♀, Province of Segovia, Cuéllar, embryo, Giemsa-stained **m** ♂, Province of León, Algadefe, mid gut, Giemsa-stained **n, o** *H. browni* ♂, China, Heilongjiang, Qitaihe, mid gut **n** Giemsa-stained **o** the same nucleus C-banded. The position of missing chromosomes is indicated by a small black disc. Scale bar: 15 μm.

H. flavipes (Fig. 10a–c). Autosomes 1–8 are metacentric, 9 is metacentric to submetacentric, 10 is acrocentric and the X chromosome, about two thirds the length of autosome 10, is subacrocentric. The y is a dot.

H. obscurus (Fig. 10d–g). All the autosomes, and the X-chromosome, are clearly biarmed, metacentric (pairs 1–4 and the X chromosome) or metacentric to submetacentric (pairs 5–10). The X chromosome is about as long as pair 10, and the y is a dot.

H. algiricus Motschulsky, 1860 (Fig. 10h, i). This species closely resembles *H. obscurus* but differs in minor aedeagal differences and in the smaller larval head. Chromosomally the only clear difference is in autosome 9, which is subacrocentric.

H. subarcuatus Rey, 1885 (Fig. 10j, k). Endemic to Corsica and Sardinia, described by Rey from Corsica but very scarce there and much commoner on Sardinia. The karyotype is similar to that of *H. algiricus* but the X chromosome is clearly not metacentric, and autosome pair 9 is more nearly metacentric.

H. seidlitzi Kuwert, 1885 (Fig. 10l, m). Endemic to Spain and Portugal where its range overlaps with those of *H. flavipes* and, in the north, *H. obscurus*. The karyotype is very similar to that of *H. flavipes*, the most obvious difference being the subacrocentric autosome 7, which is metacentric in *H. flavipes*.

H. browni McCorkle, 1970 ex Angus, 1970b (Fig. 10n, o). This Holarctic species was originally described from tundra in the Canadian Northwest Territories (Mackenzie delta) and Yukon, and Alaska. It is widespread and common in the Baikal area of East Siberia and in central Yakutia and extends to the Russian Far East (Primorye). It is scarce in Mongolia and in China is known from Nei Mongol and Heilongjiang. Angus (2019) refers to variation of the aedeagal strut length in *H. browni*, but further (as yet unpublished) data indicate that this variation is more or less random and continuous, and thus not a concern in attributing the karyotype. Autosomes 1–6 and the X chromosome are metacentric, 7 and 8 are subacrocentric and 9 and 10 are acrocentric. The X chromosome is about the same size as autosome 10. The y is a dot. Autosome 1 is markedly longer than pair 2, while autosomes 2–6 show a smaller and more even decrease in length. C-banding (Fig. 10p) shows centromeric C-bands on all the chromosomes (except the y), those on autosome 1 being particularly small.

Triploids and parthenogenesis
Fig. 11a–g

Within the Helophoridae, parthenogenesis was recorded by Angus (1970c) in Canadian *H. orientalis* Motschulsky, 1860, who established its existence by rearing females for two generations in the laboratory. No chromosome data were available. The first chromosomally proven parthenogenesis was by Angus (1992b) who found triploid female *H. brevipalpis* in the Spanish province of León, accompanied by diploids of both sexes. Fig. 11, a shows a karyotype from a parthenogenetic triploid female. The chromosomes match up in triplets without any difficulty, though it may be noted that in triplets 1 and 7 there is a progressive size decrease in the three replicates and triplets 6 and 8 each have one replicate shorter than the other two.

Fig. 11b, c shows a Giemsa-stained and C-banded karyotype from a triploid female taken by Angus at Ligonichio, Regio Emilia, Italy in 2018. Two chromosomes have been lost from this preparation, shown as missing from triplets 2 and 6. The variation in chromosome length within triplets is less than in the Spanish material, but triplets 6 and 9 appear to have one longer replicate, and triplet 7 one shorter one. The apparently shorter replicate in triplet 10 is clearly the result of the short arm not being extended. A triploid nucleus from a single female from Ponte Scipione (Parma Prov., Italy) (Fig. 11d) shows triplet 1 with a similar gradation in replicate length to that shown by the Spanish karyotype shown in Fig. 11a, and triplets 6 and 8 each have one replicate longer than the others.

Fig. 11e, f shows a karyotype from mid gut of a female *H. orientalis* from Mishan, Heilongjiang, China, Giemsa-stained and C-banded. Triplet 1 shows a gradation in replicate lengths, as in Spanish *H. brevipalpis*, but in the other triplets the replicates are more or less equal in length. It is not possible to identify the X chromosome in these *H. orientalis* preparations as in that species males are known only in the American Rockies, and from one locality near Vladivostok in Russia (Angus 1992a).

The question arises is whether these variations in replicate length within triplets result from slight random variation in rates of chromosome condensation through prophase and into metaphase of mitosis, or whether they result from a hybrid origin of these triploids (allotriploidy), which Simon et al. (2003) report as widespread in invertebrates, including, among insects, some Coleoptera, Phasmatodea and Orthoptera. The problem here is finding candidate species which might be involved in hybrid formation. *H. brevipalpis* is intriguing in this context. Angus (1985,Figs 50–56) illustrated variation in the aedeagus size of populations of *H. brevipalpis*, with specimens from northern France (the lectotype, Fig. 50) and Crete (Fig. 51) having relatively smaller aedeagi, while some, including material from the Shetland Islands (Fig. 54, *H. bulbipalpis* Kuwert, lectotype) and Khorasan, Iran (Fig. 55), (now *H. brevipalpis levantinus* Angus, 1988), have them larger. Two of Rey's names, *H. mixtus* (Fig. 52), with a smaller aedeagus and *H. insignis* (Fig. 53), with a larger one, both refer to material from Provence (southern France). There is thus appreciable variation within *H. brevipalpis*, which might indicate hitherto undetected cryptic species. It is also worth noting that the Spanish León region where triploids were discovered, is on the edge of the species' range (Millán et al. 2014).

The case of *H. orientalis* is intractable in view of the very limited distributions of bisexual populations (Angus 1992a).

One occurrence which is relevant is the chance occurrence of a triploid embryo among batches of developing eggs obtained from a female *H. aequalis* brought back to the laboratory from St Flour (Cantal), France in 1987. Fig. 11g. shows this karyotype. Triplets 1, 5, 6 and 7 each have one replicate shorter than the others. There appears to be no possibility that this is of hybrid origin, and in particular, there is no other known species of *Helophorus* s. str. with chromosomes sufficiently similar to those of *H. aequalis* to be able to produce such a convincing triploid karyotype. It is worth noting that, in the course of Ph.D. research supervised by Angus, F. Shaarawi obtained a triploid embryo from *Hydrochus elongatus* (Schaller, 1783) (Shaarawi and Angus 1992).

