

Sex-biased gene content is associated with sex chromosome turnover in Danaini butterflies

Pablo Mora¹ | Monika Hospodářská^{1,2} | Anna Chung Voleníková¹ | Petr Koutecký¹ | Jana Štundlová¹ | Martina Dalíková^{1,3} | James R. Walters³ | Petr Nguyen^{1,2} 

¹Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

²Institute of Entomology, Biology Centre of the Czech Academy of Sciences, České Budějovice, Czech Republic

³Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, Kansas, USA

Correspondence

Petr Nguyen, Institute of Entomology, Biology Centre of the Czech Academy of Sciences, České Budějovice, Czech Republic.

Email: petr.nguyen@entu.cas.cz

James R. Walters, Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, KS, USA.

Email: jrwalters@ku.edu

Present address

Pablo Mora, University of Jaén, Jaén, Spain

Handling Editor: Hannah Augustijnen

Abstract

Sex chromosomes play an outsized role in adaptation and speciation, and thus deserve particular attention in evolutionary genomics. In particular, fusions between sex chromosomes and autosomes can produce neo-sex chromosomes, which offer important insights into the evolutionary dynamics of sex chromosomes. Here, we investigate the evolutionary origin of the previously reported *Danaus* neo-sex chromosome within the tribe Danaini. We assembled and annotated genomes of *Tirumala septentrionis* (subtribe Danaina), *Ideopsis similis* (Amaurina), *Idea leuconoe* (Euploina) and *Lycorea halia* (Itunina) and identified their Z-linked scaffolds. We found that the *Danaus* neo-sex chromosome resulting from the fusion between a Z chromosome and an autosome corresponding to the *Melitaeta cinxia* chromosome (McChr) 21 arose in a common ancestor of Danaina, Amaurina and Euploina. We also identified two additional fusions as the W chromosome further fused with the synteny block McChr31 in *I. similis* and independent fusion occurred between ancestral Z chromosome and McChr12 in *L. halia*. We further tested a possible role of sexually antagonistic selection in sex chromosome turnover by analysing the genomic distribution of sex-biased genes in *I. leuconoe* and *L. halia*. The autosomes corresponding to McChr21 and McChr31 involved in the fusions are significantly enriched in female- and male-biased genes, respectively, which could have hypothetically facilitated fixation of the neo-sex chromosomes. This suggests a role of sexual antagonism in sex chromosome turnover in Lepidoptera. The neo-Z chromosomes of both *I. leuconoe* and *L. halia* appear fully compensated in somatic tissues, but the extent of dosage compensation for the ancestral Z varies across tissues and species.

KEYWORDS

butterflies, dosage compensation, fusions, sex chromosomes, sex-biased genes, sexual antagonism

1 | INTRODUCTION

Sex chromosomes are a dynamic part of many genomes, often distinguished from autosomes by a wide range of features resulting from distinct evolutionary processes. Typical X and Y chromosomes (and comparably, the Z and W in female-heterogametic taxa) each differ from autosomes, sometimes dramatically, in their gene content, sequence composition, recombination rates, population sizes and sex-specific inheritance (Vicoso & Charlesworth, 2006). Accordingly, sex chromosomes do not just determine sex (Furman et al., 2020; Kiuchi et al., 2014), but play an outsized role in major biological processes such as sexual conflict (Lund-Hansen et al., 2021; Mank et al., 2014), adaptation (Lasne et al., 2017; McAllister et al., 2008) and speciation (Payseur et al., 2018). Therefore, discovering and describing the sex chromosome content in diverse taxa is a long-standing objective for both genomics and evolutionary biology alike.

Importantly, such efforts may identify turnovers in sex chromosomes, which present excellent opportunities to investigate their evolutionary dynamics (Bachtrog, 2013; Palmer et al., 2019; Pan et al., 2021; Sigeman et al., 2019). Often such transitions involve a fusion of an autosome with a sex chromosome, resulting in so-called neo-sex chromosomes that recapitulate the X and Y (or Z and W) differentiation through restriction of recombination and pseudogenization and accumulation of repetitive sequences in non-recombining region (Wright et al., 2016). As such, they offer novel insights into this complex evolutionary process. Moreover, the neo-sex chromosomes are suggested to promote species diversification (Filatov, 2018; Graves, 2016; Kitano & Peichel, 2012; Yoshido et al., 2020).

Sex chromosome-autosome fusions can spread in populations due to genetic drift, meiotic drive or selective advantage (Charlesworth & Charlesworth, 1980; Lucek et al., 2023). Selective advantage may be conferred by new linkage between sex and sexually dimorphic traits (Provaznikova et al., 2023; Rice, 1984; Sigeman et al., 2019). These traits have different fitness optima in males and females and thus are likely to evolve under sexually antagonistic selection, which is also a cornerstone of the classical model of sex chromosome evolution. The model postulates that differentiation of sex chromosomes results from selection for linkage between the sex determining locus and sexually antagonistic mutations, which can mitigate the sexual conflict (Rice, 1987; Wright et al., 2016; Furman et al., 2020; but cf. Perrin, 2021). An intriguing corollary to this theory is that certain autosomes might also be preferentially involved in forming neo-sex chromosomes if they harbour an excess of genes experiencing sexually antagonistic selection (Charlesworth & Charlesworth, 1980; Kitano et al., 2009; O'Meally et al., 2012; Ross et al., 2009; Sigeman et al., 2019). Robustly assessing which loci are sexually antagonistic is challenging, as doing so would ultimately require assessing sex-specific fitness in a wide range of contexts (Cheng & Kirkpatrick, 2016; Dagilis et al., 2022; Innocenti & Morrow, 2010; Kirkpatrick & Guerrero, 2014; Mank & Ellegren, 2009; Veltsos et al., 2020). However, since sexual dimorphism results from sex-biased genes, i.e., genes differentially

expressed between sexes, these are regularly used as a tractable, if admittedly imperfect, indicator of sexual antagonism (Grath & Parsch, 2016; Mank, 2017). Thus, this corollary prediction could be investigated by examining the relative degree of sex-biased expression for autosomes known to be involved with fused neo-sex chromosome (cf. Lichilin et al., 2021).

