



OPEN

## Large-scale comparative analysis of cytogenetic markers across Lepidoptera

Irena Provazníková<sup>1,2,3</sup>, Martina Hejníčková<sup>1,2</sup>, Sander Visser<sup>1,2,4</sup>, Martina Dalíková<sup>1,2</sup>, Leonela Z. Carabajal Paladino<sup>1,5</sup>, Magda Zrzavá<sup>1,2</sup>, Anna Voleníková<sup>1,2</sup>, František Marec<sup>1,2</sup> & Petr Nguyen<sup>1,2</sup>✉

Fluorescence in situ hybridization (FISH) allows identification of particular chromosomes and their rearrangements. Using FISH with signal enhancement via antibody amplification and enzymatically catalysed reporter deposition, we evaluated applicability of universal cytogenetic markers, namely 18S and 5S rDNA genes, U1 and U2 snRNA genes, and histone H3 genes, in the study of the karyotype evolution in moths and butterflies. Major rDNA underwent rather erratic evolution, which does not always reflect chromosomal changes. In contrast, the hybridization pattern of histone H3 genes was well conserved, reflecting the stable organisation of lepidopteran genomes. Unlike 5S rDNA and U1 and U2 snRNA genes which we failed to detect, except for 5S rDNA in a few representatives of early diverging lepidopteran lineages. To explain the negative FISH results, we used quantitative PCR and Southern hybridization to estimate the copy number and organization of the studied genes in selected species. The results suggested that their detection was hampered by long spacers between the genes and/or their scattered distribution. Our results question homology of 5S rDNA and U1 and U2 snRNA loci in comparative studies. We recommend the use of histone H3 in studies of karyotype evolution.

Cytogenetic studies aim at characterization of genome organization and its changes. Previously indispensable for the identification of genes of interest, cytogenetics may seem to struggle in the post-genomic era as it lags behind the resolution of molecular biology and genomics. Yet it remains crucial for genomic research. Cytogenetic data such as genome size and chromosome number allow for an informed choice of sequencing strategies and provide hypothetical framework for genomic studies, context to bioinformatic analyses, and physical evidence for results produced in silico<sup>1–3</sup>. Recent efforts, such as the Earth BioGenome project that aspire to characterize genomes of all eukaryotic biodiversity<sup>4</sup>, will without a doubt lead to further cytogenetic research. As a result, the new field integrating cytogenetics and genomics has recently been proposed under the term chromosomics (coined by Claussen<sup>5</sup> but repurposed later by Graphodatsky<sup>6</sup> and Deakin et al.<sup>3</sup>).

There are several approaches to distinguish individual chromosomes within a karyotype. Classical techniques such as orcein or Giemsa staining as well as various banding methods can produce chromosome-specific patterns. These techniques work very well in mammals including humans<sup>7,8</sup>, other vertebrates<sup>9</sup>, some invertebrate taxa<sup>10–12</sup> and plants<sup>13,14</sup>. However, classical staining and banding techniques have failed in some organisms, such as moths and butterflies<sup>15,16</sup>.

Lepidoptera with more than 160,000 described species and great ecological diversity<sup>17</sup> represent an excellent model system to study karyotype evolution and the role of changes in genome architecture in evolutionary processes. In Lepidoptera, chromosomal rearrangements such as inversions, fusions, and fissions play an important role in speciation<sup>18</sup> and adaptation, such as resistance to insecticides<sup>19,20</sup> and baculoviruses<sup>21</sup> and detoxification of plant secondary metabolites<sup>22</sup> and xenobiotics<sup>23</sup>. However, comparative cytogenetic studies are scarce in Lepidoptera due to the many peculiarities of lepidopteran chromosomes. Mitotic complements of both Lepidoptera and their sister group Trichoptera typically consist of a high number of small and morphologically uniform holokinetic chromosomes<sup>24,25</sup>. Since they lack a primary constriction, i.e. the centromere, its position cannot be used in chromosome identification<sup>26</sup>. Thus, cytogenetic analyses of lepidopteran karyotypes were challenging for years before molecular cytogenetic tools were introduced<sup>25,27</sup> and applied on meiotic pachytene chromosomes rather than mitotic chromosomes<sup>28</sup>. However, broader comparative cytogenetic studies, which would help us to

<sup>1</sup>Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic. <sup>2</sup>Institute of Entomology, Biology Centre CAS, České Budějovice, Czech Republic. <sup>3</sup>Present address: European Molecular Biology Laboratory, Heidelberg, Germany. <sup>4</sup>Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands. <sup>5</sup>The Pirbright Institute, Surrey, UK. ✉email: petr.nguyen@prf.jcu.cz

understand major trends in karyotype evolution of moths and butterflies, are few<sup>29,30</sup> due to a lack of appropriate cytogenetic markers that can be used on this scale.

Mapping specific sequences on chromosome preparations by means of fluorescence in situ hybridization (FISH) allows us to identify particular chromosomes, study their potential rearrangements and origin, and their behaviour during cell divisions<sup>29,31,32</sup>. Various tandemly arrayed genes have been established as suitable markers for cytogenetic comparative studies. These universal markers have proved to be useful in a wide range of non-model species due to their conserved nature and ease of visualization by FISH methods<sup>33–35</sup>. The most commonly used markers are genes for major ribosomal RNAs (rDNA). Genes for 18S, 5.8S, and 28S ribosomal RNA form a transcription unit organized in clusters, which can contain hundreds or thousands of copies<sup>36,37</sup>. Genes for the 5S ribosomal RNA are also used<sup>38–41</sup>. 5S ribosomal RNA is distributed independently from the major rDNA array and is used as an independent marker. As with the major rDNA array, 5S rDNA can be localized in clusters containing tens to thousands of copies<sup>38,42,43</sup> but can also be found as singular copies scattered throughout the genome<sup>43</sup>. Abundant data on the number and localization of both 5S and the major rDNA gene clusters in animals and plants are available in public databases<sup>44,45</sup>. Finally, another group of markers used in cytogenetic studies includes the uridine-rich small nuclear RNA (U-rich snRNA) genes, which are an important part of the spliceosome. For cytogenetic purposes, U1 and U2 snRNA genes have been used. U1 snRNA gene clusters have been mapped in only a few species of Orthoptera<sup>42,46</sup>, Isopoda<sup>47</sup>, and fish<sup>48</sup>. U2 snRNA has been used only in a few fish species (e.g. Refs. <sup>49–51</sup>). U1 and U2 snRNA genes are relatively new markers often used in combination with other markers as major rDNA.

Despite their easy visualization and universality, rDNA and snRNA markers also have some limitations. Their evolution is highly dynamic, and changes in their distribution do not always reflect chromosome rearrangements<sup>46,52</sup>. They have been compared with mobile elements and in several cases have actually been found to be associated with transposons<sup>48,53,54</sup>. FISH experiments using 18S and 28S rDNA genes as probes successfully revealed concealed karyotype variation between populations and closely related species of both plants (e.g. Ref. <sup>55</sup>) and animals (e.g. Refs. <sup>56,57</sup>). Therefore, rDNA and snRNA genes might be good markers for chromosome evolution between closely related species or even intra-species evolution but are less informative with increasing evolutionary scale. To study such large-scale chromosome evolution patterns, additional markers should be developed that evolve less erratically.

Despite their great potential, histone genes have rarely been used in cytogenetic studies. Histone genes encode H1, H2A, H2B, H3, and H4 proteins, which have a strong affinity for DNA. Together, the histone proteins and DNA form a nucleosome, the basic unit of chromatin<sup>58</sup>. Histone genes usually form tandem arrays, as this facilitates efficient transcription<sup>58,59</sup>. The histone genes are conserved in their protein sequence and also in the distribution of their clusters in the genome<sup>60,61</sup>. This makes them ideal chromosomal markers<sup>62</sup> as differences in their number and position genuinely reflect chromosomal rearrangements<sup>60</sup>. Some examples of successful application of histone genes to fish<sup>61</sup>, Bivalvia<sup>63</sup>, and insects<sup>34,39</sup>, including Lepidoptera<sup>64,65</sup>, show their applicability in various organisms.

In this study, we analysed the chromosomal distribution of several universal cytogenetic markers, namely 18S and 5S rDNAs, U1 and U2 snRNA genes, and histone H3 genes, in 29 species of Lepidoptera to evaluate their applicability and resolution in the study of karyotype evolution. We found that some of the markers can be used successfully in all species, while others cannot be detected in certain species. To determine the reason for the unsuccessful detection of markers by FISH, we used quantitative PCR and Southern hybridization to estimate copy numbers and distribution patterns in different species. The obtained results provide not only information on the use of various markers in Lepidoptera, but also on trends in changes in the architecture of lepidopteran genomes.