Figure 11. a–g Triploid females, Giemsa-stained **a–d** *H. brevipalpis* **a** Spain, Province of León, Algadefe **b, c** Italy, Sologno **b** Giemsa-stained **c** the same nucleus C-banded **d** Italy, Ponte Scipione **e, f** *H. orientalis*, China, Heilongjiang, Qitahe **e** Giemsa-stained **f** the same nucleus C-banded. No male *H. orientalis* was available so the X chromosome cannot be identified **g** *H. aequalis*, a solitary triploid embryo found among numerous normal diploids from egg cocoons from France, Cantal, St Flour. The positions of missing chromosomes are indicated by small black discs. Scale bar: 15 μm.
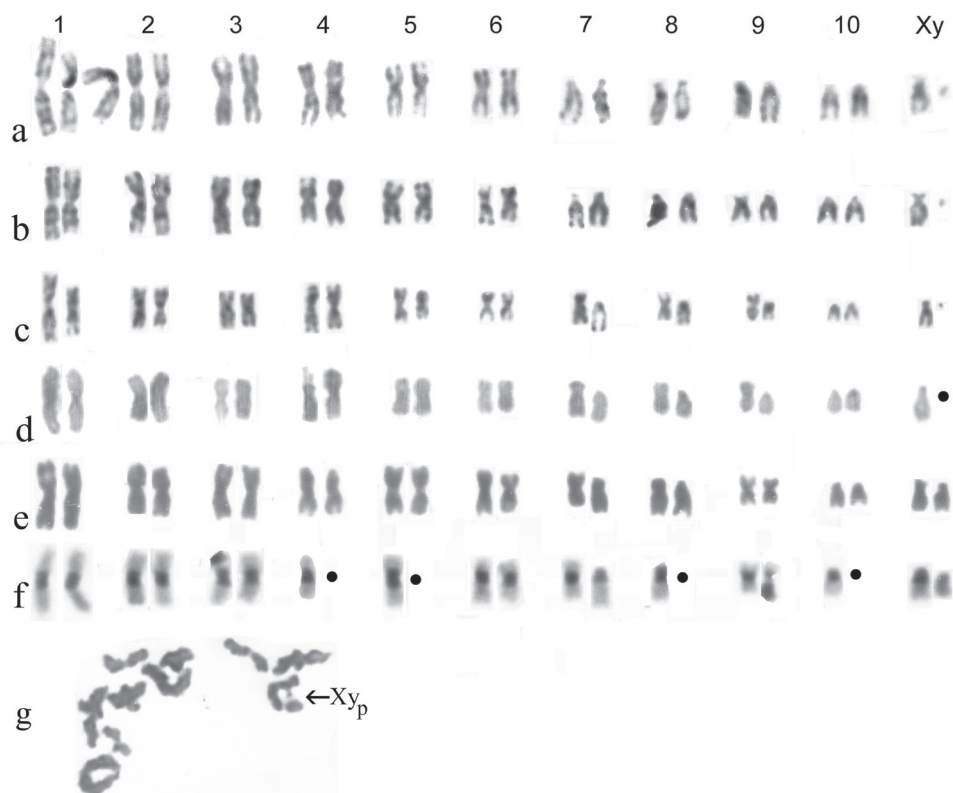
Experimental hybrids
Fig. 12a–f

♂ hybrid, *H. lapponicus* ♀ lab-reared from Karasuk X ♂ *H. paraminutus*, from Karasuk (Fig. 12a, b). This cross was originally undertaken with a view to obtaining karyotypes in which the condensation of the chromosomes through prophase into metaphase of mitosis was completely synchronised, to see if any minor differences could be found

between the apparently identical karyotypes of the parent species. In fact this was not at all what happened. Both the hybrid karyotypes show serious irregularities in chromosome condensation. Autosome pair 1 shows serious differences in the lengths of the replicates, as does pair 4 in Fig. 12, a, and pair 6 in both karyotypes. The smaller chromosomes perhaps have less scope for showing irregularities, but pair 8 in Fig 12b and pair 9 in Fig. 12 a both show obvious differences. Angus (1986) suggested that there might be differences in the ease with which paternal chromosomes could incorporate non-histone proteins from a predominantly maternal cytoplasm. But now I am less convinced as by the time the embryos were sufficiently developed for chromosome preparations to be made, the cytoplasm would be of hybrid origin. Some of these hybrids were reared through to adulthood and were apparently able to produce functional meiosis with no failures of chromosomes pairing up during prophase (Fig. 12g).

♂ hybrid, *H. minutus* ♀ lab-reared from Egham, Surrey X ♂ *H. paraminutus*, wild-caught, Austria (Fig. 12), c Giemsa-stained, d phase-contrast. As with the *H. lapponicus* X *paraminutus* cross, there is a mismatch of the replicates of chromosome 1 and no *H. paraminutus* chromosome matches the NOR-bearing *H. minutus* chromosome 4. Angus (1986) shuffled the order of the *H. paraminutus* chromosomes 7–9, to get a better match with those of *H. minutus*. Partly this was a by-product of the reversal of *H. minutus* chromosomes 8 and 9 used in that paper, but with the chromosomes now being arranged on-screen with Photoshop the standard arrangement of the *H. paraminutus* karyotype is used. Chromosomes 7–9 of the two species differ by a pericentric inversion, 7 metacentric in *H. minutus*, acrocentric in *paraminutus*, 8 subacrocentric in *H. minutus*, acrocentric in *paraminutus*, and 9 metacentric in *H. minutus*, acrocentric, possibly with a terminal NOR in the short arm in *paraminutus*. It should at this stage be stressed that this represents the minimum number of differences between homologous chromosomes of the two species. The true extent of the differences may be greater!