Studies of sex chromosomes, and especially neo-sex chromosomes, have historically been concentrated in male-heterogametic taxa (φ XX/ δ XY system and its variants) (Ellegren, 2011). Furthermore, until recently, neo-sex chromosomes formed by autosomal fusion were thought to be scarce in taxa with female heterogamety (δ ZZ/ φ ZW system and its variants) (Pennell et al., 2015; Pokorna et al., 2014). Also, a null model of sex chromosome-autosome fusions assuming all chromosomes fuse with equal probability predicts that sex chromosome-autosome fusions should be rare in clades with high chromosome numbers (Anderson et al., 2020). For these reasons, sex chromosome fusion events may be expected to be rare among Lepidoptera (moths and butterflies), a female-heterogametic taxon with chromosome count $n=31$ being both most frequent and ancestral (Ahola et al., 2014; Wright et al., 2023). Yet, there have been numerous reports of lepidopteran neo-sex chromosomes (Fraïsse et al., 2017; Mongue et al., 2017; Nguyen et al., 2013; Šichová et al., 2015; Yoshido et al., 2020) and in some taxa, repeated sex chromosome-autosome fusions have been observed (Carabajal Paladino et al., 2019; Mongue et al., 2017; Smith et al., 2016; Yoshido et al., 2020). Notably, a recent comparative analysis of karyotypes in >200 lepidopteran species identified distinctly more fusions for the Z chromosome than any autosome (Wright et al., 2023). This suggests that fused neo-sex chromosomes may well be favoured by selection (Anderson et al., 2020) and that Lepidoptera presents a rich taxon for further investigations of sex chromosome evolution.

Among Lepidoptera, the butterfly genus *Danaus* stands out for containing two prominent examples of fused neo-sex chromosomes, one quite young neo-W and another much older neo-Z. The neo-W occurs in a population of the African queen, *D. chrysippus*, which harbours a recently arisen W-autosome fusion that is linked with sexually antagonistic phenotypes: wing colour pattern and susceptibility to a maternally inherited, male-killing bacterium, *Spiroplasma ixodeti* (Martin et al., 2020; Smith et al., 2016). Additionally, *D. chrysippus* and all other *Danaus* species share a fusion between the Z and another autosome that is sufficiently old to have lost any readily detectable sequence similarity between the neo-Z and the corresponding neo-W region, if the latter was not lost completely in this species (Gu et al., 2019; Mongue et al., 2017). Indeed, it remains an outstanding question how old this *Danaus* neo-Z is, and whether it occurs outside the genus, in other members of the tribe Danaini.

Identifying the age of the *Danaus* neo-Z would be valuable for contextualizing some notable differences between the new and ancestral segments of the neo-Z chromosome, particularly in regard to their masculinization and dosage compensation. Examined in the monarch butterfly, *D. plexippus*, both the ancestral and new neo-Z

segments exhibit gene expression that is balanced between sexes. However, expression from the ancestral segment is reduced relative to autosomes (i.e. incompletely compensated), while expression from the new segment is comparable to autosomes (i.e. completely compensated) (Gu et al., 2019; Ranz et al., 2021). Also, the ancestral segment exhibits an excess of male-biased genes in agreement with the sexual antagonism hypothesis predicting masculinization of Z chromosome, while the new segment does not (Mongue & Walters, 2018), perhaps because it is not yet sufficiently old to have been strongly impacted by this process.

Here, we report our efforts to further elucidate the evolutionary origin of the *Danaus* neo-Z chromosome. We sequenced, de novo assembled, and annotated genomes of *Tirumala septentrionis* (subtribe Danaina), *Ideopsis similis* (Amaurina), *Idea leuconoe* (Euploeina) and *Lycorea halia* (Itunina) (Figure 1) and identified Z-linked scaffolds by comparison of read-depth between sexes. While these genomes are not assembled to chromosomes, we could nonetheless assign scaffolds to ancestral linkage groups, based on orthology with the butterfly *Melitaea cinxia*, which retains the ancestral karyotype of $n=31$ (Ahola et al., 2014; Smolander et al., 2022). Doing so not only resolved the phylogenetic origin of the *Danaus* neo-Z but also revealed two additional sex chromosome-autosome fusions to have evolved among Danaini butterflies. This presented the opportunity to examine the role of sexually antagonistic selection in fixation of neo-sex chromosomes by analysing genomic distributions of sex-biased genes, which we did in *I. leuconoe* and *L. halia*, indeed finding that autosomes involved in sex chromosome fusions were enriched for sex-biased genes. Since dosage effects and idiosyncrasies of meiotic gene regulation complicate empirical testing of sex linkage and

sexual antagonism using sex-bias gene expression, we further analysed patterns of sex chromosome dosage compensation and balance in these two species. Dosage compensation patterns were very similar to monarch butterfly, even in the case of an independently evolved neo-Z chromosome.

2 | MATERIALS AND METHODS

2.1 | Sample collection, DNA extraction and genome sequencing

Pupae of *T. septentrionis*, *I. similis*, *I. leuconoe* and *L. halia* were obtained from Stratford Butterfly Farm (Stratford-upon-Avon, UK). The material was barcoded as described by Herbert et al. (2004).

A thorax of a single specimen, female in *I. leuconoe* and *T. septentrionis* and male in *I. similis* and *L. halia*, was used for high-molecular-weight DNA extraction using the Nanobind Tissue Big DNA Kit with supplemental buffers for insects (PacBio, Menlo Park, USA) following the supplier protocol optimized for insects. The Short Read Eliminator Kit (PacBio) was used according to the manufacturer's instructions to dispose of the short fragments. The resulting DNA was quantified and quality checked using the Qubit dsDNA BR Assay Kit (Invitrogen, Eugene, USA). Library preparation and Oxford Nanopore (ONT) sequencing on the PromethION platform were performed by Novogene (Hong Kong, China). Moreover, three males and females, including specimens used for ONT sequencing, were sequenced separately by the Illumina technology by Novogene to produce short reads for correction of ONT reads and coverage analyses.