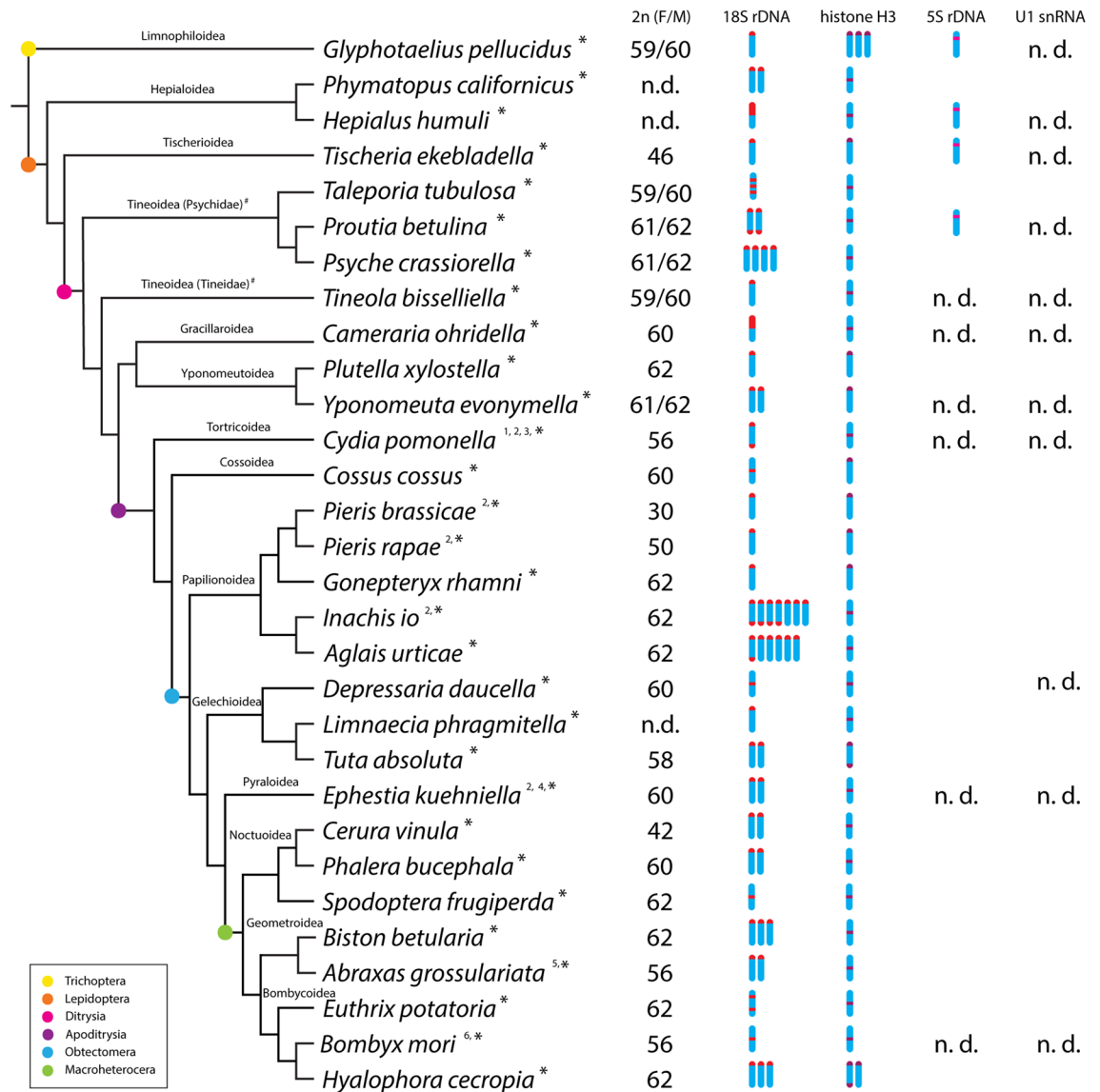
## Results

**Localization of 18S rDNA and histone H3 genes.** To visualize clusters of the major rRNA genes, we used FISH with a partial sequence of 18S rDNA from the codling moth, *Cydia pomonella* (Tortricidae) as a probe<sup>25</sup>. Since the nucleotide sequence of 18S rDNA is highly conserved, the probe successfully hybridized onto chromosomal preparations of all studied species sampled across the order Lepidoptera, as well as the representative of their sister order Trichoptera. Major rDNA clusters were detected at a terminal position in 22 out of 30 species. Only 6 species with interstitial clusters were documented. Multiple, up to 11, clusters were observed in approximately half of the studied species.

Although histone genes are known for their highly conserved protein sequence, they can differ significantly at the nucleotide level due to the degeneracy of codons. To ensure optimal hybridization, a fragment of the histone H3 gene was amplified, sequenced, and used as a specific probe from each species studied (Supplementary Table S1) except for few (for details see “Materials and methods”). To increase sensitivity of the FISH detection, we employed TSA-FISH which can detect unique sequences > 1300 bp<sup>27</sup>. In total, we successfully mapped the distribution of histone gene clusters in all studied species. In the vast majority, a single cluster was detected, located interstitially or terminally. Multiple histone clusters (2–3) were observed only in two lepidopteran species, *Tuta absoluta* and *Hyalophora cecropia*, and in the outgroup species *Glyptotaelius pellucidus* (Trichoptera).

All results from the mapping of 18S rDNA and histone H3 genes are summarized in Fig. 1 and Supplementary Table S2. For a complete overview we also added information on chromosome numbers and the distribution of 18S rDNA and histone H3 genes available to date in other Lepidoptera.

**Trichoptera and non-Ditrysia.** The diploid chromosome number of the caddis fly, *Glyptotaelius pellucidus* (Limnephiloidea)  $2n_{\text{♀}} = 59$ ,  $Z0/2n_{\text{♂}} = 60$ , ZZ, was described previously<sup>71</sup>. FISH experiments using 18S rDNA probe revealed a pair of terminal signals on one autosomal bivalent in this species (Supplementary



**Figure 1.** Overview of the number and position of 18S rDNA, histone H3, 5S rDNA, and U1 snRNA markers in haploid genomes of studied species. Phylogenetic relationships are based on Refs.<sup>66–68</sup>. #Tineoidea are considered paraphyletic. Data was obtained: \*in this study; 1–<sup>25</sup>, 18S rDNA; 2–<sup>29</sup>, 18S rDNA; 3–<sup>64</sup> histone H3; 4–<sup>69</sup>, 18S rDNA; 5–<sup>70</sup>, 18S rDNA; 6–<sup>31</sup>, 18S rDNA. n.d.—not detected. F/M—female and male diploid chromosome numbers, if different. A complete list of all species analysed so far for the distribution of studied markers, including their chromosomal numbers and references, is given in Supplementary Table S2. The figure was created in Adobe Illustrator 2020, version 24.0 ([www.adobe.com](http://www.adobe.com)).

Fig. S1a). Interestingly, various strong heterochromatin blocks in almost all chromosome bivalents were visible after staining with DAPI. Some of these heterochromatin patterns could potentially be used for chromosome identification. Hybridization of histone H3 probe revealed three terminal clusters of histone genes on three different bivalents (Supplementary Fig. S1b). This is one of three cases in our study where we observed multiple histone clusters (Fig. 1).

Two species of the superfamily Hepialoidea were examined, namely the ghost moth, *Hepialus humuli*, and the lupine ghost moth, *Phymatopus californicus*. Diploid chromosome numbers of these two species have not been described yet. Due to the lack of mitotic nuclei, we were not able to determine chromosomal numbers in this study. In *H. humuli*, the 18S rDNA probe highlighted approximately half of one pachytene bivalent (Supplementary Fig. S1c). Hybridization signals colocalized with a DAPI-positive heterochromatin block. In *P. californicus*, two chromosomal bivalents were detected, each bearing an rDNA cluster at the chromosome terminus (Supplementary Fig. S1e). The histone H3 probe revealed one bivalent with an interstitial cluster of histone genes in both *H. humuli* and *P. californicus* (Supplementary Fig. S1d, f).

The oak leaf miner, *Tischeria ekebladella*, was examined as a representative of the Tischerioidea superfamily. Its diploid chromosomal number 2n = 46 (for both sexes) was determined previously<sup>72</sup>. Indeed, n = 23 was confirmed in this study (Supplementary Fig. S2a). After DAPI staining, a strong heterochromatin block with terminal

or subterminal location was visible in three pachytene bivalents. The subterminal heterochromatin block was adjacent to a terminal rDNA cluster highlighted by the 18S rDNA probe on one of the longer bivalents in male pachytene nuclei (Supplementary Fig. S1g). A single histone gene cluster was localized at the end of another bivalent (Supplementary Fig. S1h).

**Basal Ditrysia.** Three bagworm species from the family Psychidae, namely *Taleporia tubulosa* ( $2n_{\text{♀}} = 59$ ,  $Z0/2n_{\text{♂}} = 60$ , ZZ), *Proutia betulina* ( $2n_{\text{♀}} = 61$ ,  $Z0/2n_{\text{♂}} = 62$ , ZZ), and *Psyche crassiorella* ( $2n_{\text{♀}} = 61$ ,  $Z0/2n_{\text{♂}} = 62$ , ZZ)<sup>73–75</sup> and one species from the family Tineidae, the common clothes moth *Tineola bisselliella* ( $2n_{\text{♀}} = 59$ ,  $Z0/2n_{\text{♂}} = 60$ , ZZ)<sup>76</sup>, were studied. The 18S rDNA probe revealed a distinct hybridization pattern in each species. In pachytene nuclei of *T. tubulosa*, an extraordinary pattern of three strong interstitial rDNA signals located on a single bivalent with regular spacing was observed (Supplementary Fig. S3a). In pachytene nuclei of *P. betulina*, one bivalent with signals on both ends and two bivalents bearing one terminal signal each were observed (Supplementary Fig. S3c). In *P. crassiorella*, four terminal signals were located on four bivalents (Supplementary Fig. S3e). In *T. bisselliella* male pachytene nuclei, a single rDNA locus was detected in a subterminal region of a pachytene bivalent (Supplementary Fig. S3g). Only one interstitial cluster of histone genes was observed in all four species, *T. tubulosa*, *P. betulina*, *P. crassiorella*, and *T. bisselliella* (Supplementary Fig. S3b,d,f,h).

The diploid chromosome number  $2n = 60$  of the horse-chestnut leaf miner, *Cameraria ohridella* (Gracillariodea), was determined previously<sup>77</sup>. The 18S rDNA probe hybridized to a terminal region of one bivalent in male pachytene nuclei (Supplementary Fig. S4a). A strong, yet discontinuous signal covered approximately one fourth of the bivalent in a pattern similar to the one observed in *H. humuli* (see above). Mapping of the histone H3 gene showed one interstitial histone cluster on the rDNA bearing bivalent (Supplementary Fig. S4b).

Two species from the superfamily Yponomeutoidea were examined, i.e. the diamondback moth *Plutella xylostella* (Plutellidae;  $2n = 62$ <sup>78</sup>) and the bird-cherry ermine moth *Yponomeuta evonymella* (Yponomeutidae;  $2n_{\text{♀}} = 61$ ,  $Z_1Z_2W/2n_{\text{♂}} = 62$ ,  $Z_1Z_1Z_2Z_2$ ; Ref.<sup>79</sup> and references therein). FISH experiments carried out on pachytene nuclei of *P. xylostella* revealed a single terminal cluster of rDNA genes (Supplementary Fig. S4c). The histone H3 probe revealed one terminal cluster which colocalized with a strong heterochromatic block (Supplementary Fig. S4d). On chromosomal preparations of *Y. evonymella*, the 18S rDNA probe showed two bivalents with terminal signals of similar size (Supplementary Fig. S4e). One terminal histone cluster was observed in pachytene nuclei of *Y. evonymella* (Supplementary Fig. S4f).

**Apoditrysia.** In a representative of the Cossioidea superfamily, the goat moth *Cossus cossus*, we determined the diploid male chromosome number  $2n = 60$  (Supplementary Fig. S5a). FISH with the 18S rDNA probe on male pachytene nuclei revealed one chromosomal pair bearing an interstitial cluster which colocalized with a small block of DAPI-positive heterochromatin (Supplementary Fig. S5a). The histone H3 probe labelled one cluster at the end of one chromosome bivalent (Supplementary Fig. S5b). However, it should be noted that due to the lack of material, our FISH experiments were performed on only one male *C. cossus* larva. The karyotype of the codling moth, *Cydia pomonella* (Tortricidae), was already described as  $2n = 56$  by Ref.<sup>80</sup> and later verified by Ref.<sup>25</sup>. Our results of mapping of the 18S rDNA and histone H3 probes (Supplementary Fig. S5c,d) confirmed previously published data, i.e. two rDNA clusters at both ends of a single chromosome bivalent and another bivalent bearing one interstitial histone cluster<sup>25,64</sup>.