♀ Giemsa-stained (Fig. 12e) and ♂ C-banded (Fig. 2f) hybrid embryos, *H. minutus* ♀ lab-reared from Egham, Surrey X ♂ *H. calpensis*, wild-caught, Tarifa, Spain. The Giemsa-stained female karyotype shows no obvious mismatches in autosomes 1–6, with the NOR-bearing chromosome placed as pair 4 in both species. In pair 9 the metacentric *H. minutus* chromosome is longer than the *H. calpensis* submetacentric one. Pairs 7 and 8 show the expected differences in centromere position, but their sizes match quite well. Autosome 10 and the X chromosome both match well. Turning to the incomplete C-banded ♂ karyotype (Fig. 12f), autosome 1 pairs up well, the tip of 1 replicate of autosome 3 lies over another chromosome (perhaps the one shown as *calpensis* pair 9), pairs 4, 5, 8 and 10 are shown as represented by one chromosome each and the longer metacentric *H. calpensis* autosome 9 is suggested to be the one which overlapped something else, possibly the tip of the *H. minutus* autosome 3. The acrocentric and largely heterochromatic y chromosome of *H. calpensis* is clearly recognisable. In general, there is a good overall resemblance in the sequences of sizes of the chromosomes of the two species, but the differences between them may be greater than this suggests.

Figure 12. a–f experimental Hybrids **a, b** ♂ hybrid embryos, *H. lapponicus* ♀ lab-reared from Karasuk X ♂ *H. paraminutus*, from Karasuk, with the shorter replicate of autosome 1 shown in its natural position (on the right) and "straightened" (centre) **c, d** ♂ hybrid embryos, *H. minutus* ♀, lab-reared from Egham, Surrey X ♂, *H. paraminutus*, wild-caught, Austria **e, f** hybrid embryos **e** ♂ **f** ♀, *H. minutus* ♀, lab-reared from Egham, Surrey X ♂, *H. calpensis*, wild-caught, Tarifa, Spain. The suggested positions of missing chromosomes are indicated by small black discs **g** Meiosis, first metaphase from a ♀ *H. lapponicus* X ♂*H. paraminutus* hybrid, showing 10 bivalents + Xy$_p$ sex chromosomes (labelled). Scale bar: 15 μm.

General comments

Helophorus chromosomes show useful interspecies variation which is very helpful in delimiting species. They show variation in the size and extent of the C-bands and the distribution of NORs. Where this has been investigated, they show extensive rather fine-grained and fairly uniform chromomeric banding, perhaps equivalent of G-banding. It can be useful in showing where translocations have occurred but this is difficult to demonstrate convincingly, and from the point of view of cytotaxonomy, probably not worth the effort. The obvious polymorphisms encountered result from pericentric inversions, with acrocentric and metacentric versions of the chromosomes involved, and from interpolated heterochromatin (C-bands) into chromosome arms, as in the long variant of the *H. grandis* X chromosome (Fig. 2i, j).

One notable feature of *Helophorus* karyotypes is the frequent occurrence of a particular type of X chromosome–usually not quite the smallest in the nucleus, and sub-acrocentric to submetacentric. Angus (1989) referred to this as the *H. minutus* pattern of X chromosome. It is shown by some members of the subgenus *Helophorus* s. str.–*H. thauma*, *aequalis*, *grandis* (short form) (Fig. 2e–i) and *H. hammondi* (Fig. 4c, d). Both species of subgenus *Eutrichelophorus* have it, *H. micans* (Fig. 4l, m) and *H. oxygonus* (Fig. 4n). In the subgenus *Rhopalohelophorus* it widespread, occurring in the three *H. leontis* group species (Fig. 6g–j), and *H. nevadensis* (Fig. 6k) in the *Atractohelophorus* group.

Among the other *Rhopalohelophorus* it occurs in *H. nanus* (Fig. 7a–c), *H. pallidipennis* (Fig. 7g, h), *H. minutus* Fig. 8a–c), *H. atlantis* (Fig. 8d, e), *H. calpensis* (Fig. 8f–j), *H. paraminutus* (Fig. 8k, l), *H. lapponicus* (Fig. 8m–p), *H. griseus* (Fig. 9f), *H. jocoteroi* (Fig. 9j, k), *H. strigifrons* (Fig. 9l), *H. asperatus* (Fig. 9m), *H. pumilio* (Fig. 9n, o), *H. croaticus* (Fig. 9p), *H. flavipes* (Fig. 10a–c) (but not in *H. obscurus* and *H. algiricus*, Fig. 10d–i), *H. subarcuatus* (Fig. 10j, k), *H. seidlitzi* (Fig. 10l–n) and *H. browni* (Fig. 10o, p). Other forms of X-chromosomes tend to be metacentric, sometimes a bit larger, but may actually be the longest in the nucleus, as in *H.* (*Gephelophorus*) *sibiricus* (Fig. 4h, i), *H. glacialis* (Fig. 6e, f) and *H. redtenbacheri* (Fig. 7, d).

Chromosome polymorphisms include pericentric inversions, apparently relatively unusual in Coleoptera but present in *Melolontha melolontha* Linnaeus, 1758 (Scarabaeidae) where one autosome is polymorphic for an inversion, resulting in both metacentric and acrocentric forms. In a secod autosome pair pericentric inversion is suggested as the cause of its departure from the ancestral dinastine metacentric arrangement to its present acrocentric form (Giannoulis et al. 2011).

B-chromosomes may be present, normally small and sometimes difficult to distinguish from the y chromosome, as in *H. nevadensis* (Fig. 6k) and *H. griseus* (Fig. 9g), but may be rather larger, as in *H. jocoteroi* (Fig. 9j).

Parthenogenesis is apparently rare and is currently known in only two species, *H. brevipalpis* (Fig. 11a–d) and *H. orientalis* (Fig. 11e, f), and, as far as is known, is always associated with triploidy.

This contrasts with the situation in *Anacaena lutescens* (Stephens, 1829) (Hydrophilidae) where diploid parthenogenetic females, in populations where males are unknown are always heterozygous for deletion of a small distal portion, beyond a secondary constriction, of autosome pair 8. In some of these populations there are also triploids and these show variation indicating that the triploidy has arisen on separate occasions, after the development of parthenogenesis (Shaarawi and Angus 1991).

The data reported here are summarised in Table 1 which gives the earliest references for material already published. The general karyotype formulae are given for each subgenus and only variations are listed (where they occur) for individual species. The sex chromosomes are listed as Xy. In cases where meiosis is known, it is Xy_p, the usual polyphagan arrangement, and I have seen nothing to suggest any deviations.

Table 1. Summary of the data.