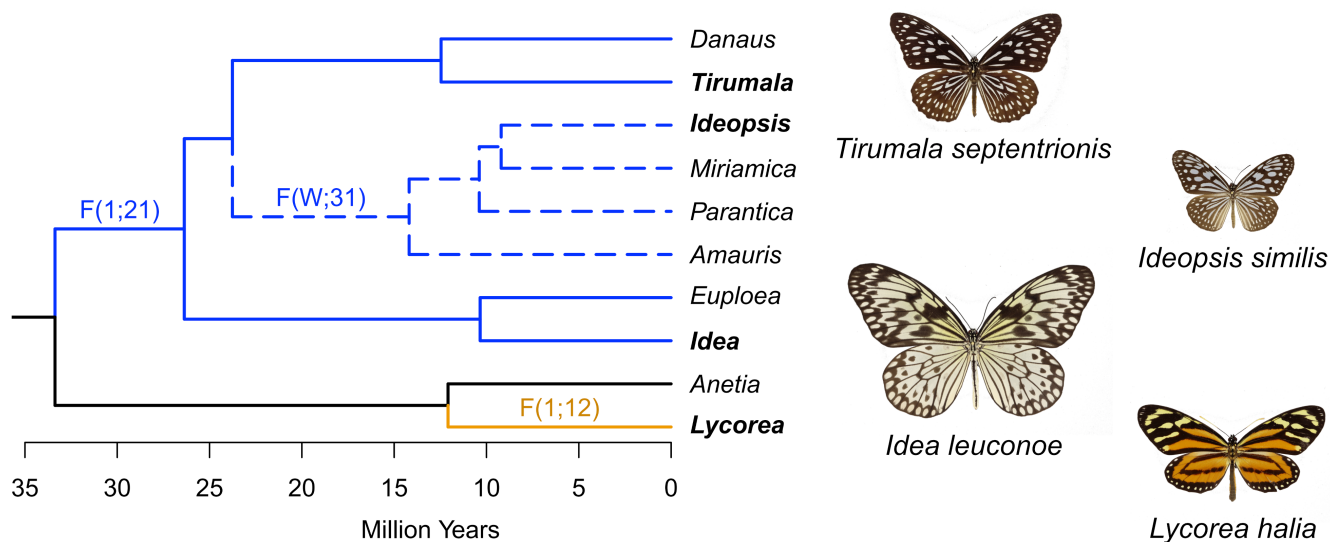


FIGURE 1 A genus-level dated phylogeny for the tribe Danaini, with inferred transitions in sex chromosome content, as explained in Results and Discussion. Fusions of the ancestral chromosome Z designated as chromosome 1 with autosomes 21 and 12 are indicated as F(1;21) (blue) or F(1;12) (orange), respectively, and placed above the branches indicating our current inference of where they arose. We propose that the observed transition of chromosome 31 into a neo-sex chromosome resulted from a fusion of its homologue with the W. The timing of this event is currently ambiguous, so a dotted line is used to indicate possible lineages carrying this change, and we placed “F(W;31)” on the earliest possible branch given our sampling. Genera containing species assayed in this study are in bold. Phylogeny and dating taken from Chazot et al. (2019).

2.2 | Read control quality and genome size estimation using GenomeScope and flow cytometry

Raw data from Illumina sequencing were inspected with FastQC v0.11 (Anderson et al., 2020) and trimmed using Trimmomatic v0.39 (Bolger et al., 2014). The ONT reads were quality checked using the NanoPack tool (De Coster et al., 2018), more specifically NanoPlot was used for the quality check and NanoFilt for filtering and trimming. Briefly, reads <10kb and of quality <10 were filtered out. The Ratatosk v0.1 tool (Holley et al., 2021) was used with default setting to correct the ONT reads using the accurate Illumina reads from the same specimen.

For genome size estimate, GenomeScope (Vurture et al., 2017) with the Jellyfish v2.1.3k-mer counting tool (Marçais & Kingsford, 2011) was used. K-mers ranging from 11 to 33 were tried with $k=21$ selected for its best model fit value.

Genome size was also determined by flow cytometry using frozen heads ($N=3$) from the sequenced specimens, females for both *I. leuconoe* and *T. septentrionis* and males for both *I. similis* and *L. halia*. A linden bug *Pyrrhocoris apterus* with $1C=1.34$ pg (this study, based on calibration against *E. kuehniella* males with $1C=440$ Mbp, Buntrock et al., 2012) was used as an internal standard. The tissues were shredded in 500 μ L of the nuclear isolation buffer (Buntrock et al., 2012; 0.1 M Tris-HCl pH=7.5, 2 mM MgCl₂, 1% Triton X-100) with a razor blade to release nuclei. The resulting suspension was filtered and another 500 μ L of isolation buffer was added. The nuclei were then stained for 20 min with 50 μ g/mL propidium iodide (Sigma-Aldrich, St. Louis, USA). The samples were measured with the Sysmex CyFlow Space flow cytometer (Sysmex Partec, Munster, Germany) equipped with a 100 mW 532 nm (green) solid-state laser. The side scattered light (SSC) and the fluorescence of about 8000 nuclei or more were analysed using FlowJo 10 software (TreeStar, Inc., Ashland, OR, USA). Mean, coefficient of variation and number of nuclei were documented for 2C peaks of both the standard and the sample. Finally, the ratio standard/sample of mean fluorescence was calculated.

2.3 | Genome de novo assembly and its quality evaluation

The corrected ONT reads were assembled by Flye v2.8 (Kolmogorov et al., 2019) using the “--nano-raw” option and an appropriate genome size estimated by GenomeScope. The assemblies were subjected to one round of long-read polishing using medaka (<https://github.com/nanoporetech/medaka>) and then one round of short-read polishing using NextPolish (Hu et al., 2020). The purge_dups v1.0.1 tool was used to remove haplotypic duplicates (Guan et al., 2020). The assemblies were evaluated by BUSCO v5 using the Lepidoptera database (odb_10) (Manni et al., 2021). BlobTools v1.0 (Laetsch & Blaxter, 2017) was used to check for contamination. Contigs belonging to non-target organisms were removed using the “seqfilter” function.

2.4 | Genome annotation

Both functional and structural annotations were done through GenSAS v6.0 pipeline (Humann et al., 2019). Repetitive sequences were identified by RepeatModeler2 (Flynn et al., 2020) with the RMBlast search engine and the TFR v4.09, RECON, and RepeatScout v1.0.5 modules. Moreover, we used TAREAN (Novák et al., 2017) for the satDNA annotation. All consensus sequences annotated as satellites by TAREAN (both high and low confidence) were included in a custom database as dimers in order to get a better annotation of satellite DNA. RepeatMasker v4.1.1 (Smit et al., 2013–2015, available at <http://www.repeatmasker.org>) with the NCBI/RMBlast search engine was used for the annotation of repeats using a combination of the new repeats retrieved by RepeatModeler2 and the custom database with the satDNA sequences from TAREAN. For *L. halia* and *I. leuconoe*, total RNA from gonads (testes or ovaries), head and thorax dissected from 1-day-old imagoes of both sexes were extracted by TRI Reagent (Sigma-Aldrich) and used to produce Illumina RNA-Seq libraries (Novogene) with 450bp inserts. The 150-bp Illumina reads were mapped to the genomes using STAR v2.7.7 (Dobin & Gingeras, 2015). First, the genome index was generated with options “--runMode genomeGenerate --genomeSAindexNbases 13”. Then, the mapping was carried out with the basic options as listed in the manual. The resulting SAM file was transformed into BAM format using SAMtools suite (v1.11) (Li et al., 2009). The BAM file was used for gene prediction using BRAKER2 with default options which include Augustus and GeneMark-EP (Brůna et al., 2021; Lomsadze et al., 2014). The *T. septentrionis* and *I. similis* assembled genomes were used as input for Augustus v3.3.1 (Stanke & Morgenstern, 2005) and GeneMark-ES directly with no RNA-Seq evidence to guide the annotation. Both tRNA and rRNA sequences were identified in all the assemblies using tRNAscan-SE v2.0.7 (Chan & Lowe, 2019) and RNAmmer v1.2 (Lagesen et al., 2007) respectively.