**Obtectomera.** Within the superfamily Papilionoidea we studied three species belonging to the family Pieridae and two species of the Nymphalidae family. The three studied pierids, namely the small cabbage white *Pieris rapae* ( $2n = 50$ <sup>81</sup>), the cabbage white *Pieris brassicae* ( $2n = 30$ <sup>81</sup>), and the common brimstone *Gonepteryx rhamni* ( $2n = 62$ <sup>81</sup>), differ in chromosomal numbers, however, mapping of the 18S rDNA and histone H3 genes revealed common hybridization patterns for both markers (Supplementary Fig. S6). Consistent with previous reports<sup>29</sup>, we identified one bivalent bearing a terminal rDNA cluster in *P. rapae* (Supplementary Fig. S6a) and *P. brassicae* (Supplementary Fig. S6c). We also observed this pattern in autosome pair of *G. rhamni* (Supplementary Fig. S6e). In pachytene nuclei of *P. brassicae*, small DAPI-positive blocks of heterochromatin were observed at the ends of several bivalents (Supplementary Fig. S6c). In *G. rhamni*, only one block of heterochromatin was visible, which colocalized with the 18S rDNA signal (Supplementary Fig. S6e). The histone H3 probe highlighted the terminal region in one chromosome pair in all three species (Supplementary Fig. S6b,d,f). Moreover, histone bearing chromosomes clearly correspond to autosomes in *P. brassicae*, in which the sex chromosome bivalent was identified by a typical pairing of W and Z chromosomes (Supplementary Fig. S6d).

From the family Nymphalidae, the small tortoiseshell *Aglais urticae* and the peacock butterfly *Inachis io* were examined. Both species have a chromosome number  $2n = 62$ , reported previously<sup>81</sup> and confirmed in this study (Supplementary Fig. S2b,c). In male pachytene nuclei of *A. urticae*, six to seven small rDNA clusters were observed (Supplementary Fig. S7a). FISH with the histone H3 probe revealed one interstitial cluster colocalizing with a heterochromatin block (Supplementary Fig. S7b). Mapping of 18S rDNA genes in *I. io*, which was done previously<sup>29</sup>, revealed up to 11 small terminal clusters in pachytene nuclei, three bivalents bearing one terminal signal and four bivalents carrying terminal signals at both ends. To increase the sensitivity of detection, we repeated this experiment using TSA-FISH. Our data confirm the previous identification and distribution of eleven 18S rDNA clusters in *I. io* (Supplementary Fig. S7c). Similar to *A. urticae*, we mapped a single histone cluster to an interstitial region of a bivalent, which colocalized with a block of heterochromatin (Supplementary Fig. S7d).

We studied three species from the superfamily Gelechioidea, namely the cosy flat body moth, *Depressaria daucella* (Depressariidae), the shy cosmet moth, *Limnaecia phragmitella* (Cosmopterigidae), and the tomato leafminer, *Tuta absoluta* (Gelechiidae). The diploid chromosome number of  $2n = 60$  in *D. daucella* was reported recently<sup>22</sup>. The 18S rDNA probe revealed one interstitial cluster of major rDNA (Supplementary Fig. S8a).



Similarly, we detected one interstitial histone cluster (Supplementary Fig. S8b). In *L. phragmitella*, the number of chromosomes was not determined previously, and we also failed to determine it due to the lack of mitotic chromosomes. However, using FISH mapping on pachytene chromosomes, we successfully identified one terminal rDNA cluster colocalized with a DAPI-positive block of heterochromatin (Supplementary Fig. S8c) and one interstitial cluster of histone genes (Supplementary Fig. S8d). In the *T. absoluta* strain used in this study, a diploid chromosome number of  $2n = 58$  was described previously<sup>82</sup> and confirmed in another study<sup>22</sup>. The 18S rDNA probe highlighted two clusters in terminal regions of two autosomal bivalents (Supplementary Fig. S8e). Histone gene clusters were detected at both ends of a pachytene bivalent (Supplementary Fig. S8f), which makes *T. absoluta* one of only two lepidopteran species with multiple histone gene clusters described so far.

The only representative of the superfamily Pyraloidea included in our study was the Mediterranean flour moth, *Ephestia kuehniella*. Its diploid chromosome number of  $2n = 60$  was described previously<sup>83</sup>. Two terminal rDNA clusters present on two chromosome bivalents were identified by Ref.<sup>69</sup>, which was later confirmed by means of FISH<sup>29</sup>. To complete the dataset, we additionally mapped the histone H3 probe on male pachytene nuclei, which revealed one chromosome bivalent bearing a single interstitial cluster of histone genes colocalizing with a block of heterochromatin (Supplementary Fig. S10a).

**Macroheterocera.** Three members of the Notodontidae and Noctuidae families within the superfamily Noctuoidea were examined. In male pachytene complements of the puss moth, *Cerura vinula* (Notodontidae), we verified the diploid chromosome number of  $2n = 42$  (Supplementary Fig. S2d) previously reported<sup>81</sup> and observed two autosomal bivalents carrying a terminal rDNA cluster (Supplementary Fig. S9a). One interstitial cluster of histone genes was detected by the histone H3 probe (Supplementary Fig. S9b). In female pachytene nuclei of the buff-tip, *Phalera bucephala* (Notodontidae), with a diploid number of chromosomes  $2n = 60$ <sup>81</sup>, the same hybridization pattern for both markers as in *C. vinula* was observed (Supplementary Fig. S9c,d). In the fall armyworm, *Spodoptera frugiperda* (Noctuidae), with a diploid chromosome number of  $2n = 62$ <sup>81</sup>, only one bivalent bearing an interstitial cluster of rDNA genes was identified by FISH (Supplementary Fig. S9e). Similar to the other two species, one interstitial histone cluster was detected in one of the bivalents in male pachytene nuclei (Supplementary Fig. S9f).

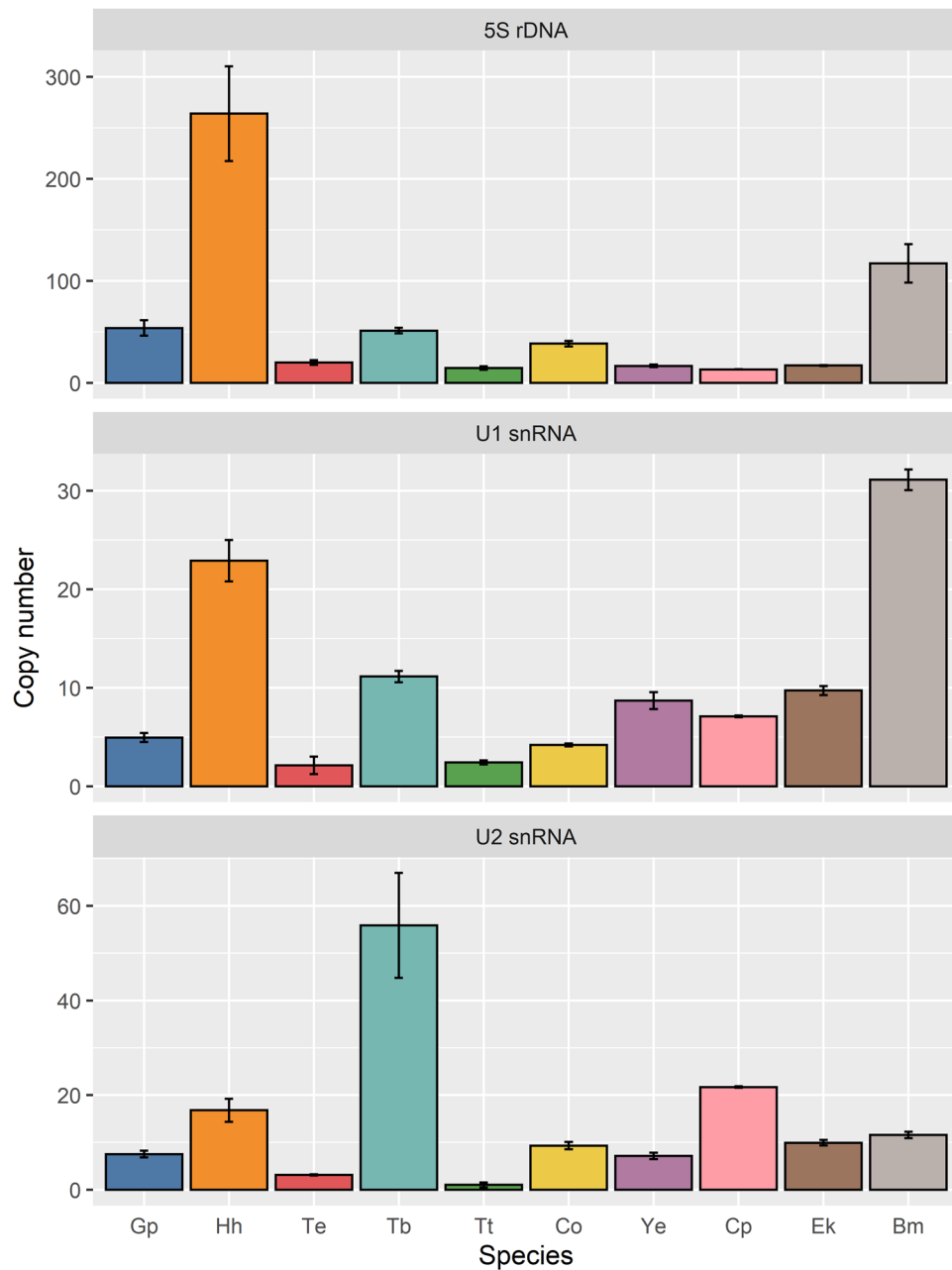
Two representatives of the superfamily Geometroidea were included in our study, the peppered moth *Biston betularia* and the magpie moth *Abraxas grossulariata* (both Geometridae). In *A. grossulariata*, the diploid number of chromosomes  $2n = 56$  was reported in an earlier study<sup>70</sup>, which also detected a single terminal rDNA cluster on W and Z sex chromosomes. On chromosomal preparations of *A. grossulariata*, we identified one interstitial cluster of histone genes (Supplementary Fig. S10b). Moreover, numerous strong DAPI-positive heterochromatin blocks were detected (Supplementary Fig. S10b), which is also in agreement with previous observations<sup>70</sup>. The diploid chromosome number of  $2n = 62$  was previously reported for *B. betularia*<sup>84</sup> and the same material was used in this study. In male pachytene nuclei, we identified three bivalents bearing a single small terminal rDNA cluster each (Supplementary Fig. S10c). Using the histone H3 probe, a single interstitial histone cluster was detected on one of the autosomal bivalents in female pachytene nuclei (Supplementary Fig. S10d).

Species from three different families were explored within the superfamily Bombycoidea, the drinker moth *Euthrix potatoria* (Lasiocampidae), the silkworm *Bombyx mori* (Bombycidae), and the cecropia silkworm *Hyalophora cecropia* (Saturniidae). A diploid chromosome number of  $2n = 62$  was previously described in *E. potatoria*<sup>81</sup> and was confirmed by our results (Supplementary Fig. S2e). Hybridization of the 18S rDNA probe revealed an interesting distribution of rDNA genes, namely two interstitially located clusters within one pachytene bivalent (Supplementary Fig. S11a). The histone H3 probe uncovered one interstitial cluster of histone genes (Fig. S11b). The diploid karyotype of *B. mori* consists of  $2n = 56$  chromosomes<sup>85</sup>. Distribution of rDNA was previously reported as a single interstitial rDNA cluster<sup>29</sup>. We confirmed this in male pachytene preparations (Supplementary Fig. S11c). Moreover, the FISH experiments with the histone H3 probe also revealed an interstitial position of a single histone gene cluster in one of the chromosomal pairs (Supplementary Fig. S11d).