Subgenus/species	Karyotype/peculiarities	Reference
Helophorus s. str.	$2n = 16 + Xy_p$	
H. aquaticus		Angus 1982
H. thauma		Angus and Toledo 2010
H. aequalis		Angus 1982
H. grandis	Autosome 5 metacentric or acrocentric, polymorphic for a pericentric inversion. X chromosome with a length polymorphism associated with an interstitial C-band. 1 or 2 B-chromosomes	Angus 1983
H. liguricus		Angus 1989
H. maritimus		Angus 1983
H. occidentalis		Angus 1983
H. milleri	Autosome 8 metacentric or acrocentric, polymorphic for a pericentric inversion.	Angus 1989
H. syriacus		Angus 1989
H. oscillator		Angus 1989
H. hammondi		Angus 2015
H. jaechi		This paper
H. (*Gephelophorus*)	$2n = 16 + Xy$	
H. auriculatus		Angus 2015
H. sibiricus	Autosome 7 with a length polymorphism associated with interstitial heterochromatin	Angus 2019; This paper
H. (*Eutrichelophorus*)	$2n = 16 + Xy$	
H. micans		Angus 2015
H. oxygonus		Angus 2015
H. (*Empleurus*)	$2n = 20 + Xy$	
H. nubilus		Angus 2015
H. rufipes		Angus 2015
H. (*Trichohelophorus*)	$2n = 20 + Xy$	
H. alternans		Angus 1989
H. (*Lihelophorus*)	$2n = 20 + Xy$	
H. lamicola		Angus et al. 2916
H. ser		Angus et al. 2016
H. yangae		Angus et al. 2016
H. (*Rhopalohelophorus*)	$2n = 20 + Xy$	
H. brevipalpis	Diploid: Autosome 5 metacentric or acrocentric, polymorphic for a pericentric inversion. Triploid ♀♀: $3n = 30 + 3X$	Angus 1992b
H. montenegrinus		This paper
H. glacialis		This paper
H. leontis		Angus et al. 2005
H. dixoni		Angus et al. 2005
H. biltoni		Angus et al. 2005
H. nevadensis	B-chromosomes	This paper
H. korotyaevi		This paper
H. lewisi		This paper
H. nanus		Angus 1989, 2015
H. redtenbacheri		Angus 1989
H. pallidus		Angus 1989
H. villosus		Angus 1989
H. pallidipennis		Angus 1998
H. kervillei		Angus 1998
H. minutus		Angus 1986
H. atlantis		Angus and Aouad 2009
H. calpensis		Angus 1988
H. paraminutus		Angus 1986

Subgenus/species	Karyotype/peculiarities	Reference
H. lapponicus		Angus 1986
H. fulgidicollis		Angus 1989
H. asturiensis		Angus 1989
H. kirgisicus		Angus 1989
H. similis		Angus 1989
H. griseus		Angus 1989
H. granularis		Angus 1989
H. jocoteroi	2 B-chromosomes	Angus and Diaz Pazos 1990
H. strigifrons		Angus 1989
H. asperatus		Angus 1989
H. pumilio		This paper
H. croaticus		This paper
H. cincticollis		This paper
H. flavipes		Angus 1989, 1996
H. obscurus		Angus 1989, 1996
H. algiricus		Angus 1996
H. subarcuatus		Angus 1996
H. seidlitzi		Angus 1989, 1996
H. browni		Angus 2019
H. orientalis		Angus and Jia 2020
Hybrids		
♀*H. lapponicus* X ♂*H. paraminutus*	2n = 20 + Xy	Angus 1986
♀*H. minutus* X ♂*H. paraminutus*	2n = 20 + Xy	Angus 1986
♀*H. minutus* X ♂ *H.calpensis*	2n = 20 + Xy	Angus 1988

Acknowledgements

meticulous care in going through the manuscript and for insisting that I provided improved versions of some of the original illustrations. And finally, my ongoing research is made possible by my position as a Scientific Associate at the Natural History Museum in London. I thank them for this and the use of the research equipment involved.

References

Angus RB (1970a) A revision of the beetles of the Genus *Helophorus* F. (Coleoptera: Hydrophilidae) subgenera *Orpheloporus* d'Orchymont, *Gephelophorus* Sharp and *Meghelophorus* Kuwert. Acta Zoologica Fennica 129: 1–62

Angus RB (1970b) Revisional studies on East Palaearctic and some Nearctic species of *Helophorus* F. (Coleoptera: Hydrophilidae). Ergebnisse der Zoologischen Forschungen von Dr. Z. Kaszab in der Mongolei (No. 226). Acta Zoologica Academiae Scientarum Hugaricae 16: 249–290.

Angus RB (1970c) *Helophorus orientalis* (Coleoptera: Hydrophilidae), a parthenogenetic water beetle from Siberia and North America, and a British Pleistocene fossil. Canadian Entomologist 102: 129–143. https://doi.org/10.4039/Ent102129-2

Angus RB (1982) Separation of two species standing as *Helophorus aquaticus* (L.) (Coleoptera, Hydrophilidae) by banded chromosome analysis. Systematic Entomology 7: 265–281. https://doi.org/10.1111/j.1365-3113.1982.tb00444.x

Angus RB (1983) Separation of *Helophorus grandis*, *maritimus* and *occidentalis* sp. n. (Coleoptera, Hydrophilidae) by banded chromosomal analysis. Systematic Entomology 8: 1–13. https://doi.org/10.1111/j.1365-3113.1983.tb00462.x

Angus RB (1985) K Revizii Palearkticheskikh Vodolyubov roda *Helophorus* F. (Coleoptera, Hydrophilidae). 2. Entomologicheskoye Obozrenie 4: 716–747 . English version: Towards a Revision of the Palaearctic Species of *Helophorus* F. (Coleoptera, Hydrophilidae). 2. Entomological Review 4: 128–162.

Angus RB (1989) Towards an atlas of *Helophorus* chromosomes. Balfour-Browne Club Newsletter 44: 13–22.

Angus, RB (1986) Revision of the Palaearctic species of the *Helophorus minutus* group (Coleoptera: Hydrophilidae), with chromosome analysis and hybridization experiments. Systematic Entomology 11: 133–163. https://doi.org/10.1111/j.1365-3113.1986.tb00173.x

Angus RB (1992a) Insecta Coleoptera Hydrophilidae Helophorinae. Süßwasserfauna von Mitteleuropa 20/10-2. Jena, 144 pp.

Angus RB (1992b) A chromosomal investigation of *Helophorus brevipalpis* Bedel (Coleoptera: Hydrophilidae), with triploid Spanish females a possible source of American parthenogenetic material. The Entomologist 111: 56–60.

Angus RB (1996) A re-evaluation of the *Helophorus flavipes* group of species (Coleoptera, Hydrophiloidea), based on chromosomal analysis, larvae and biology. Nouvelle Revue d'Entomologie (n. s.) 13: 111–122.

Angus RB (1998) *Helophorus pallidipennis* Mulsant & Wachanru and *H. kervillei* d'Orchymont as good species. (Coleoptera: Helophoridae). Koleopteroplogische Rundschau 68: 189–196.