2.5 | Coverage analyses to identify sex linkage

Well-differentiated sex chromosomes should produce diagnosable differences in sequencing coverage that can be used to assess sex linkage in genome assemblies (Palmer et al., 2019). To this end, for each species, we aligned Illumina short reads from three male and female samples to the reference assembly using Bowtie2 (v2.3.5.1) (Langmead & Salzberg, 2012) with the “--very-sensitive-local”, “--no-mixed” and “--no-discordant” options, compressing the output to BAM format using SAMtools suite (v1.11) (Li et al., 2009). The resulting BAM files were then parsed using the “genomecov” and “groupby” utilities from the Bedtools suite (v2.25.0) (Quinlan & Hall, 2010) to obtain the median coverage depth for each scaffold; reads mapping to regions corresponding to repetitive sequences annotated by RepeatModeler2 were excluded using the Bedtools “subtract” utility. In each sample, coverage depths for each scaffold were normalized by median coverage across scaffolds, averaged within sex and compared between sexes, formulated as the Log₂ of

the male:female (M:F) coverage ratio. An alternative, window-based, coverage analysis conducted using IndexCov (Pedersen et al., 2017) was also employed (without masking of repeats) to support visualizations in comparative analyses, with coverage in windows averaged and compared between sexes as noted above. Autosomal scaffolds should present a $\text{Log}_2(\text{M:F})=0$ as they are expected to be present in equal proportion in both sexes, while the Z-linked scaffolds should ideally present a $\text{Log}_2(\text{M:F})=1$ due to its double representation in males. We considered scaffolds or regions of assembly with $\text{Log}_2(\text{M:F}) > 0.5$ to be Z-linked.

2.6 | Differential expression analysis

RNA-Seq libraries of *L. halia* and *I. leuconoe* were used also for the differential expression analysis, with three (single individual) replicates from each sex for head, thorax and gonads. The Illumina reads were quality checked with FastQC (Andrews, 2010) and filtered with “--nextseq-trim = 20” option using cutadapt v1.15 (Martin, 2011) and trimmed with Trimmomatic v0.36 (Bolger et al., 2014) with following parameters: “SLIDINGWINDOW:5:20”, “MINLEN:50”, “HEADCROP:14” and “CROP:134”. The rRNA was filtered out using SortMeRNA v4.3.6 (Kopylova et al., 2012). The trimmed and filtered reads were mapped to the reference genome of the given species using STAR v2.5.2b (Dobin & Gingeras, 2015). The annotation gff file was used with the option “--sjbGTFtagExonParentTranscript Parent”. The maximum intron length was specified at 130,000bp (“--alignIntronMax 130,000”). The alignments were sorted by coordinate and output directly in binary BAM format (“--outSAMtype BAM SortedByCoordinate”).

Read counts were generated with R/Rsubread featureCounts package v2.8.2 (Liao et al., 2014) using an annotation created with the easyRNASeq package v2.30.0 (Delhomme et al., 2012) that contained synthetic transcripts, i.e., a transcript combining every single exon of a gene into a single abiological splice variant to avoid counting unique mRNA fragments multiple times. Differential expression analysis between males and females (in different tissue types separately) was performed with R/Bioconductor DESeq2 package v1.34.0 (Love et al., 2014). For visual exploration of sample relationship, we transformed counts with the implemented variance-stabilizing transformation (VST). The differentially expressed genes were filtered following Schurch et al. (2016) recommendations by lowering the false discovery rate threshold to 0.01 (“alpha = 0.01”) and by raising the Log_2 fold change threshold to 0.5 (“lfcThreshold = 0.5”).

2.7 | Dosage compensation analysis

Read counts were transformed to FPKM-normalized expression values using library sizes calculated via the TMM algorithm (Oshlack et al., 2010; Robinson & Oshlack, 2010). For each tissue, expression values were then averaged within sex to produce a mean FPKM

value for each gene in each sex. Genes with a sex-specific average FPKM value < 0.01 were considered unexpressed and were removed from analysis in that sex. Distributions of FPKM values were compared among chromosomes within sex, contrasting three partitions: genes found on the ancestral Z, neo-Z or autosomes. A similar comparison was made among chromosomal groupings using the Log_2 transformed male:female ratios of expression, considering only genes expressed in both sexes. Significant variation among the chromosomal partitions was statistically evaluated with a Kruskal-Wallis non-parametric test, followed by post-hoc pairwise contrasts via Dunn's test as warranted, implemented in the R package PMCMRplus (Pohlert, 2022).

2.8 | Chromosomal distribution of sex-biased genes

To support comparative analyses across species, scaffolds in each species were assigned to homeologous chromosomes in the Glanville fritillary, *Melitaea cinxia* (Nymphalidae), which represent the ancestral lepidopteran genome organization with $n=31$ chromosomes (Smolander et al., 2022) (We adopt the term “homeologous” here to indicate evolutionarily homologous chromosomes derived from separate species and to avoid ambiguity when referring to genetically homologous chromosomes, i.e., those that pair during meiosis). Chromosomal assignments were based on maximum count per scaffold of orthologous genes identified by BUSCO (v5) using the Lepidoptera database (odb_10) (Manni et al., 2021). For the scaffold that bridged the transition between neo- and ancestral Z, we used PROmer (Kurtz et al., 2004) to align the scaffold with the two corresponding *M. cinxia* chromosomes, manually inspected the results to estimate the breakpoint between neo-Z and anc-Z chromosomal segments and assigned genes accordingly.

With most genes assigned to ancestral linkage group, we could then investigate the chromosomal distribution of sex-biased genes in *Idea leuconoe* and *Lycorea halia*. We used a chi-squared test with Bonferroni correction to assess each chromosome for significant differences in the proportion of male-biased or female-biased genes relative to the combined count from the remainder of the genome. Because this assessment of sex-biased expression is intended to serve as a proxy to indicate which genes experience sexually antagonistic evolution, we performed this analysis twice, once considering all genes with significant differential expression (DE) between sexes, and again excluding genes which were expressed in only one sex, based on the criteria of FPKM > 0.01 applied during dosage compensation analysis (see above). By default, significant DE genes may include genes with expression only in one sex. Such sex-limited expression may reflect the resolution of prior sexually antagonistic evolution, and so it is reasonable to include sex-limited genes in an assay of sexual antagonism. But we cannot exclude the possibility that such genes were always sex limited and thus did not experience further sexually antagonistic evolution. So, a more conservative version of the analysis can be

accomplished by considering only DE genes with evidence for expression in both sexes. We primarily present the analysis excluding sex-specific genes and report the analysis with all DE genes in the supplement.

2.9 | Genomic in situ hybridization

For *I. leuconoe*, *T. septentrionis* and *L. halia*, spread chromosome preparations were made from ovaries of female pupae following Mediouni et al. (2004) with slight modifications detailed in Šichová et al. (2013). The preparations were dehydrated in an ethanol series (70, 80 and 100%, 30s each) and stored at -80°C till further use.