The karyotype of *H. cecropia* consists of  $2n = 62$  chromosomes, as previously reported<sup>86</sup>, which is corroborated also by our observation (Supplementary Fig. S2f). Three terminal clusters of rDNA genes in three different bivalents were mapped by the 18S rDNA probe (Supplementary Fig. S11e). Histone H3 mapping revealed two histone gene clusters at both ends of one bivalent and another cluster at one end of another bivalent colocalizing with strong blocks of heterochromatin. Bivalents bearing the histone clusters were almost exclusively associated in pachytene complements forming a specific configuration (Supplementary Fig. S11f).

**Mapping of 5S rDNA and U1 and U2 snRNA genes.** The 5S rDNA gene and U1 and U2 snRNA genes have never been used as cytogenetic markers in the order Lepidoptera. Therefore we decided to test their suitability for comparative analysis within this order. We chose nine species from different families across the whole order Lepidoptera with a focus on basal groups, namely *H. humuli* (Hepialidae), *T. ekebladella* (Tischeriidae), *T. bisselliella* (Tineidae), *T. tubulosa* (Psychidae), *C. ohridella* (Gracillariidae), *Y. evonymella* (Yponomeutidae), *C. pomonella* (Tortricidae), *E. kuehniella* (Pyralidae), *B. mori* (Bombycidae), and one outgroup species, *G. pellucidus* (Limnephilidae), from the sister order Trichoptera. We amplified and labelled species-specific probes for the 5S rDNA, U1 and U2 snRNA genes, and used them in FISH experiments in the respective species (Supplementary Table S1).

Although we used an optimized TSA-FISH protocol to maximize sensitivity of the FISH experiments, we successfully mapped 5S rDNA only in four species, namely *G. pellucidus*, *H. humuli*, *T. ekebladella*, and *T. tubulosa*. In all four species, we detected one subterminal cluster of 5S rDNA genes on one chromosomal pair (Supplementary Fig. S12). In the other species, no clear hybridization signals were identified. The U1 and U2 snRNA genes did not show any hybridization signals in any of the ten species studied (summary of results in Fig. 1). The negative



**Figure 2.** Estimated copy numbers of 5S rDNA, U1 snRNA, and U2 snRNA genes per haploid genome in selected species. Gp, *G. pellucidus* (Trichoptera, outgroup); Hh, *H. humuli*; Te, *T. ekebladella*; Tb, *T. bisselliella*; Tt, *T. tubulosa*; Co, *C. ohridella*; Ye, *Y. evonymella*; Cp, *C. pomonella*; Ek, *E. kuehniella*; Bm, *B. mori*. For a summary of the qPCR results, see Supplementary Table S3.

results of the FISH experiments suggest that the genomic arrangement of these three genes is not suitable for FISH mapping in Lepidoptera. For example, these genes may occur in low numbers in tandem arrays or may be scattered throughout the genome, rather than clustered. To test these hypotheses, we carried out quantitative PCR (qPCR) and Southern hybridization.

**qPCR experiments.** Quantitative PCR was carried out to estimate relative copy number of 5S rDNA, U1 snRNA, and U2 snRNA genes in ten representatives probed for these genes by TSA-FISH (see above). Estimated copy numbers are summarized in Fig. 2 and Supplementary Table S3. In the case of 5S rDNA, estimated copy numbers ranging from 13 to 264 copies (mean  $60.602 \pm 14.506$ , median 29.198). The 5S rDNA copy number was higher (>50 copies) in three of the four species in which the 5S rDNA locus was detected by TSA-FISH, namely *G. pellucidus*, *H. humuli*, and *T. tubulosa*. Using TSA-FISH we also localized 5S rDNA in *T. ekebladella*, although its copy number was much lower (~20 copies) and comparable with other species in which 5S rDNA

could not be localized. However, we were unable to localize 5S rDNA by TSA-FISH in *C. ohridella* and *B. mori* with 5S rDNA copy numbers of 38 and > 100, respectively. For U1 snRNA and U2 snRNA, the results obtained showed low copy numbers in all species ranging from 2 to 31 copies (mean 10.436 SD  $\pm$  0.595, median 7.885) and from 1 to 56 copies (mean 14.387 SD  $\pm$  3.331, median 9.611) per haploid genome, respectively.

**Southern hybridization.** To test whether 5S rDNA and U1 snRNA genes are organized in tandem arrays, we performed Southern hybridization in ten selected species (see above). The U2 snRNA was excluded from this analysis due to difficulties in preparing digoxigenin-labelled probes.

Southern hybridization of the 5S rDNA probe was successful in all species tested (Supplementary Fig. S13). Results revealed multiple DNA fragments mostly > 2000 bp bearing the target sequence in all species examined. The intensity of hybridization signals was mostly uniform, although stronger bands correlating with multiple gene copies were identified e.g. in *G. pellucidus*, *H. humuli*, and *T. bisselliella*. Strong bands of smaller size, which presumably correspond to identical repeat units derived from tandem arrays, were observed only in *G. pellucidus*. Given the low copy numbers indicated by qPCR, the Southern hybridization results suggest the 5S rDNA gene copies are either scattered throughout the genome or loosely associated, i.e. individual copies are separated by varying spacers longer than 2000 bp.

Hybridization patterns using the U1 snRNA probe were similar to the 5S rDNA patterns in all species tested (Supplementary Fig. S14). Multiple bands of mostly weak intensity were observed, which implies that the 10 copies determined on average by qPCR are mostly either separated by long spacers or scattered across the genome in all species studied. Stronger bands corresponding to DNA fragments bearing multiple U1 snRNA copies were detected e.g. in *C. ohridella* and *C. pomonella* (Supplementary Fig. S14d,f). In the latter, however, the stronger bands can comprise multiple bands due to insufficient separation of long fragments (Supplementary Fig. S14f). In *T. ekebladella* and *T. tubulosa*, we were not able to successfully perform Southern hybridization, probably due to the low quality of the input gDNA and/or insufficiently labelled probes.

## Discussion

In this study, we tested whether commonly used cytogenetic markers, namely 18S rDNA, histone H3, 5S rDNA, and U1 and U2 snRNA genes, are applicable and informative for studies of karyotype evolution in Lepidoptera. We employed fluorescence in situ hybridization techniques, rDNA-FISH and TSA-FISH, which enhance hybridization signals by antibody amplification and enzymatically catalysed reporter deposition, respectively. We complemented our FISH results by estimating the copy number of the markers by qPCR and characterizing their genomic organization using Southern hybridization.

Nguyen et al.<sup>29</sup> reviewed available data on the distribution of major rDNA in Lepidoptera and mapped 18S rDNA in 18 ditrysian species from 4 superfamilies (Pyraloidea, Bombycoidea, Papilionoidea, Noctuoidea). The results suggested that in karyotypes with one locus, rDNA was usually localized interstitially, whereas in karyotypes with two or more clusters, rDNA loci were detected at chromosome ends. It was hypothesized that rDNA can spread to terminal chromosome regions by ectopic recombination between subtelomeric repetitive sequences. However, missing data from non-ditrysian and early diverging ditrysian families did not allow inferring an ancestral rDNA distribution. To fill these gaps, we carried out FISH with the 18S rDNA probe in 27 moth and butterfly species with a special focus on early diverging taxa. We also investigated one trichopteran species as an outgroup.

The results of this study (summarized in the Fig. 1) suggest that one terminal cluster of rDNA genes is an ancestral state, as it is present in the outgroup and across all lepidopteran families. In species with multiple chromosomes bearing rDNA, these clusters are usually located terminally and there is a trend towards an increase in rDNA loci in Lepidoptera. Interestingly, the rDNA loci also multiplied in the early diverging ditrysian lineage Psychidae (Supplementary Fig. S3). The highest numbers of rDNA clusters, 11 and 7, were detected in two nymphalid species, *I. io* and *A. urticae*, respectively (Supplementary Fig. S7). Other karyotype features, such as chromosomal number  $n = 31$  and presence of a single interstitial histone cluster (Supplementary Fig. S7), which are both considered ancestral traits (Refs.<sup>72,87</sup>; this study, see below), do not point to any large-scale chromosomal rearrangements in these nymphalids. Thus, the multiplication of rDNA clusters in nymphalids concurs with the ectopic recombination-driven spread of rDNA into new loci. Remarkably, multiple interstitial rDNA clusters present on a single bivalent were also documented in some species. Three interstitial rDNA clusters within a single bivalent were detected in *T. tubulosa* (Supplementary Fig. S3a) and two interstitial clusters within one bivalent were observed in *E. potatoria* (Bombycoidea) (Supplementary Fig. S11a). In both cases, the multiple clusters probably originated from intrachromosomal rearrangements such as inversions of a region containing part of the rDNA cluster (cf. Ref.<sup>88,89</sup>). In the case of *H. humuli* (Hepialoidea) (Supplementary Fig. S1c) and *C. ohridella* (Gracillarioidea) (Supplementary Fig. S4a), rDNA covers almost half of the chromosome. In addition, in *H. humuli* the rDNA cluster colocalizes with a strong heterochromatin block indicating the presence of repetitive sequences potentially associated with rDNA. More detailed research is needed to determine the mechanism of rDNA spread in these two species. Our data show that the multiplication of the major rDNA cluster occurs in multiple lepidopteran families and via different mechanisms, without any clear evolutionary pattern. This erratic behaviour makes the major rDNA an uninformative marker for the study of karyotype evolution in Lepidoptera.

Histone H3 genes have previously been mapped in several lepidopteran species<sup>64,65,90</sup>. One interstitial cluster of histone H3 genes was identified consistently in five species of the family Tortricidae<sup>64</sup>. Histone H3 genes were also localized in four *Leptidea* spp. (Pieridae)<sup>65,90</sup>. The position of the histone gene cluster was stable in *L. amurensis*, but in the other three *Leptidea* species, the number and position varied even among the offspring of one female. The karyotype evolution of *Leptidea* butterflies is known to be dynamic, characterized by unstable chromosome numbers<sup>65</sup>, and the distribution of histone gene clusters thus reflects this instability<sup>65,91</sup>. To analyse

common trends in histone cluster repatterning across Lepidoptera, we mapped histone H3 genes in 29 moth and butterfly species and one caddisfly outgroup.