Angus RB (2015) Some *Helophorus* karyotypes. Latissimus 36: 2–3.

Angus RB (2019) Once more in the Middle Kingdom! Latissimus 44: 1–6.

Angus RB, Diaz Pazos JA (1990) *Helophorus jocoteroi* n. sp., from Northwest Spain. Nouvelle Revue d'Entomologie (N. S.) 7: 419–422.

Angus RB, Jia F, Chen Z-N, Zhang Y, Vondráček D, Fikáček M (2016) Taxonomy, larval morphology and cytogenetics of *Lihelophorus*, the Tibetan endemic subgenus of *Helophorus* (Coleoptera: Hydrophiloidea). Acta Entomologica Musei Nationalis Pragae 56(1): 109–148.

Angus RB, Jia F (2020) Triploidy in Chinese parthenogenetic *Helophorus orientalis* Motschulsky, 1860, further data on parthenogenetic *H. brevipalpis* Bedel, 1891 and a brief discussion of parthenogenesis in Hydrophiloidea (Coleoptra). Comparative Cytogenetics 14(1): 1–10. https://doi.org/10.3897/CompCytogen.v14i1.47656

Angus RB, Litovkin SV, Jia F (2019) Notes on *Helophorus* (s. str.) *kozlovi* Zaitzev, 1908, with description of two new species, re-evaluation of *Helophorus* s. str. Fabricius 1775 and *Trichohelophorus* Kuwert, 1886, and revised keys to the subgenera of *Helophorus* and to the species of *Helophorus* s. str. (Coleoptera: Helophoridae). Koleopterologiche Rundschau 89: 127–150.

Angus RB, Mahdizadeh S, Hosseinie SO (2005) A re-evaluation of the *Helophorus leontis* complex (Coleoptera: Helophoridae) based on chromosomal analysis, with description of *H. biltoni* sp. nov. from Iran. Aquatic Insects 27: 193–198. https://doi.org/10.1080/01650420500170510

Bigger TRL (1975) Karyotypes of some Lepidoptera chromosomes and changes in their holokinetic organization as revealed by new cytological techniques. Caryologia 40: 713–726. https://doi.org/10.1508/cytologia.40.713

Crozier RH (1968) An acetic acid dissociation, air-drying technique for insect chromosomes, with aceto- lactic orcein staining. Stain Technology 43: 171–173. https://doi.org/10.3109/10520296809115063

Dutrillaux AM, Moulin S, Dutrillaux B (2006) Use of meiotic pachytene stage of spermatocytes for karyotypic studies in insects. Chromosome Research 14: 549–557. https://doi.org/10.1007/s10577-006-1052-7

Dutrillaux B, Dutrillaux A-M, McClure M, Elias M, Bed'hom B (2022) Improved basic cytogenetics challenges holocentricity of butterfly chromosomes. Cytogenetic and Genome Research 162(5): 262–272. https://doi.org/10.1159/000526034

Giannoulis T, Dutrillaux A-M, Mamuris Z, Montreuil O, Stamatis C, Dutrillaux B (2011) Evolution of European Cockchafers (Melolonthinae: Scarabaeidae: Coleoptera:): a morphological, molecular and chromosomal study of intra- and inter-specific variations. Bulletin of Entomological Research 101: 345–352. https://doi.org/10.1017/S0007485310000568

John B, Lewis KR (1960) Nucleolar Controlled segregation of the Sex Chromosomes in Beetles. Heredity: 431–439. https://doi.org/10.1038/hdy.1960.107

Juan C, Pons J, Petitpierre E (1993) Localisation of tandemly repeated DNA sequences in beetle chromosomes by fluorescent *in situ* hybridization. Chromosome Research 1: 167–176. https://doi.org/10.1007/BF00710770

Luciani JM, Morazzini MR, Stahl A (1975) Identification of pachytene bivalents in human male meiosis using G-banding technique. Chromosoma 52: 275–282. https://doi.org/10.1007/BF00332116

Maudlin I (1974) Giemsa banding of metaphase chromosomes in triatomine bugs. Nature 252: 392–393. https://doi.org/10.1038/252392a0

Millán A, Sánchez-Fernández D, Abellán P, Picazo F, Carbonell JA, Lobo JM, Ribera I (2014) Atlas de los Coleopteros Acuáticos de España Peninsular. Madrid, 819 pp.

O'Brien SJ, Nash WG, Wildt, DE, Bush ME, Beneveniste RE (1985) A molecular solution to the riddle of the giant panda's phylogeny. Nature 317: 140–144. https://doi.org/10.1038/317140a0

Pearson PL (1997) Banding Patterns, Chromosome Polymorphism, and Primate Evolution. Chapter 8. In: Yunis JJ (Ed.) Molecular Structure of Human Chromosomes, New York, 267–293. [326 pp.] https://doi.org/10.1016/B978-0-12-775168-9.50013-0

Rønne M (1977) In vitro induction of G bands with cycloheximide. Hereditas 86: 107–110. https://doi.org/10.1111/j.1601-5223.1977.tb01217.x

Rønne M, Andersen O (1978) Effect of 5-fluorouracil and 5-fluorouridine on metaphase structure in human lymphoid cells. Hereditas 88: 127–130. https://doi.org/10.1111/j.1601-5223.1978.tb01612.x

Rothfels KH, Siminovitch L (1958) An air-drying technique for flattening chromosomes in mammalian cells grown *in vitro*. Stain Technology 13(2): 73–77. https://doi.org/10.3109/10520295809111827

Shaarawi FAI, Angus RB (1992) Chromosomal analysis of some European species of the genera *Georissus* Latreille, *Spercheus* Illiger and *Hydrochus* Leach (Coleoptera: Hydrophiloidea). Koleopterologische Rundschau 62: 127–135.