Genomic DNA (gDNA) was isolated separately from thoraces of both sexes by MagAttract HMW DNA Kit (Qiagen, Hilden, Germany). Due to the obtained small gDNA yield in all males and *L. halia* female, we amplified their gDNA by illustra GenomiPhi HY DNA Amplification Kit (GE Healthcare, Milwaukee, WI). Hybridization probes were prepared by Nick Translation from female gDNA and labelled by fluorochrome Cy3-dUTP (Jenna Bioscience, Jena, Germany). The 20- μL reaction contained 1 μg of DNA, 1 $\times$ nick translation buffer (5mM Tris-HCl, 0.5mM MgCl_2 , 0.0005% BSA; pH 7.5), 10mM β -mercaptoethanol, 50 μM dATP, dCTP and dGTP, 10 μM dTTP, 20 μM Cy3-dUTP, 20U DNA polymerase I (Thermo Fisher Scientific, Waltham, MA, USA) and 0.005U DNase I (RNase-free, Thermo Fisher Scientific). The reaction was incubated at 15°C for 3h and finally stopped by adding 1 $\times$ loading buffer (50% glycerol, 250mM EDTA, 5.9mM bromophenol blue).

To detect W sex chromosomes, genomic in situ hybridization (GISH) was carried out following the procedure of Sahara et al. (1999) with slight modifications. Chromosomal preparations were removed from the freezer, passed through the ethanol series and air-dried. The chromosomal preparations were pre-treated with 100 $\mu\text{g}/\text{mL}$ RNase A for 1h at 37°C and washed twice for 5min in 2 $\times$ SSC. Afterwards, chromosomes were denatured for 3min 30s in 70% formamide in 2 $\times$ SSC and then immediately dehydrated in cold ethanol series. For one slide, 10 μL of hybridization mix contained 300 ng labelled female genomic DNA, 3 μg unlabelled competitor DNA, 25 μg sonicated salmon sperm DNA (Sigma-Aldrich, Munich, Germany), 50% formamide and 10% dextran sulphate in 2 $\times$ SSC. The fragmented competitor DNA was prepared from male genomic DNA by heating (99°C for 20min). Hybridization was carried out for 3days at 37°C in a moist chamber. Then, the slides were washed at 62°C for 5min in 1% Triton X-100 in 0.1 $\times$ SSC and 2min in Ilford (Ilford) at room temperature. Finally, the slides were counterstained with 0.5mg/mL DAPI (4',6-diamidino-2-phenylindole; Sigma-Aldrich, Munich, Germany) in antifade based on DABCO (1,4-diazabicyclo(2.2.2)-octane; Sigma-Aldrich, Munich, Germany).

Results were documented using a Leica DM6 B Fluorescence Microscope (Pragolab, Prag, Czech Republic) equipped with the appropriate filter set. Images were captured with a Leica sCMOS Monochrome Camera DFC9000GT equipped with the Leica

Application Suite X (LAS X) imaging software v3.7.3.23245. The images were pseudocoloured and superimposed with Adobe Photoshop CC 2019.

3 | RESULTS

3.1 | Genome size estimation

Genome sizes were estimated bioinformatically using GenomeScope as well as cytologically using flow cytometry. Size estimates from GenomeScope ranged between 301 and 338 Mbp, with estimated heterozygosity ranging between 1 and 1.75%, and repeat content between 24 and 32% (Table 1, Figure S1). Size estimates from flow cytometry (Table 1, Figure S2) were about 30–50% greater than those from GenomeScope in all species. For both methods, *T. septentrionis* was estimated to have the largest genome size, consistent with having the largest portion of repetitive content as estimated by GenomeScope.

3.2 | Genome assemblies and annotation

Details of raw read counts and lengths for genome sequencing samples are given in Text S1 and Table S1. A summary of genome assembly results for each species is presented in Table 1. Assemblies from three species had N50 values >2 Mbp, but the *T. septentrionis* assembly was notably less contiguous, perhaps due to greater repeat content and heterozygosity in this species. Nonetheless, BUSCO analysis showed that $>95\%$ of conserved orthologues were found to be complete and unique in all four species (Table 1). Thus, all the assemblies provide high-quality references to support comprehensive gene annotations (Table 1). Total length of genome assemblies corresponded very closely to the genome sizes estimated by GenomeScope, as did the proportion of repetitive sequences. The genome assembly and annotation files were deposited in the Dryad repository (Mora et al., 2023).

3.3 | Identifying Z-linked regions of genomes

Putative Z-linked scaffolds were identified in the assemblies by comparing sequencing depth between males and females. In taxa with female heterogamety ($\delta\text{ZZ}/\eta\text{ZW}$) such as Lepidoptera, males should have double the coverage of Z-linked sequences compared to females, while coverage of autosomal sequences should be comparable. All four assemblies showed one or a few large scaffolds with male:female coverage ratios falling almost exactly in line with the expected two-fold difference (Figure 2), strongly indicating that these correspond to regions of the Z chromosome that are well differentiated from any W chromosome. Notably, in all four assemblies, the largest scaffold is apparently Z-linked. Ratio and length data plotted in Figure 2 are provided in Table S2.

TABLE 1 Summary of genome assemblies.

	<i>Idea leuconoe</i>	<i>Ideopsis similis</i>	<i>Lycorea halia</i>	<i>Tirumala septentrionis</i>
Genome size estimation				
Genome Scope	312 Mbp	332 Mbp	301 Mbp	338 Mbp
Heterozygosity (%)	1.58	0.99	1.26	1.74
Repeats (%)	26.8	24.9	24.0	32.6
Flow Cytometry	400 Mbp	475 Mbp	424 Mbp	508 Mbp
Assembly summary				
Total Length (Mbp)	314	325.5	291.7	381.4
N Fragments	474	693	1085	2616
N50 (Kbp)	3420.6	2913.3	2095.3	679.8
Largest Contig (Mbp)	13.56	19.63	15.76	8.63
GC content (%)	31.64	31.99	31.74	32.13
Total Repeats (%)	27.06	29.04	19.12	33.20
Tandem Repeats (%)	2.11	2.47	2.45	2.50
Interspersed Repeats (%)	25.02	26.59	16.55	30.57
BUSCO (C; D; F; M) (%)	98.4; 0.3; 0.4; 0.9	97; 0.3; 0.5; 2.2	97.7; 0.7; 0.3; 1.3	95.3; 3.1; 0.5; 1.1
N genes	20,094	18,754	17,652	21,918