In the vast majority of species, TSA-FISH with the histone H3 probe revealed a single bivalent bearing the histone gene cluster (summarized in Fig. 1). This pattern was conserved in several superfamilies, such as Hepialoidea, Tineoidea, Geometroidea, Noctuoidea, and Bombycoidea. In the superfamily Papilionoidea, one interstitial histone gene cluster was observed in nymphalids (Supplementary Fig. S7), whereas in representatives of the family Pieridae, the cluster was identified at the terminal region of a bivalent (Supplementary Fig. S6). This difference in position is most likely the result of an inversion as no additional clusters were identified. This inversion can be one of many chromosomal rearrangements which seem to be typical for the genus *Pieris*<sup>85,92</sup>. A single terminal histone gene cluster was also characteristic of the superfamilies Tischerioidea, Yponomeutoidea, and Cossoidea, although more species need to be tested in these taxa. Multiple clusters were observed only in three species. The caddisfly *G. pellucidus* had three terminal clusters on different bivalents (Supplementary Fig. S1b), whereas *T. absoluta* (Gelechioidea) had two clusters on both ends of a single bivalent (Supplementary Fig. S8f). In *H. cecropia* (Bombycoidea), three terminal clusters were present on two bivalents (Supplementary Fig. S11f). Taken together, the ancestral state of histone genes is probably a single interstitially located cluster. In some taxa, the cluster moved to the chromosome end, allowing its further spread to terminal regions of the same or other chromosomes, probably due to ectopic recombination (cf. Ref.<sup>29</sup>). The localization of the histone gene cluster seems to be very conserved in Lepidoptera, with the exception of *Leptidea* spp.<sup>65</sup>, and its changes indicate chromosomal rearrangements such as inversions, translocations, or chromosomal fusions and fissions (cf. Ref.<sup>65</sup>). Therefore, the histone H3 gene cluster is a good marker to study karyotype evolution in Lepidoptera.

Genes for 5S rRNA and U1 and U2 snRNAs have not yet been localized in lepidopteran genomes. Nine species sampled across Lepidoptera, namely *H. humuli* (Hepialidae), *T. ekebladella* (Tischeriidae), *T. bisselliella* (Tineidae), *T. tubulosa* (Psychidae), *C. ohridella* (Gracillariidae), *Y. evonymella* (Yponomeutidae), *C. pomonella* (Totricidae), *E. kuehniella* (Pyrilidae), and *B. mori* (Bombycidae) were analysed along with a trichopteran outgroup, *G. pellucidus* (Limnephilidae). The genes and corresponding probes were very short ( $\leq 140$  bp, Supplementary Table S4). Therefore, we used TSA-FISH, which allows the detection of single-copy genes  $\geq 1300$  bp<sup>27</sup>. Despite the optimization of the protocol, we were unable to localize the U1 and U2 snRNA genes in any of the species studied. The 5S rDNA clusters were detected only in the caddisfly *G. pellucidus* and representatives of early diverging lepidopteran lineages, namely *H. humuli*, *T. ekebladella*, and *T. tubulosa*. In all these species, TSA-FISH revealed a single interstitial 5S rDNA cluster (Fig. 1, Supplementary Fig. S12). It is tempting to speculate that the observed phylogenetic pattern could reflect a genome reorganization in Ditrysia, i.e. a lineage comprising 98% of extant moths and butterflies<sup>17</sup>, in which *Hox* gene amplification occurred<sup>93</sup> and over 1000 novel gene families emerged<sup>94</sup>. However, more data on the distribution of 5S rDNA in early diverging lineages is needed to confirm whether this pattern is consistent.

To find out why 5S rDNA, U1 and U2 snRNAs were not detected by FISH, we determined the copy number of the genes and tested whether the gene copies are arranged in tandem. Quantitative PCR revealed that copy numbers of 5S rRNA genes vary greatly between species (Fig. 2). An upper limit of the 5S rRNA gene copy number was observed in *H. humuli*, which may correlate with its likely large genome size. Although the genome size of *H. humuli* is unknown, the C-value of other hepialids, *Thitarodes* (*Hepialus*) sp. and *Triodia sylvina*, is 2.92 Gb<sup>95</sup> and 1.8 Gb<sup>96</sup>, respectively. However, copy number alone cannot explain the detectability of 5S rDNA in Lepidoptera. In *T. ekebladella*, we found approximately 20 copies of 5S rDNA genes, which we were able to detect by TSA-FISH, while we were not able to detect more than 100 copies of 5S rDNA genes in *B. mori*. In all species examined, the total length of the 5S rDNA cluster should be above the detection threshold of 1300 bp<sup>27</sup>, if all the copies are arranged in tandem. However, results of Southern hybridization revealed multiple bands with fragment length  $> 2000$  bp in all species (Supplementary Fig. S13), which suggests that the gene copies are scattered throughout the genome. Indeed, Vierna et al.<sup>43</sup> reported the presence of ten 5S rDNA clusters in the *B. mori* genome based on the analysis of genomic data. Alternatively, the copies can be only loosely clustered, i.e. separated by long spacers varying both in size and sequence. Clusters of 5S rDNA genes have been successfully mapped in many taxa<sup>38,42,97,98</sup>. However, the presence of multiple loci, which remain undetected even by TSA-FISH, questions homology of detected clusters and the usefulness of this marker in studies on karyotype evolution.

Copy number estimates for U1 and U2 snRNA genes by qPCR revealed much lower numbers than for 5S rDNA genes. These differences in copy number between 5S and U1 and U2 snRNAs seem to be consistent in Metazoa<sup>87,99–101</sup>. Based on our Southern hybridization results, the organization of U1 snRNA copies was quite similar to 5S rDNA, as multiple long fragments bearing the studied genes were observed (Supplementary Fig. S14). This means that successful detection of U1 and U2 snRNA clusters by FISH in some taxa<sup>35,46,48,102</sup> is the exception rather than the rule, and these genes are not universally applicable cytogenetic markers.

Taken together, 5S rDNA, U1 and U2 snRNA genes are not suitable markers for comparative cytogenetic studies in Lepidoptera. With a few exceptions, no clear cluster organization was detected by in situ hybridization. Their scattered organization and/or the presence of long spacer sequences between the genes does not allow for the observation of specific hybridization patterns and thereby the reconstruction of karyotype evolution. On the contrary, hybridization of 18S rDNA and histone H3 genes revealed a clustered organization of these genes in all species studied. Mapping of 18S rDNA showed rather dynamic evolution of the major rDNA, which does not always reflect chromosomal changes. However, various patterns, numbers, and locations of rDNA clusters could provide information on the evolution of repetitive sequences in lepidopteran genomes. Even though the mapping of histone H3 genes requires a species-specific probe preparation, hybridization patterns seem to genuinely reflect chromosomal rearrangements that occurred during the evolution of lepidopteran species. Our study shows that the evaluation of cytogenetic markers can significantly contribute to research focused on comparative cytogenetics and evolutionary genetics not only in Lepidoptera, but in all eukaryotic species.



## Material and methods

**Insects.** Examined lepidopteran species and one representative of caddisflies (Trichoptera), which was used as an outgroup, were either collected in the field or obtained from laboratory stocks. Some species were dissected immediately after collection. In the other species, captured females were left to lay eggs in plastic containers with host plants. Hatched larvae were then reared on their host plants or artificial diet. For a list of studied species, their origin, and details of rearing see Table S5.

**Chromosome preparations.** Meiotic and mitotic chromosomes from all the studied species were obtained from female and male gonads of 4th or 5th instar larvae. The only exception was *Gonepteryx rhamni*, which was dissected as young imago. Chromosomal preparations were made by spreading technique as described previously<sup>103</sup>. Briefly, dissections were performed in physiological solution<sup>104</sup>. The dissected gonads were hypotonized for 10 min (0.075 M KCl) and fixed in Carnoy's fixative (ethanol, chloroform, acetic acid; 6:3:1) for 15 min. They were then dissociated using tungsten needles in a drop of 60% acetic acid on a slide and spread using a heating plate set at 45 °C. Chromosome preparations were passed through an ethanol series (70%, 80% and 100% ethanol; 30–60 s each) and stored at – 20 °C or – 80 °C until further use.

**FISH with 18S rDNA probe.** A partial sequence of 18S rDNA was generated by PCR from male genomic DNA (gDNA) of the codling moth, *Cydia pomonella*, using a pair of specific primers as described previously<sup>25</sup> (Supplementary Table S4). This fragment was ligated into Promega pGem T-Easy Vector (Promega, Madison, WI, USA), cloned, purified by NucleoSpinPlasmid kit (Macherey–Nagel, Düren, Germany), verified by sequencing (SEQme, Dobříš, Czech Republic), and reamplified by PCR from plasmid. The reamplified 18S rDNA fragment was purified by the Wizard SV Gel and PCR Clean-Up System (Promega) and labelled by nick translation using Nick Translation Kit (Abbott Molecular Inc., Des Plaines, IL, USA) for 105 min at 15 °C. The 25 µL labelling reaction contained 500 ng DNA, 40 µM dATP, 40 µM dCTP, 40 µM dGTP, 14.4 µM dTTP, and 25.6 µM biotin-16-dUTP (Roche Diagnostics GmbH, Mannheim, Germany).

FISH experiments were carried out according to the previous study<sup>25</sup> with some modifications. Briefly, chromosome preparations were removed from the freezer, dehydrated in ethanol series, and air-dried. Preparations were treated with 100 µg/mL RNase A for 1 h at 37 °C to remove RNA and subsequently blocked in 5× Denhardt's solution for 30 min at 37 °C. In the next step, the slides were denatured in 70% formamide in 2×SSC for 3.5 min at 68 °C. After denaturation for 5 min at 90 °C, a probe mixture containing 25 ng of biotin-labelled 18S rDNA probe, 25 µg of sonicated salmon sperm, 50% deionized formamide, 10% dextran sulphate in 2×SSC in a total volume of 10 µL was applied to the slide and hybridized overnight at 37 °C. The biotin-labelled probe was detected by Cy3-conjugated streptavidin (diluted 1:1000 with blocking solution) (Jackson ImmunoRes. Labs. Inc, West Grove, PA, USA). Signals were amplified with biotinylated anti-streptavidin (diluted 1:25 with blocking solution) (Vector Labs. Inc, Burlingame, CA, USA), which was again detected by Cy3-conjugated streptavidin (diluted 1:1000 with blocking solution). The preparations were counterstained with 0.5 µg/mL of DAPI (4',6'-diamidino-2-phenylindole) and mounted in antifade containing DABCO (1,4-diazabicyclo[2.2.2]octane).

**FISH with tyramide signal amplification (TSA-FISH).** To obtain specific histone H3, 5S rDNA, and U1 and U2 snRNA probes for each species or family, fragments of the respective genes were amplified by PCR using degenerate primers (Supplementary Table S4) and gDNA of each individual species as a template, as detailed previously<sup>64</sup>. Species-specific amplified gene fragments were cloned and verified by sequencing (SEQme) (Supplementary Table S1). The verified plasmids were purified by NucleoSpin Plasmid kit (Macherey–Nagel) and used as template DNA to prepare a labelled probe by PCR. Each 25 µL labelling reaction contained 1–10 ng template DNA, 1× Ex Taq buffer, 1 mM each dATP, dCTP, and dGTP; 0.36 mM dTTP; 0.64 mM of fluorescein-12-dUTP (PerkinElmer, Waltham, MA, USA), 5 µmol of each primer, and 0.25 U TaKaRa Ex Taq DNA polymerase (TaKaRa, Otsu, Japan). Labelled probes were purified using Sephadex (Illustra Sephadex G-50 fine DNA grade). In *Inachis io* and *Tuta absoluta*, rDNA clusters were also mapped by TSA FISH with 18S rDNA probe labelled by fluorescein instead of FISH with biotin-labelled probe described above. Species-specific probes were generated for most species, except for *Phymatopus californicus*, to which we hybridized a probe from *Hepialus humuli*, *Psyche crassiorella* with a probe from *Taleporia tubulosa*, and *Pieris brassicae* with a probe from *Pieris rapae*.