Shaarawi FAI, Angus RB (1991) A chromosomal investigation of five European species of *Anacaena* Thomson (Coleoptera, Hydrophilidae). Entomoligica Scandinavica 21: 415–426. https://doi.org/10.1163/187631290X00319

Simon J-C, Delmotte F, Rispe C, Crease T (2003) Phylogenetic relationships between parthenogens and their sexual relatives: the possible routes to parthenogenesis in animals. Biological Journal of the Linnean Society 79: 151–163. https://doi.org/10.1046/j.1095-8312.2003.00175.x

Smith SG (1960) Chromosome numbers of Coleoptera, 2. Canadian Journal of Genetics and Cytology 2: 66–68. https://doi.org/10.1139/g60-007

ORCID

Robert B. Angus https://orcid.org/0000-0002-3860-5617

RESEARCH ARTICLE

COMPARATIVE
Cytogenetics
A peer-reviewed open-access journal
International Journal of Plant & Animal Cytogenetics,
Karyosystematics, and Molecular Systematics

Chromosomes of the genus *Arge* Schrank, 1802 (Hymenoptera, Argidae): new data and review

Vladimir E. Gokhman[1]

1 *Botanical Garden, Moscow State University, Moscow 119234, Russia*

Corresponding author: Vladimir E. Gokhman (vegokhman@hotmail.com)

Citation: Gokhman VE (2023) Chromosomes of the genus *Arge* Schrank, 1802 (Hymenoptera, Argidae): new data and review. Comparative Cytogenetics 17: 327–333. https://doi.org/10.3897/compcytogen.17.115485

Abstract

Results of the chromosome study of 12 sawfly species of the genus *Arge* Schrank, 1802 are reviewed, including new data on the karyotypes of *A. ciliaris* (Linnaeus, 1767) and *A. enodis* (Linnaeus, 1767) with n = 10. Moreover, the same chromosome number, n = 10, is found in *A. ustulata* (Linnaeus, 1758), for which n = 8 was previously reported. In addition, n = 8 is confirmed in *A. gracilicornis* (Klug, 1814). The results of the morphometric analysis of chromosome sets of these four species are given. In the genus *Arge*, haploid chromosome numbers of n = 8, 10, 11 and 13 were found. Among these sawflies, n = 8 appeared to be the most frequent chromosome number, followed by n = 10. The known data of the chromosome study of these insects are summarized and discussed in the light of phylogeny and taxonomy of the genus *Arge*.

Keywords

Chromosome morphometry, karyotypes, sawflies

Introduction

Arge Schrank, 1802 is the most speciose genus of the family Argidae, which is, in turn, the second largest group of its kind among sawflies (Symphyta) (Taeger et al. 2018). The genus *Arge* currently includes more than 400 described species, with about 180 members of the genus occurring in the Palaearctic (Taeger et al. 2018). To date, chromosomal data for Argidae are known only for ten *Arge* species (Naito 1982; Westendorff 2006). For most of them, certain additional information on the karyotype structure is also available. In the present paper, I have recently examined karyotypes

of several members of this genus, including two newly studied species. In another two species, either the existing chromosome number was confirmed, or, unexpectedly, a different n value was found. Since some members of the genus *Arge* appeared to have superficially similar karyotypes, morphometric analysis of the chromosome sets, which could find some hidden interspecific differences, was also undertaken. The existing results of the chromosome study of the genus *Arge* are summarized and discussed in the light of phylogeny and taxonomy of these sawflies (see below).

Material and methods

Adult female sawflies of the genus *Arge* were collected by the author in the wild, mostly on the flowers of umbelliferous plants (Apiaceae) in Ozhigovo, Moscow, Russia (55°28'N, 36°52'E) in 2022–2023 (Table 1). The sawflies were initially identified by the author, the identifications were then checked by Sergey A. Basov (Zoological Institute, Russian Academy of Sciences, St. Petersburg, Russia). Voucher specimens are deposited in the collection of the Zoological Museum of Moscow State University (Moscow, Russia).

Chromosomal preparations were obtained from embryos forming inside the developing eggs, generally following the protocols used by Naito (1982) and Imai et al. (1988) with a few modifications. Specifically, mature eggs were extracted from adult females and put inside small Petri dishes on a filter paper soaked with distilled water. These eggs were kept for about three days at room temperature. During that time, sawfly embryos developed inside these eggs. These embryos were first dissected in 0.5% hypotonic sodium citrate solution containing 0.005% colchicine, and then transferred to a fresh portion of hypotonic solution and incubated for 30 min at room temperature. After that, the material was transferred onto a pre-cleaned microscope slide using a Pasteur pipette and then gently flushed with Fixative I (glacial acetic acid: absolute ethanol: distilled water 3:3:4). The tissues were disrupted using dissecting needles in an additional drop of Fixative I. Another drop of Fixative II (glacial acetic acid: absolute ethanol 1:1) was applied to the center of the area, and the more aqueous phase was blotted off the edges of the slide. The same procedure was then performed with Fixative III (glacial acetic acid). The slides were dried for approximately half an hour and stored at room temperature. The preparations were stained overnight with a freshly prepared 3% Giemsa solution.

Haploid mitotic divisions were studied and photographed using an optic microscope Zeiss Axioskop 40 FL fitted with a digital camera Axiocam 208 color (Carl Zeiss, Germany). To produce illustrations, the resulting images were handled with image processing programs ZEN version 3.0 (blue edition) and GIMP version 2.10. Chromosomes were measured on ten metaphase plates of all studied species using KaryoType software version 2.0 and then classified according to the guidelines provided by Levan et al. (1964), i.e., as metacentrics (M), submetacentrics (SM), subtelocentrics (ST) and acrocentrics (A).

Results

Arge gracilicornis (Klug, 1814) (n = 8). Seventeen embryos obtained from four females were examined. Most chromosomes are metacentric/submetacentric, but the shortest one is an acrocentric (Fig. 1A, Table 2). The first chromosome is very large, about twice as long as the second one, which is, in turn, more than twice as long as the last acrocentric (Table 1). All chromosomes, except the largest and the smallest, form a continuous gradation in length.

   A. enodis (Linnaeus, 1767) (n = 10). Two embryos obtained from a single female were examined. All chromosomes are obviously biarmed, either metacentric or submetacentric (Fig. 1B, Table 2). However, unlike the karyotype of the previous species, length of the first chromosome only slightly exceeds that of the second one (Table 1). The remaining chromosomes gradually decrease in size.