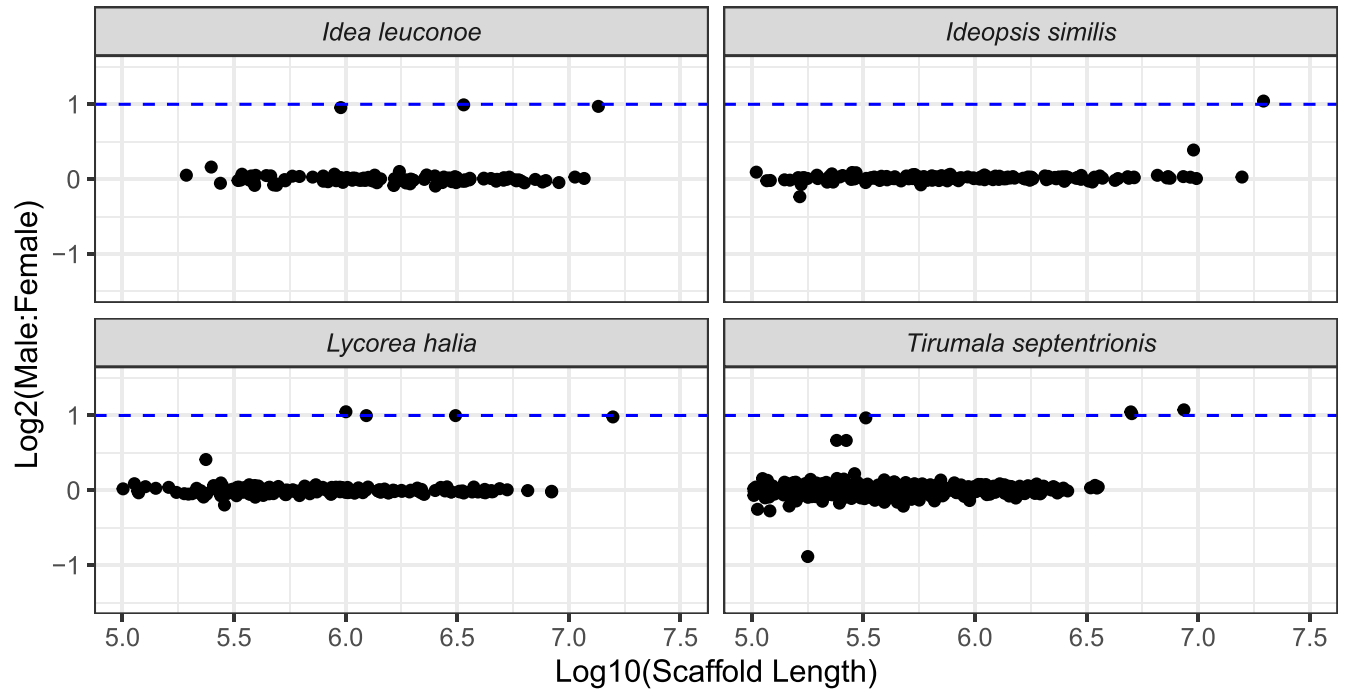


FIGURE 2 Patterns of male versus female average median sequencing coverage for genome scaffolds. Repetitive regions were excluded from coverage calculations. Only scaffolds larger than the N90 scaffold length are shown. The dotted blue line at 1 indicates the expected value for Z-linked scaffolds.

The assemblies of *I. similis*, *L. halia* and *T. septentrionis* also yielded one or two scaffolds with intermediate values of sex-specific coverage. Such intermediate values may result from variation in repetitive DNA content among re-sequenced individuals compared to the reference assembly. It may also reflect variable levels of differentiation between neo-sex chromosomes, or indicate a fusion between a well-differentiated Z chromosome and an autosome. Such a fusion may

be real (i.e. it exists in the organism's genome), but often it reflects an error in genome assembly producing a chimeric scaffold (cf. Mongue et al., 2017). Inspecting sex-specific coverage in windows along the scaffold can often be informative in such cases; hence, we assessed these few “intermediate” scaffolds using IndexCov (Pedersen et al., 2017), which provides normalized coverage patterns for each sample in ~16 Kbp windows.

In *L. halia* and *T. septentrionis*, the scaffolds (one and two respectively) are each ~250Kbp and do not show consistent patterns of coverage between sexes or across the scaffolds (Figure S3). We, therefore, assume that these intermediate “average” coverage patterns reflect variation in genome structure unrelated to sex (e.g. differences in repeat content) and that these scaffolds are not Z-linked. In *I. similis*, the one scaffold showing intermediate coverage is much larger, at ~9.5Mbp, and does show a clear difference in sex-specific coverage indicating that a portion of this scaffold is Z-linked (Figure S3). We assess this pattern results from a mis-assembly, as direct inspection of ONT reads mapped back to the assembly shows a region with a distinct shift in coverage and only one presumably chimeric read spanning position 5,604,305 in scaffold 46:1.0–9461643.0_np1212 (Figure S4). Accordingly, we manually split this scaffold at this position, and all subsequent analyses use this modified assembly.

T. septentrionis also exhibited one fragment (contig_36:1.0–178234.0_np1212) with strongly female-biased coverage, which likely corresponds to W-linked sequence, since the sequenced individual was female (Figure 2).

3.4 | Ancestral linkage groups and neo-Z chromosomes

We used counts of orthologous genes to assign scaffolds in each assembly to a homeologous chromosome from *M. cinxia*. Because *M. cinxia* retains the ancestral lepidopteran karyotype of $n=31$ (Ahola et al., 2014; Wright et al., 2023), this provided a common framework for comparative analysis of sex linkage among the analysed species. Broadly speaking, gene content per chromosome appears well conserved between *M. cinxia* and all four Danaini species analysed, with very few genes moving between homeologous chromosomes (Figure 3; Figure S5). Nonetheless, it is known that karyotypes and chromosome numbers have changed substantially in at least some Danaini species relative to the ancestral lepidopteran karyotype (Brown et al., 2004). Since we do not have direct assessments of synteny of autosomal genes in these species (see below), nor fully chromosomal genome assemblies, the relationships between autosomes depicted in Figure S5 are meant only to facilitate comparisons among species.

Yet, despite having imperfect knowledge of karyotypes among these four Danaini species, it is nonetheless clear that at least three major transitions in Z-linkage have occurred in this group relative to the ancestral lepidopteran karyotype (Figure 3; Figure S5). The fusion of homeologues of McChr1 (corresponding to the ancestral Z) and McChr21 was reported previously in monarch butterfly (Mongue et al., 2017). This Z-autosome fusion is shared among *T. septentrionis*, *I. similis* and *I. leuconoe*, but is absent in *L. halia*. However, *L. halia* carries a different Z-autosome fusion involving the homeologue of McChr12. Finally, in addition to the fusion between McChr1 and 21, *I. similis* exhibits another autosome to Z transition involving the homeologue of McChr31. The scaffold in *I. similis* containing this McChr31 is the single large ~9.5 Mbp scaffold with “intermediate” coverage discussed above. The region containing Z-like coverage patterns is entirely consistent

with patterns of homology to McChr31. Thus, given the assembly error that resulted in connecting the McChr31 homeologue to the section of the same scaffold with autosomal coverage (corresponding to McChr5; Figure S6), we have no evidence that the *I. similis* McChr31 homeologue is fused to another chromosome. Notably, in all three cases of neo-Z chromosomes formed from autosomes, the coverage patterns suggest very substantial divergence between the Z and any homologous W chromosome that may exist (Figures 2 and 3).