TSA-FISH was performed according to the published protocol<sup>27</sup> with some modifications. Briefly, frozen chromosome preparations were dehydrated using an ethanol series. After drying, slides were treated with 10 mM HCl for 10 min at 37 °C to remove cytoplasm and incubated in 1% hydrogen peroxide for 30 min at room temperature to quench endogenous peroxidase activity. Preparations were digested with 100 µg/mL RNase A for 1 h at 37 °C and blocked with 5× Denhardt's solution for 30 min at 37 °C. Thereafter, a 50 µL probe mixture containing 10–30 ng of labelled specific probe in 50% deionized formamide and 10% dextran sulfate in 2×SSC was added to the slide, and the probe and chromosomes were simultaneously denatured for 5 min at 70 °C. Hybridization took place overnight at 37 °C. Hybridization signals were enhanced by Antifluorescein-HRP conjugate (PerkinElmer) diluted 1:1000 and incubated with tyramide solution (TSA Plus Fluorescein system, PerkinElmer) for 10–15 min for 5S rDNA and U1 and U2 snRNA and 5–7 min for histone H3. The preparations were counterstained and mounted in antifade containing DABCO with 0.5 µg/mL of DAPI.

**Microscopy and image processing.** Observation of chromosome preparations from FISH experiments was performed with a Zeiss Axioplan 2 microscope (Carl Zeiss, Jena, Germany) equipped with appropriate fluorescence filter sets. An Olympus CCD monochrome camera XM10 equipped with cellSens 1.9 digital imaging software (Olympus Europa Holding, Hamburg, Germany) was used to record and capture black-and-white

pictures. Images were captured separately for each fluorescent dye and then pseudocoloured and superimposed with Adobe Photoshop CS4, version 11.0.

**Quantitative analysis of gene doses.** Quantitative PCR (qPCR) was used to estimate relative copy numbers of three target genes, namely U1 and U2 snRNA and 5S rDNA, in *Glyphotaelius pellucidus*, *Hepialus humuli*, *Tischeria ekebladella*, *Taleporia tubulosa*, *Tineola bisselliella*, *Cameraria ohridella*, *Yponomeuta evonymella*, *Ephestia kuehniella*, *Cydia pomonella*, and *Bombyx mori*<sup>19,76</sup>. By comparing the genes of interest to a single-copy autosomal reference gene (*Acetylcholinesterase 2*, *Ace2*), their relative copy numbers were estimated based on a target to reference gene dose ratio formula (Ref.<sup>105</sup>; see below). The reference gene and genes of interest were analysed simultaneously in technical triplicates of three independent biological replicas. Due to small body size of some species, namely *T. ekebladella*, *T. tubulosa*, *T. bisselliella*, and *C. ohridella*, 5–10 individuals were pooled for gDNA extraction carried out using NucleoSpin Tissue kit (Macherey–Nagel), DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany), or NucleoSpin DNA Insect kit (Macherey–Nagel) following manufacturer's instructions. One individual per biological replica was used for the other species.

The qPCR contained 1–10 ng of gDNA, an optimized concentration of primers per species (details in Supplementary Table S6) and Xceed qPCR SG Mix Lo-ROX (Institute of Applied Biotechnologies, Prague, Czech Republic) in a total volume of 10  $\mu$ L. Amplification efficiencies ( $E$ ) for each gene and species were determined by 0 $\times$ , 5 $\times$ , 25 $\times$ , and 125 $\times$  dilutions of pooled gDNA of all biological replicas. For all three markers in *C. pomonella* and 5S rDNA in *B. mori* SYBR Premix Ex Taq II Perfect Real Time (1 $\times$ ; TaKaRa) was used and amplification efficiencies were determined by 0 $\times$ , 10 $\times$ , 100 $\times$  and 1000 $\times$  dilutions (details in Supplementary Table S6). The experiments were carried out using the C1000 Thermal cycler CFX96 Real-Time System (Bio-Rad, Hercules, CA, USA) and data were analysed using software Bio-Rad CFX Manager 3.1. The target to reference gene dose ratio was calculated for each biological sample according the formula  $R = [(1 + E_{\text{Reference}})^{Ct_{\text{Reference}}}] / [(1 + E_{\text{Target}})^{Ct_{\text{Target}}}]$ , where  $R$  is a relative copy number of target gene,  $E$  is the primer efficiency and  $Ct =$  cycle threshold<sup>105</sup>.

**Southern hybridization.** Southern hybridization was performed to independently estimate the copy number of U1 snRNA and 5S rDNA and to test whether the genes are tandemly arranged in the genomes of ten selected species. Cloned fragments of studied genes were reamplified by PCR using degenerate primers (Supplementary Table S4) and the products were used as template for labelling with digoxigenin-11-dUTP (Roche Diagnostics GmbH). Labelling and purification of the probes were done as for TSA-FISH probes (see above).

High-molecular-weight gDNA of the studied species was extracted by standard phenol–chloroform<sup>106</sup> or by cetyltrimethylammonium bromide (CTAB)<sup>107</sup> extraction. Three pairs of restriction enzymes with no restriction sites within the target sequences were selected (Supplementary Table S7) and digestion of gDNA was carried out overnight at 37  $^{\circ}$ C. Enzymes were inactivated by addition of loading buffer (50% glycerol, 250 mM EDTA, 5.9 mM bromophenol blue) or Gel Loading Dye, Purple (6 $\times$ ) (New England Biolabs, Ipswich, MA, USA) (for details see Supplementary Table S7). Five micrograms of digested DNA per well was separated using a 1% agarose gel in 1 $\times$  TBE buffer by horizontal electrophoresis at 5 V/cm. Southern hybridizations were carried out according to the published protocol<sup>108</sup> with some modifications. Briefly, after electrophoretic separation, DNA was denatured and transferred onto an Amersham Hybond-N + nylon membrane (GE Healthcare, Buckinghamshire, UK) by capillary flow. Hybridization of labelled probes (100 ng) was done overnight at 42  $^{\circ}$ C and the stringent washes on the subsequent day were performed at 68  $^{\circ}$ C. Probes were detected using Anti-Digoxigenin-AP (75 mU/mL; Roche Diagnostics GmbH) incubated with CDP-*Star* ready-to-use (Roche Diagnostics GmbH). Resulting chemiluminescence was recorded with a LAS-3000 Lumi-Imager (Fuji Photo Film Europe GmbH, Düsseldorf, Germany).

### Data availability

All data generated or analysed during this study are included in this published article [and its Supplementary Information files]. Partial sequences of genes under study were deposited in GenBank under the acc. nos. MW149037–MW149046, MW194851–MW194870, and MW558903–MW558929.

Received: 26 March 2021; Accepted: 24 May 2021

Published online: 09 June 2021

### References

1. Traut, W., Ahola, V., Smith, D. A. S., Gordon, I. J. & French-Constant, R. H. Karyotypes versus genomes: The nymphalid butterflies *Melitaea cinxia*, *Danaus plexippus*, and *D. chrysippus*. *Cytogenet. Genome Res.* **153**, 46–53 (2017).
2. Heng, H. H. *et al.* A postgenomic perspective on molecular cytogenetics. *CG* **19**, 227–239 (2018).
3. Deakin, J. E. *et al.* Chromosomics: Bridging the gap between genomes and chromosomes. *Genes* **10**, 627. <https://doi.org/10.3390/genes10080627> (2019).
4. Lewin, H. A. *et al.* Earth BioGenome Project: Sequencing life for the future of life. *PNAS* **115**, 4325–4333 (2018).
5. Clausen, U. Chromosomics. *Cytogenet. Genome Res.* **111**, 101–106 (2005).
6. Graphodatsky, A. S. Comparative chromosomics. *Mol. Biol.* **41**, 361–375 (2007).
7. Dutrillaux, B., Couturier, J., Richer, C.-L. & Viegas-Péquignot, E. Sequence of DNA replication in 277 R- and Q-bands of human chromosomes using a BrdU treatment. *Chromosoma* **58**, 51–61 (1976).
8. Bickmore, W. A. Karyotype analysis and chromosome banding. In *Encyclopedia of Life Sciences* (ed. Wiley) a0001160 (Wiley, 2001).
9. King, M. C-Banding studies on Australian hyliid frogs: Secondary constriction structure and the concept of euchromatin transformation. *Chromosoma* **80**, 191–217 (1980).
10. Steiniger, G. E. & Mukherjee, A. B. Insect chromosome banding: technique for G- and Q-banding pattern in the mosquito *Aedes albopictus*. *Can. J. Genet. Cytol.* **17**, 241–244 (1975).
11. Brum-Zorrilla, N. & Postiglioni, A. Karyological studies on Uruguayan spiders I. Banding pattern in chromosomes of *Lycosa* species (Araneae-Lycosidae). *Genetica* **54**, 149–153 (1980).