   A. ciliaris (Linnaeus, 1767) (n = 10). Ten embryos, also obtained from a single female, were examined. As in the previous species, all chromosomes are clearly biarmed, either metacentric or submetacentric (Fig. 1C; Table 2). Similarly to *A. gracilicornis*, the first chromosome is very large, about four times longer than the last one (Table 1).
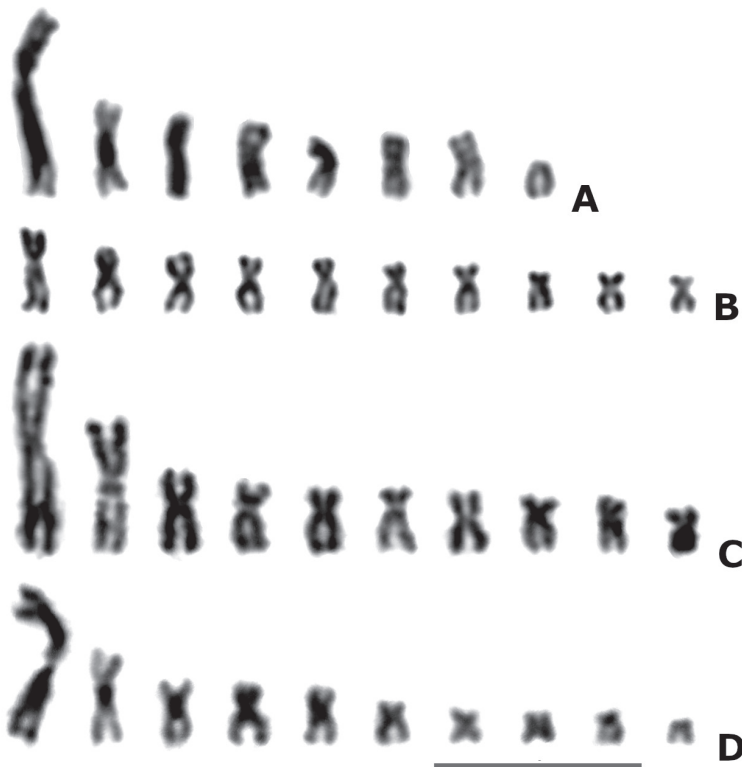
Figure 1. Haploid karyograms of *Arge* species **A** *A. gracilicornis* **B** *A. enodis* **C** *A. ciliaris* **D** *A. ustulata*. Scale bar: 10 μm.

Table 1. Relative lengths (RLs) and centromeric indices (CIs) of chromosomes of four *Arge* species (mean ± SD).

Chromosome no.	*A. gracilicornis*		*A. enodis*		*A. ciliaris*		*A. ustulata*	
	RL	CI	RL	CI	RL	CI	RL	CI
1	27.07 ± 1.32	41.80 ± 3.88	15.04 ± 0.87	40.68 ± 3.37	23.22 ± 0.92	42.08 ± 4.00	25.98 ± 1.34	43.66 ± 4.35
2	13.59 ± 0.54	44.52 ± 4.10	12.68 ± 0.88	41.10 ± 4.45	13.23 ± 2.13	41.45 ± 4.31	14.49 ± 0.82	40.81 ± 4.23
3	12.31 ± 0.60	46.94 ± 1.76	11.39 ± 0.58	41.32 ± 4.80	10.41 ± 0.34	42.60 ± 3.77	11.18 ± 0.41	43.49 ± 3.31
4	11.67 ± 0.46	45.37 ± 4.19	10.66 ± 0.35	40.03 ± 4.48	9.25 ± 0.54	34.34 ± 3.35	10.52 ± 0.47	45.44 ± 3.72
5	10.38 ± 0.64	42.80 ± 3.12	9.56 ± 0.47	37.81 ± 4.78	8.48 ± 0.45	38.79 ± 2.98	9.14 ± 0.36	42.41 ± 4.13
6	9.82 ± 0.69	42.64 ± 3.94	9.22 ± 0.47	37.46 ± 4.63	8.28 ± 0.51	37.20 ± 4.64	7.12 ± 0.57	41.96 ± 4.93
7	9.08 ± 0.48	46.40 ± 2.75	8.67 ± 0.47	41.74 ± 3.48	7.57 ± 0.49	38.67 ± 3.96	6.03 ± 0.27	43.26 ± 4.69
8	6.08 ± 0.33	0	8.06 ± 0.36	38.74 ± 4.05	7.09 ± 0.50	36.38 ± 4.83	5.48 ± 0.33	42.58 ± 3.96
9	–	–	7.56 ± 0.31	44.31 ± 3.40	6.49 ± 0.54	41.01 ± 3.70	5.25 ± 0.35	43.86 ± 3.70
10	–	–	7.16 ± 0.37	39.28 ± 4.79	5.98 ± 0.42	43.79 ± 3.09	4.81 ± 0.40	44.45 ± 4.91

Table 2. Karyotypes of *Arge* species.

Species	n(2n)	Chromosomal formula, n	Region	Reference
A. ciliaris (Linnaeus, 1767)	10	3M + 7M/SM	European Russia	Present paper
A. clavicornis (Fabricius, 1781)	8	8M[†]	Eastern Canada	Maxwell 1955, 1958
A. cyanocrocea (Förster, 1771)	11	6M + 5SM	Eastern Germany	Westendorff and Taeger 2002
A. enodis (Linnaeus, 1767)	10	2M + 8M/SM	European Russia	Present paper
A. gracilicornis (Klug, 1814)	8	7M + 1ST	Eastern Germany	Westendorff and Taeger 2002
	8	7M + 1A	European Russia	Present paper
A. jonasi (Kirby, 1882)	10	5M + 5M/SM[†]	Japan	Naito 1976
A. melanochra (Gmelin, 1790)	10	5M + 5SM	Eastern Germany	Westendorff and Taeger 2002
A. nigripes (Retzius, 1783)	13	4M + 4M/SM + 5ST/A[†]	Eastern Germany	Westendorff and Taeger 2002
A. nigronodosa (Motschulsky, 1860)	8	4M + 3M/SM + 1A[†]	Japan	Naito 1982
A. pagana (Panzer, 1798)	(16)	7M + 1M/SM	Eastern Germany	Westendorff and Taeger 2002
A. pectoralis (Leach, 1817)	8	8M[†‡]	Eastern Canada	Maxwell 1955, 1958
	8(16)	?	Scotland, UK	Greenshields 1937
A. ustulata (Linnaeus, 1758)	10	8M + 2M/SM	European Russia	Present paper

[†] Extrapolated from images given in the references.

[‡] Although both images given by Maxwell (1955) for this species show n = 7, they strongly differ in the relative size of the largest chromosome. However, this author defines both *A. clavicornis* and *A. pectoralis* as "cytologically homogeneous" (Maxwell 1955, 1958), and I therefore follow Westendorff (2006) who suggested n = 8 for the latter species.

In turn, the second chromosome is approximately 1.8 times shorter than the preceding one, all other elements more or less gradually decreasing in length. On most metaphase plates, a secondary constriction can be clearly seen in the pericentromeric region of the longer arm of the second chromosome.

A. ustulata (Linnaeus, 1758) (n = 10). Seven embryos obtained from four females were studied. The karyotype generally resembles that of *A. ciliaris* (Fig. 1D, Table 2). As in the previous species, most chromosomes, except for the first and second ones, form a continuous gradation in size, but the fifth chromosome is visibly longer than the remaining elements (Table 1).