Placing these inferences in a phylogenetic context, we can start to clarify the origins of the sex chromosome transitions we identified (Figure 1). It appears that the fusion of the ancestral Z and autosome homeologous to McChr21 is shared by all assayed species (and *Danaus*) other than *L. halia*. Based on the dated phylogeny from Chazot et al. (2019), this likely took place about 30 Mya. We have less ability to resolve when the transition of the McChr31 homeologue to a sex chromosome occurred because we do not know the status of this chromosome in genera *Miriamica*, *Parantica* or *Amauris*. Thus, this transformation could be nearly as old as the McChr21 fusion, or potentially rather recent, if it occurs only in *Ideopsis* species. In any case, the consistent two-fold difference in coverage between males and females for this region of the genome suggests that the neo-Z is strongly diverged from its neo-W counterpart, or is hemizygous in females. It is also possible that the fusion and transformation of McChr12 observed in *L. halia* is as old as that for McChr21; we currently lack information concerning the status of McChr12 in its sister genus *Anetia*, and the next oldest branching is at the base of the Danaini.

In addition to its transformation into a Z chromosome in *I. similis*, the homeologue of McChr31 showed notable patterns in the assemblies for *I. leuconoe* and *L. halia*. However, in both cases we judge these to be analytical artefacts, and not biologically significant. In *I. leuconoe*, there is a small region with Z-like M:F coverage ratios, suggestive of sex-specific differentiation as might be associated with a sex-determining region. However, this pattern reflects comparison of averages within sex as individual samples are actually highly variable in this region and do not show a consistent two-fold coverage difference between sexes (Figure S7). Thus, we deem this apparent two-fold difference to be an artefact of averaging across samples in a region with highly variable individual coverage, likely arising from a copy-number variant segregating among samples at this locus.

In *L. halia*, our genome assembly initially yielded a single scaffold (1585_1.0–3721850.0_np1212) that merged chromosomes homeologous to McChr28 and McChr31. Together with the fusion of McChr1 and McChr12 forming the neo-Z, this would suggest a chromosome count of $n=29$, which is one fewer than the published karyotype of $n=30$ for this species based on microscopic observation of meiosis (Brown et al., 2004). We, therefore, suspected this apparent fusion may be a mis-assembly. Using PROmer alignments with *M. cinxia*, we estimated the transition point between regions corresponding to McChr28 and McChr31 occurred at about position 2.2Mbp. We then inspected ONT reads mapped back to the assembly at this point and observed it contained an ~10kbp region with exceptionally low coverage, poorly mapped reads and numerous

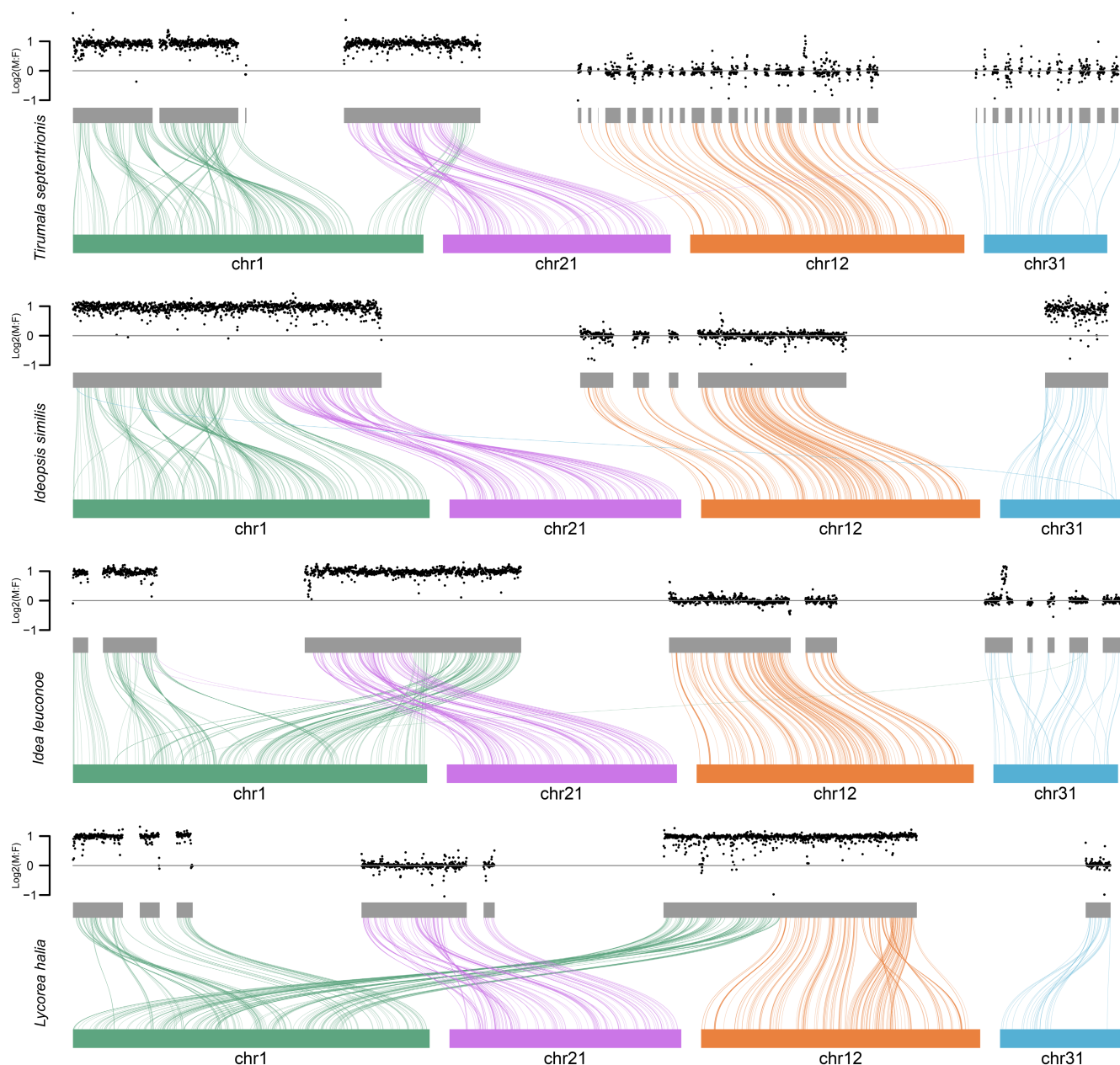


FIGURE 3 Patterns of shared orthology and sex-specific coverage for sex-linked regions among four Danaini butterflies under study. Coloured lines connect the respective positions of single-copy BUSCO orthologues in chromosomes from the common reference species *Melitaea cinxia* (coloured blocks) to scaffolds in the focal genome assembly (grey blocks). The plot is limited to only scaffolds showing sex linkage in Danaini and the corresponding reference chromosomes. Scaffolds in each focal species are arranged and oriented based on the maximum count and positions of orthologues in reference chromosome. The $\text{Log}_2(\text{Male}:\text{Female})$ coverage is plotted in ~16 Kbp windows above each focal cluster, in positions corresponding to the focal scaffold. The horizontal grey line in coverage plots indicates the expected value for autosomes. For *I. leuconoe*, a small region corresponding to chr31 appeared to have Z-like coverage, but was specifically investigated and deemed to be autosomal (see main text). A plot showing all reference chromosomes and homeologous regions of study species is given in [Figure S5](#).

indels. Dotplot analysis also revealed a large repetitive structure at this position ([Figure S8](#)). Thus, we judged this apparent fusion to be a mis-assembly, split the scaffold at 2.2 Mbp and subsequently updated analyses of ancestral chromosomal linkage using this manually modified assembly. A similar examination of the fusion point between McChr1 and 12, forming the *L. halia* neo-Z, showed no such anomalies in coverage ([Figure S9](#)).