12. Cabrero, J. & Camacho, J. P. M. Cytogenetic studies in gomphocerine grasshoppers. I. Comparative analysis of chromosome C-banding pattern. *Heredity* **56**, 365–372 (1986).
13. Chen, R., Song, W., Li, X. & An, Z. Chromosome G-banding in plants by inducing with trypsin and urea. *Cell Res.* **4**, 79–87 (1994).
14. Song, Y. C. *et al.* Comparisons of G-banding patterns in six species of the Poaceae. *Hereditas* **121**, 31–38 (2004).
15. Bedo, D. G. Karyotypic and chromosome banding studies of the potato tuber moth, *Phthorimaea operculella* (Zeller) (Lepidoptera, Gelechiidae). *Can. J. Genet. Cytol.* **26**, 141–145 (1984).
16. Prins, J. D. & Saitoh, K. Karyology and sex determination. In *Band 4: Arthropoda, 2 Hälfte: Insecta, Lepidoptera, Moths and Butterflies, Teilband/Part 36, Vol 2: Morphology, Physiology, and Development* (ed. Kükenthal, W.) (DE GRUYTER, 2003).
17. Van Nieuwerkerken, E. J. *et al.* Order Lepidoptera Linnaeus, 1758. In: Zhang, Z.-Q. (Ed.) Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa* **3148**, 212. <https://doi.org/10.11646/zootaxa.3148.1.41> (2011).
18. Yoshido, A. *et al.* Evolution of multiple sex-chromosomes associated with dynamic genome reshuffling in Leptidea wood-white butterflies. *Heredity* **125**, 138–154 (2020).
19. Nguyen, P. *et al.* Neo-sex chromosomes and adaptive potential in tortricid pests. *PNAS* **110**, 6931–6936 (2013).
20. Picq, S. *et al.* Insights into the structure of the spruce budworm (*Choristoneura fumiferana*) genome, as revealed by molecular cytogenetic analyses and a high-density linkage map. *G3* **8**, 2539–2549 (2018).
21. Asser-Kaiser, S. *et al.* Rapid emergence of baculovirus resistance in codling moth due to dominant, sex-linked inheritance. *Science* **317**, 1916–1918 (2007).
22. Carabajal Paladino, L. Z. *et al.* Sex chromosome turnover in moths of the diverse superfamily Gelechioidea. *Genome Biol. Evol.* **11**, 1307–1319 (2019).
23. Labbé, R., Caveney, S. & Donly, C. Genetic analysis of the xenobiotic resistance-associated ABC gene subfamilies of the Lepidoptera. *Insect Mol. Biol.* **20**, 243–256 (2010).
24. Wolf-I, K. W., Novák, K. & Marec, F. Kinetic organization of metaphase I bivalents in spermatogenesis of Lepidoptera and Trichoptera species with small chromosome numbers. *Heredity* **79**, 35–143 (1997).
25. Fuková, I., Nguyen, P. & Marec, F. Codling moth cytogenetics: Karyotype, chromosomal location of rDNA, and molecular differentiation of sex chromosomes. *Genome* **48**, 1083–1092 (2005).
26. Gan, Y. *et al.* Individual chromosome identification, chromosomal collinearity and genetic-physical integrated map in *Gossypium darwinii* and four D genome cotton species revealed by BAC-FISH. *Genes Genet. Syst.* **87**, 233–241 (2012).
27. Carabajal Paladino, L. Z., Nguyen, P., Šichová, J. & Marec, F. Mapping of single-copy genes by TSA-FISH in the codling moth, *Cydia pomonella*. *BMC Genet.* **15**, S15. <https://doi.org/10.1186/1471-2156-15-S2-S15> (2014).
28. Traut, W. Pachytene mapping in the female silkworm, *Bombyx mori* L. (Lepidoptera). *Chromosoma* **58**, 275–284 (1976).
29. Nguyen, P., Sahara, K., Yoshido, A. & Marec, F. Evolutionary dynamics of rDNA clusters on chromosomes of moths and butterflies (Lepidoptera). *Genetica* **138**, 343–354 (2010).
30. Vershina, A. O., Anokhin, B. A. & Lukhtanov, V. A. Ribosomal DNA clusters and telomeric (TTAGG)<sub>n</sub> repeats in blue butterflies (Lepidoptera, Lycaenidae) with low and high chromosome numbers. *CCG* **9**, 161–171 (2015).
31. Yoshido, A., Bando, H., Yasukochi, Y. & Sahara, K. The *Bombyx mori* karyotype and the assignment of linkage groups. *Genetics* **170**, 675–685 (2005).
32. Lukhtanov, V. A. *et al.* Versatility of multivalent orientation, inverted meiosis, and rescued fitness in holocentric chromosomal hybrids. *PNAS* **115**, E9610–E9619 (2018).
33. Lee, Y.-I. *et al.* Satellite DNA in *Paphiopedilum* subgenus *Parvisepalum* as revealed by high-throughput sequencing and fluorescent in situ hybridization. *BMC Genom.* **19**, 578. <https://doi.org/10.1186/s12864-018-4956-7> (2018).
34. Cabral-de-Mello, D. C., Martins, C., Souza, M. J. & Moura, R. C. Cytogenetic mapping of 5S and 18S rRNAs and H3 histone genes in 4 ancient Proscopiidae grasshopper species: Contribution to understanding the evolutionary dynamics of multigene families. *Cytogenet. Genome Res.* **132**, 89–93 (2011).
35. García-Souto, D., Troncoso, T., Pérez, M. & Pasantes, J. J. Molecular cytogenetic analysis of the European hake *Merluccius merluccius* (Merlucciidae, Gadiformes): U1 and U2 snRNA gene clusters map to the same location. *PLoS One* **10**, e0146150. <https://doi.org/10.1371/journal.pone.0146150> (2015).
36. Kobayashi, T., Heck, D. J., Nomura, M. & Horiuchi, T. Expansion and contraction of ribosomal DNA repeats in *Saccharomyces cerevisiae*: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. *Genes Dev.* **12**, 3821–3830 (1998).
37. Prokopowich, C. D., Gregory, T. R. & Crease, T. J. The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **46**, 48–50 (2003).
38. Cabral-de-Mello, D. C., Moura, R. C. & Martins, C. Chromosomal mapping of repetitive DNAs in the beetle *Dichotomius geminatus* provides the first evidence for an association of 5S rRNA and histone H3 genes in insects, and repetitive DNA similarity between the B chromosome and A complement. *Heredity* **104**, 393–400 (2010).
39. Cabral-de-Mello, D. C., Oliveira, S. G., de Moura, R. C. & Martins, C. Chromosomal organization of the 18S and 5S rRNAs and histone H3 genes in *Scarabaeinae coleopterans*: Insights into the evolutionary dynamics of multigene families and heterochromatin. *BMC Genet.* **12**, 88. <https://doi.org/10.1186/1471-2156-12-88> (2011).
40. Badaeva, E. D. *et al.* A set of cytogenetic markers allows the precise identification of all A-genome chromosomes in diploid and polyploid wheat. *Cytogenet. Genome Res.* **146**, 71–79 (2015).
41. Wang, W. *et al.* Remarkable variation of ribosomal DNA organization and copy number in gnetophytes, a distinct lineage of gymnosperms. *Ann. Bot.* **123**, 767–781 (2019).
42. Palacios-Gimenez, O. M., Castillo, E. R., Martí, D. A. & Cabral-de-Mello, D. C. Tracking the evolution of sex chromosome systems in Melanoplinae grasshoppers through chromosomal mapping of repetitive DNA sequences. *BMC Evol. Biol.* **13**, 167. <https://doi.org/10.1186/1471-2148-13-167> (2013).
43. Vierna, J., Wehner, S., Höner zu Siederdisen, C., Martínez-Lage, A. & Marz, M. Systematic analysis and evolution of 5S ribosomal DNA in metazoans. *Heredity* **111**, 410–421 (2013).
44. Garcia, S., Garnatje, T. & Kovařík, A. Plant rDNA database: Ribosomal DNA loci information goes online. *Chromosoma* **121**, 389–394 (2012).
45. Sochorová, J., Garcia, S., Gálvez, F., Symonová, R. & Kovařík, A. Evolutionary trends in animal ribosomal DNA loci: Introduction to a new online database. *Chromosoma* **127**, 141–150 (2018).
46. Anjos, A. *et al.* U1 snDNA clusters in grasshoppers: Chromosomal dynamics and genomic organization. *Heredity* **114**, 207–219 (2015).
47. Barzotti, R., Pelliccia, F. & Rocchi, A. Identification and characterization of U1 small nuclear RNA genes from two crustacean isopod species. *Chromosome Res.* **11**, 365–373 (2003).
48. Cabral-de-Mello, D. C., Valente, G. T., Nakajima, R. T. & Martins, C. Genomic organization and comparative chromosome mapping of the U1 snRNA gene in cichlid fish, with an emphasis in *Oreochromis niloticus*. *Chromosome Res.* **20**, 279–292 (2012).
49. Merlo, M. A. *et al.* Analysis of the histone cluster in Senegalese sole (*Solea senegalensis*): Evidence for a divergent evolution of two canonical histone clusters. *Genome* **60**, 441–453 (2017).