Discussion

Up to now, karyotypes of 12 members of the genus *Arge* have been studied. In these sawflies, haploid chromosome numbers of n = 8, 10, 11 and 13 were found (Table 2). Among these species, n = 8 appeared to be the most frequent chromosome number, followed by n = 10. Within chromosome sets of *Arge* species, metacentrics and submetacentrics usually predominate (Table 2, Westendorff and Taeger 2002), although most members of the genus with the same n values differ by their karyotype structure. For example, n = 10 is characteristic of both *A. ciliaris* and *A. enodis*, but the chromosome set of the former species contains a very large metacentric, which is absent from the karyotype of *A. enodis*. Analogously, *A. gracilicornis* and *A. nigronodosa* both have chromosome sets with n = 8, again with a large first metacentric, but the second metacentric/submetacentric chromosome of the latter species is substantially longer than that of *A. gracilicornis* (Naito 1982; Westendorff and Taeger 2002; present study). Moreover, Westendorff and Taeger (2002) identified the last chromosome of *A. gracilicornis* as a subtelocentric, which can also be clearly seen on Fig. 1 of their paper, but a shorter arm of an analogous acrocentric chromosome of apparently the same species is not visible (present paper, Fig. 1A). However, it is unclear at the moment whether this feature represents an intraspecific chromosomal polymorphism or indicates the presence of cryptic species within the *A. gracilicornis* complex. In *A. ustulata*, a common European species, possible involvement of cryptic taxa is also supposed. Specifically, n = 8 and 2n = 16 were reported in the 1930s for this sawfly species in the United Kingdom (Greenshields, 1937), whereas material from central European Russia clearly shows n = 10 (present paper). Nevertheless, wrong identification of the British material cannot be completely ruled out as well.

Given the relatively high karyotypic diversity of the genus *Arge*, it is difficult to understand what the initial karyotype for the group might look like. Judging from the most frequent chromosome numbers, the ancestral n value could be close to 8 or 10. Both these numbers fall within range of putative initial values for the superfamily Tenthredinoidea and Argidae in particular, i.e., n = 7 to 10 (Gokhman 2023). Moreover, n = 8 is the only chromosome number found in different subfamilies of Pergidae, a sister group to Argidae (Boevé et al. 2018; Gokhman 2023). Nevertheless, n = 8 and 10 alternatively

predominate in two apparent *Arge* clades (Boevé et al. 2018), but the ancestral chromosome number for this group may also be substantially higher. In addition, karyotypes of various members of the genus *Arge* contain the very large first metacentric chromosome, e.g., *A. ciliaris*, *A. gracilicornis*, *A. melanochra*, *A. pagana* and *A. ustulata* (Westendorff and Taeger 2002; present paper). However, whether this chromosome represents an ancestral character state for the genus remains an open question. Analogously, little can be said at present about the possible chromosomal rearrangements underlying the process of karyotypic change within this genus. Similarly to other sawflies and Hymenoptera in general, differences between karyotypes of related *Arge* species could be explained by chromosomal fusions/fissions, deletions/duplications of the constitutive heterochromatin, translocations and/or inversions (Gokhman 2009, 2023).

Nevertheless, I believe that karyotype analysis can be successfully used in further taxonomic and phylogenetic studies of the genus *Arge* due to its high chromosomal diversity. Our results together with published karyotypic data collectively suggest that chromosome sets of most species of this group can be easily distinguished without a detailed morphometric analysis. On the other hand, this kind of analysis can be important at least in some cases, which can be judged from an example of *A. ciliaris* and *A. ustulata* (see above). This situation is generally similar to the pattern observed in other studied sawfly families, e.g., Tenthredinidae (Westendorff 2006; Gokhman 2023).

Acknowledgements

References

Boevé J-L, Nyman T, Shinohara A, Schmidt S (2018) Endogenous toxins and the coupling of gregariousness to conspicuousness in Argidae and Pergidae sawflies. Scientific Reports 8: 17636. https://doi.org/10.1038/s41598-018-35925-z

Gokhman VE (2009) Karyotypes of Parasitic Hymenoptera. Dordrecht, Springer, 183 pp. https://doi.org/10.1007/978-1-4020-9807-9

Gokhman VE (2023) Chromosomes of Symphyta (Hymenoptera): current state and perspectives of research. Russian Entomological Journal 32: 375–382.

Greenshields F (1937) Hymenoptera Symphyta on the Island of Raasay. The Scottish Naturalist 223–228: 138–144.

Imai HT, Taylor RW, Crosland MWJ, Crozier RH (1988) Modes of spontaneous chromosomal mutation and karyotype evolution in ants with reference to the minimum interaction hypothesis. Japanese Journal of Genetics 63: 159–185. https://doi.org/10.1266/jjg.63.159

Naito T (1976) Chromosomes and sex of sawflies. Forest Pests 25(5): 66–71. [In Japanese]

Naito T (1982) Chromosome number differentiation in sawflies and its systematic implication (Hymenoptera, Tenthredinidae). Kontyû 50(4): 569–587.

Levan A, Fredga K, Sandberg AA (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201–220. https://doi.org/10.1111/j.1601-5223.1964.tb01953.x

Maxwell DE (1955) Cytology and correlated morphology of the genus *Neodiprion* Rohwer (Hymenoptera: Symphyta). PhD Thesis, McGill University, Montreal, Canada, 172 pp.

Maxwell DE (1958) Sawfly cytology with emphasis upon the Diprionidae. Proceedings of the Tenth International Congress of Entomology. Montreal, August 17–25, 1956, 2. Montreal, 961–978.

Taeger A, Liston AD, Prous M, Groll EK, Gehroldt T, Blank SM (2018) ECatSym – Electronic World Catalog of Symphyta (Insecta, Hymenoptera). https://sdei.de/ecatsym [Accessed 13.10.2023]

Westendorff M (2006) Chromosomes of sawflies (Hymenoptera: Symphyta) – a survey including new data. In: Blank SM, Schmidt S, Taeger A (Eds) Recent Sawfly Research: Synthesis and Prospects. Keltern, Goecke & Evers, 39–60.

Westendorff M, Taeger A (2002) New data on chromosomes of sawflies in the families Argidae, Cimbicidae and Cephidae. Beiträge zur Entomologie 52(2): 347–352. https://doi.org/10.21248/contrib.entomol.52.2.347-352

ORCID

Vladimir E. Gokhman https://orcid.org/0000-0001-9909-7559