50. Piscor, D., Fernandes, C. A. & Parise-Maltempi, P. P. Conserved number of U2 snDNA sites in *Piabina argentea*, *Piabarchus stramineus* and two *Bryconamericus* species (Characidae, Stevardiinae). *Neotrop. Ichthyol.* **16**, e170066. <https://doi.org/10.1590/1982-0224-20170066> (2018).
51. Sember, A. *et al.* Dynamics of tandemly repeated DNA sequences during evolution of diploid and tetraploid botiid loaches (Teleostei: Cobitoidea: Botiidae). *PLoS One* **13**, e0195054. <https://doi.org/10.1371/journal.pone.0195054> (2018).
52. Poletto, A. B. *et al.* Chromosome differentiation patterns during cichlid fish evolution. *BMC Genet.* **11**, 50. <https://doi.org/10.1186/1471-2156-11-50> (2010).
53. Symonová, R. *et al.* Genome differentiation in a species pair of coregonine fishes: An extremely rapid speciation driven by stress-activated retrotransposons mediating extensive ribosomal DNA multiplications. *BMC Evol. Biol.* **13**, 42. <https://doi.org/10.1186/1471-2148-13-42> (2013).
54. de Sene, V. F. *et al.* Mapping of the retrotransposable elements *Rex1* and *Rex3* in chromosomes of *Eigenmannia* (Teleostei, Gymnotiformes, Sternopygidae). *Cytogenet. Genome Res.* **146**, 319–324 (2015).
55. Cai, Q., Zhang, D., Liu, Z.-L. & Wang, X.-R. Chromosomal localization of 5S and 18S rDNA in five species of subgenus *Strobos* and their implications for genome evolution of *Pinus*. *Ann. Bot.* **97**, 715–722 (2006).
56. Grozeva, S., Kuznetsova, V. & Anokhin, B. Karyotypes, male meiosis and comparative FISH mapping of 18S ribosomal DNA and telomeric (TTAGG)<sub>n</sub> repeat in eight species of true bugs (Hemiptera, Heteroptera). *CCG* **5**, 355–374 (2011).
57. Chirino, M. G. & Bressa, M. J. Karyotype evolution in progress: A new diploid number in *Belostoma candidulum* (Heteroptera: Belostomatidae) from Argentina leading to new insights into its ecology and evolution. *Eur. J. Entomol.* **111**, 165–174 (2014).
58. Roehrdanz, R., Heilmann, L., Senechal, P., Sears, S. & Evenson, P. Histone and ribosomal RNA repetitive gene clusters of the boll weevil are linked in a tandem array: Histone-ribosomal DNA repeats. *Insect Mol. Biol.* **19**, 463–471 (2010).
59. Schienman, J. E., Lozovskaya, E. R. & Strausbaugh, L. D. *Drosophila virilis* has atypical kinds and arrangements of histone repeats. *Chromosoma* **107**, 529–539 (1998).
60. Zhang, L., Bao, Z., Wang, S., Huang, X. & Hu, J. Chromosome rearrangements in Pectinidae (Bivalvia: Pteriomorpha) implied based on chromosomal localization of histone H3 gene in four scallops. *Genetica* **130**, 193–198 (2007).
61. Silva, D. M. Z. A. *et al.* Chromosomal organization of repetitive DNA sequences in *Astyanax bockmanni* (Teleostei, Characiformes): Dispersive location, association and co-localization in the genome. *Genetica* **141**, 329–336 (2013).
62. Maxson, R., Cohn, R., Kedes, L. & Mohun, T. Expression and organization of histone genes. *Annu. Rev. Genet.* **17**, 239–277 (1983).
63. Eirín-López, J. M. *et al.* Molecular evolutionary characterization of the mussel *Mytilus* histone multigene family: First record of a tandemly repeated unit of five histone genes containing an H1 subtype with 'Orphan' features. *J. Mol. Evol.* **58**, 131–144 (2004).
64. Šichová, J., Nguyen, P., Dalíková, M. & Marec, F. Chromosomal evolution in tortricid moths: Conserved karyotypes with diverged features. *PLoS One* **8**, e64520. <https://doi.org/10.1371/journal.pone.0064520> (2013).
65. Šichová, J. *et al.* Dynamic karyotype evolution and unique sex determination systems in *Leptidea* wood white butterflies. *BMC Evol. Biol.* **15**, 89. <https://doi.org/10.1186/s12862-015-0375-4> (2015).
66. Mutanen, M., Wahlberg, N. & Kaila, L. Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. *Proc. R. Soc. B* **277**, 2839–2848 (2010).
67. Wiemers, M., Chazot, N., Wheat, C., Schweiger, O. & Wahlberg, N. A complete time-calibrated multi-gene phylogeny of the European butterflies. *ZooKeys* **938**, 97–124 (2020).
68. Kawahara, A. Y. *et al.* Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *PNAS* **116**, 22657–22663 (2019).
69. Marec, F. & Traut, W. Synaptonemal complexes in female and male meiotic prophase of *Ephestia kuehniella* (Lepidoptera). *Heredity* **71**, 394–404 (1993).
70. Zrzavá, M. *et al.* Sex chromosomes of the iconic moth *Abraxas grossulariata* (Lepidoptera, Geometridae) and its congener *A. sylvata*. *Genes* **9**, 279. <https://doi.org/10.3390/genes9060279> (2018).
71. Kiauta, B. & Lankhorst, L. The chromosomes of the caddis-fly, *Glyptotaelius pellucidus* (Retzius, 1783) (Trichoptera: Limnephilidae, Limnephilinae). *Genetica* **40**, 1–6 (1969).
72. Lukhtanov, V. A. Sex chromatin and sex chromosome systems in nonditrysian Lepidoptera (Insecta). *J. Zool. Syst. Evol. Res.* **38**, 73–79 (2000).
73. Seiler, J. Research on the sex-chromosomes of Psychidae (Lepidoptera). *Biol. Bull.* **36**, 399–404 (1919).
74. Seiler, J. Geschlechtschromosomen-Untersuchungen an Psychiden. *Z. Indukt. Abstamm. Vererbungsl.* **31**, 1–99 (1922).
75. Hejníčková, M. *et al.* Absence of W chromosome in Psychidae moths and implications for the theory of sex chromosome evolution in Lepidoptera. *Genes* **10**, 1016. <https://doi.org/10.3390/genes10121016> (2019).
76. Dalíková, M. *et al.* New insights into the evolution of the W chromosome in Lepidoptera. *J. Hered.* **108**, 709–719 (2017).
77. Prins, J. D., Prins, W. D. & Dall'Asta, U. The karyotype of *Cameraria ohridella* (Lepidoptera: Gracillariidae). *Phegea* **301**, 5–10 (2002).
78. Kawazoé, A. The chromosome in the primitive or microlepidopterous moth-groups II. *Proc. Jpn. Acad. Ser. B* **63**, 87–90 (1987).
79. Nilsson, N.-O., Löfstedt, C. & Dävring, L. Unusual sex chromosome inheritance in six species of small ermine moths (Yponomeuta, Yponomeutidae, Lepidoptera). *Hereditas* **108**, 259–265 (1988).
80. Ortiz, E. & Templado, J. Los cromosomas de tres especies de tortricidos (Lep. Tortricidae). *EOS Rev. Esp. Entomol.* **51**, 77–84 (1976).
81. Robinson, R. *Lepidoptera Genetics*. (Pergamon Press, 1971).
82. Carabajal Paladino, L. Z. *et al.* The effect of X-rays on cytological traits of *Tuta absoluta* (Lepidoptera: Gelechiidae). *Fla. Entomol.* **99**, 43–53 (2016).
83. Schulz, H.-J. & Traut, W. The pachytene complement of the wildtype and a chromosome mutant strain of the flour moth, *Ephestia kuehniella* (Lepidoptera). *Genetica* **50**, 61–66 (1979).
84. Van't Hof, A. E. *et al.* Linkage map of the peppered moth, *Biston betularia* (Lepidoptera, Geometridae): A model of industrial melanism. *Heredity* **110**, 283–295 (2013).
85. Kawamura, N. Cytological studies on the mosaic silkworms induced by low temperature treatment. *Chromosoma* **74**, 179–188 (1979).
86. Bytinski-Salz, H. Untersuchungen an Lepidopteren-hybriden V. Die Verwandtschaft der *Platysamia*-Arten (Lepidoptera, Saturniidae) nach Untersuchungen über die Fertilität und die Chromosomenverhältnisse ihrer Bastarde. *Arch. Exp. Zellforsch* **22**, 217–237 (1938).
87. Ahola, V. *et al.* The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nat. Commun.* **5**, 4737. <https://doi.org/10.1038/ncomms5737> (2014).
88. da Silva, M., Ribeiro, E. D. & Matoso, D. A. Chromosomal polymorphism in two species of *Hypancistrus* (Siluriformes: Loricariidae): An integrative approach for understanding their biodiversity. *Genetica* **142**, 127–139 (2014).
89. Supiwong, W. *et al.* Karyotype diversity and evolutionary trends in the Asian swamp eel *Monopterus albus* (Synbranchiformes, Synbranchidae): A case of chromosomal speciation?. *BMC Evol. Biol.* **19**, 73 (2019).
90. Šichová, J. *et al.* Fissions, fusions, and translocations shaped the karyotype and multiple sex chromosome constitution of the northeast-Asian wood white butterfly, *Leptidea amurensis*. *Biol. J. Linn. Soc.* **118**, 457–471 (2016).



91. Lukhtanov, V. A., Dincă, V., Talavera, G. & Vila, R. Unprecedented within-species chromosome number cline in the Wood White butterfly *Leptidea sinapis* and its significance for karyotype evolution and speciation. *BMC Evol. Biol.* **11**, 109. <https://doi.org/10.1186/1471-2148-11-109> (2011).
92. Hill, J. *et al.* Unprecedented reorganization of holocentric chromosomes provides insights into the enigma of lepidopteran chromosome evolution. *Sci. Adv.* **5**, eaau3648. <https://doi.org/10.1126/sciadv.aau3648> (2019).
93. Ferguson, L. *et al.* Ancient expansion of the Hox cluster in Lepidoptera generated four Homeobox genes implicated in extra-embryonic tissue formation. *PLoS Genet.* **10**, e1004698. <https://doi.org/10.1371/journal.pgen.1004698> (2014).
94. Thomas, G. W. C. *et al.* Gene content evolution in the arthropods. *Genome Biol.* **21**, 15. <https://doi.org/10.1186/s13059-019-1925-7> (2020).
95. Cheng, R.-L., Yu, Y.-X., Liu, L.-X., Zhang, C.-X. & Fang, C.-X. A draft genome of the ghost moth, *Thitarodes (Hepialus)* sp., a medicinal caterpillar fungus. *Insect Sci.* **23**, 326–329 (2016).
96. Fraïsse, C., Picard, M. A. L. & Vicoso, B. The deep conservation of the Lepidoptera Z chromosome suggests a non-canonical origin of the W. *Nat. Commun.* **8**, 1486. <https://doi.org/10.1038/s41467-017-01663-5> (2017).
97. Aguilera, P. M., Debat, H. J., Scaldaferrro, M. A., Martí, D. A. & Grabielle, M. FISH-mapping of the 5S rDNA locus in chili peppers (*Capsicum-Solanaceae*). *An. Acad. Bras. Ciênc.* **88**, 117–125 (2016).
98. Volkov, R. A. *et al.* Evolutional dynamics of 45S and 5S ribosomal DNA in ancient allohexaploid *Atropa belladonna*. *BMC Plant Biol.* **17**, 21. <https://doi.org/10.1186/s12870-017-0978-6> (2017).
99. Mount, S. M., Gotea, V., Lin, C.-F., Hernandez, K. & Makalowski, W. Spliceosomal small nuclear RNA genes in 11 insect genomes. *RNA* **13**, 5–14 (2006).
100. Marz, M., Kirsten, T. & Stadler, P. F. Evolution of spliceosomal snRNA genes in metazoan animals. *J. Mol. Evol.* **67**, 594–607 (2008).
101. Mesa, A., Somarelli, J. A. & Herrera, R. J. Small nuclear RNA variants of three *Bombyx mori* strains. *Entomol. Res.* **38**, 61–68 (2008).
102. Merlo, M. A., Cross, I., Rodríguez-Rúa, A., Manchado, M. & Rebordinos, L. First approach to studying the genetics of the meagre (*Argyrosomus regius*; Asso, 1801) using three multigene families. *Aquac. Res.* **44**, 974–984 (2013).
103. Mediouni, J., Fuková, I., Frydrychová, R., Dhoubi, M. H. & Marec, F. Karyotype, sex chromatin and sex chromosome differentiation in the carob moth, *Ectomyelois ceratoniae* (Lepidoptera: Pyralidae). *Caryologia* **57**, 184–194 (2004).
104. Lockwood, A. “Ringer” solutions and some notes on the physiological basis of their ionic composition. *Comp. Biol. Phys.* **2**, 241–289 (1961).
105. Rovatsos, M., Altmanová, M., Pokorná, M. & Kratochvíl, L. Conserved sex chromosomes across adaptively radiated *Anolis* lizards. *Evolution* **68**, 2079–2085 (2014).
106. Blin, N. & Stafford, D. W. A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Res.* **3**, 2303–2308 (1976).
107. Ferguson, K. B. *et al.* Jekyll or Hyde? The genome (and more) of *Nesidiocoris tenuis*, a zoophytophagous predatory bug that is both a biological control agent and a pest. *Insect Mol. Biol.* **30**, 188–209 (2020).
108. Fuková, I. *et al.* Probing the W chromosome of the codling moth, *Cydia pomonella*, with sequences from microdissected sex chromatin. *Chromosoma* **116**, 135–145 (2007).

## Acknowledgements

We thank Marie Korchová for technical assistance. We further thank David Doležel, Peter Roessingh, Ivo Šauman, Hana Sehadová, and Atsuo Yoshido for material and Jan Provazník for help with data visualization.

## Author contributions

P.N. conceived the study. P.N., M.D., and F.M. designed experiments. I.P., M.H., S.V., A.V., L.Z.C.P., and M.Z. performed research. L.Z.C.P. contributed an optimized protocol of TSA-FISH. I.P., M.D., S.V., A.V., M.Z., and P.N. analysed data. I.P. created the figures. I.P. and P.N. wrote the manuscript. All authors read and approved the manuscript.

## Funding

This research was funded by the Grant 17-17211S of the Czech Science Foundation (CSF) to P.N. Support from CSF Grants 17-13713S and 20-13784S given to FM is also acknowledged.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-91665-7>.

**Correspondence** and requests for materials should be addressed to P.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021