3.5 | Differentially expressed genes between males and females

We used RNA-seq to assay patterns of sex-biased gene expression in two of our four study species for which relevant samples were available: *I. leuconoe* and *L. halia*. The amount of sex-biased gene expression differed substantially between tissues, with somatic

tissues yielding almost no genes with significant differential expression between sexes (not shown). We, therefore, present here only an analysis of gonads, using the criteria of an FDR adjusted p -value ($p_{\text{adj}} < .01$) and at least 50% difference in expression to demarcate differential expression. We identified 4690 differentially expressed (DE) genes in *I. leuconoe*, with 1703 female-biased (down-regulated in males) and 2987 male-biased (up-regulated in males) genes (Figure S10). Using the same criteria in *L. halia*, we identified 5203 DE genes, with 1860 female-biased and 3343 male-biased genes (Figure S10).

We then counted sex-biased genes per *M. cinxia* homeologue and tested each chromosome for significant deviations in the proportion of male- and female-biased genes (Chi-square, Bonferroni adjusted $p < .05$). Genomic regions corresponding to McChr1 and 31 were significantly enriched and McChr19 and 24 depleted in male-biased genes in *I. leuconoe*. The only chromosome significantly enriched in female-biased genes was McChr11 (Figure 4). In *L. halia*, McChr1 and 31 were enriched and McChr24 depleted in male-biased genes. Female-biased genes were significantly enriched on McChr21 and

22 but depleted on McChr1 (Figure 4). This version of the analysis excludes differentially expressed genes with sex-specific expression (see methods). Including all differentially expressed genes produced comparable results, with the addition that McChr25 and 8 were also significantly enriched in male-biased genes in *I. leuconoe* and *L. halia*, respectively (Figure S10). Lists of the DE genes were deposited in the Dryad repository (Mora et al., 2023).

3.6 | Dosage compensation

In somatic tissues (head and thorax), patterns of average gene expression for the ancestral and new segments of the neo-Z, and autosomes were grossly similar for both sexes. No significant differences were detected between the new Z segment (Znew) and autosomes, indicating that the Znew segment appears to be fully compensated relative to autosomes in both species (Figure 5; Tables S3–S5). However, there was a consistent qualitative difference in Znew: A median ratios between species, with Znew being slightly lower than

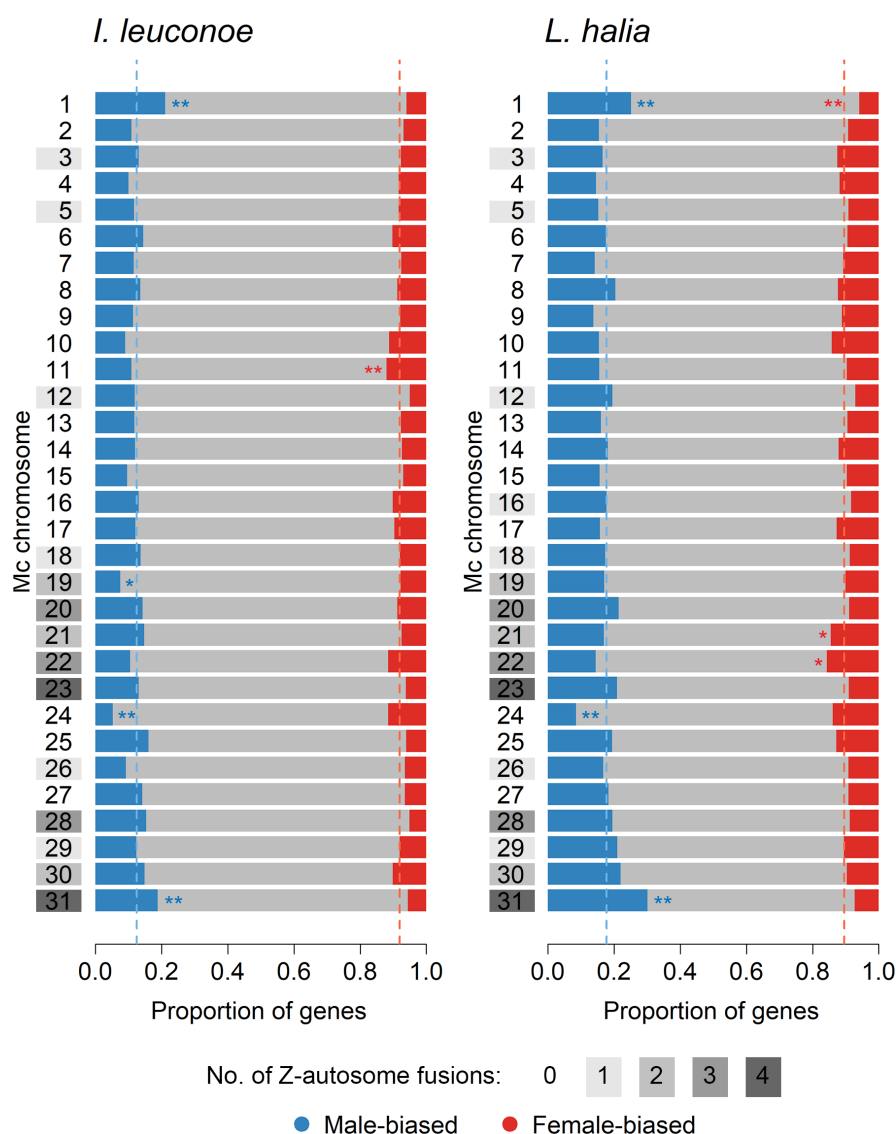


FIGURE 4 Distribution of male- and female-biased genes expressed in gonads in *Idea leuconoe* and *Lycorea halia*. Genes were assigned to ancestral linkage groups represented by chromosomes of *Melittaea cinxia*. In both *I. leuconoe* and *L. halia*, the ancestral Z genes correspond to McChr1, while genes of new Z segments were assigned to McChr21 and McChr12 respectively. Note that genes from McChr31 are sex linked in *I. similis*. No. of Z-autosome fusions shows tendency of synteny blocks to be involved in these fusions (according to Wright et al., 2023). This analysis excludes sex-specific genes; results including all differentially expressed genes are given in Figure S10. Significant difference from mean (dashed line), Chi-square, Bonferroni adjusted $p < .05$ (*), $p < .01$ (**).

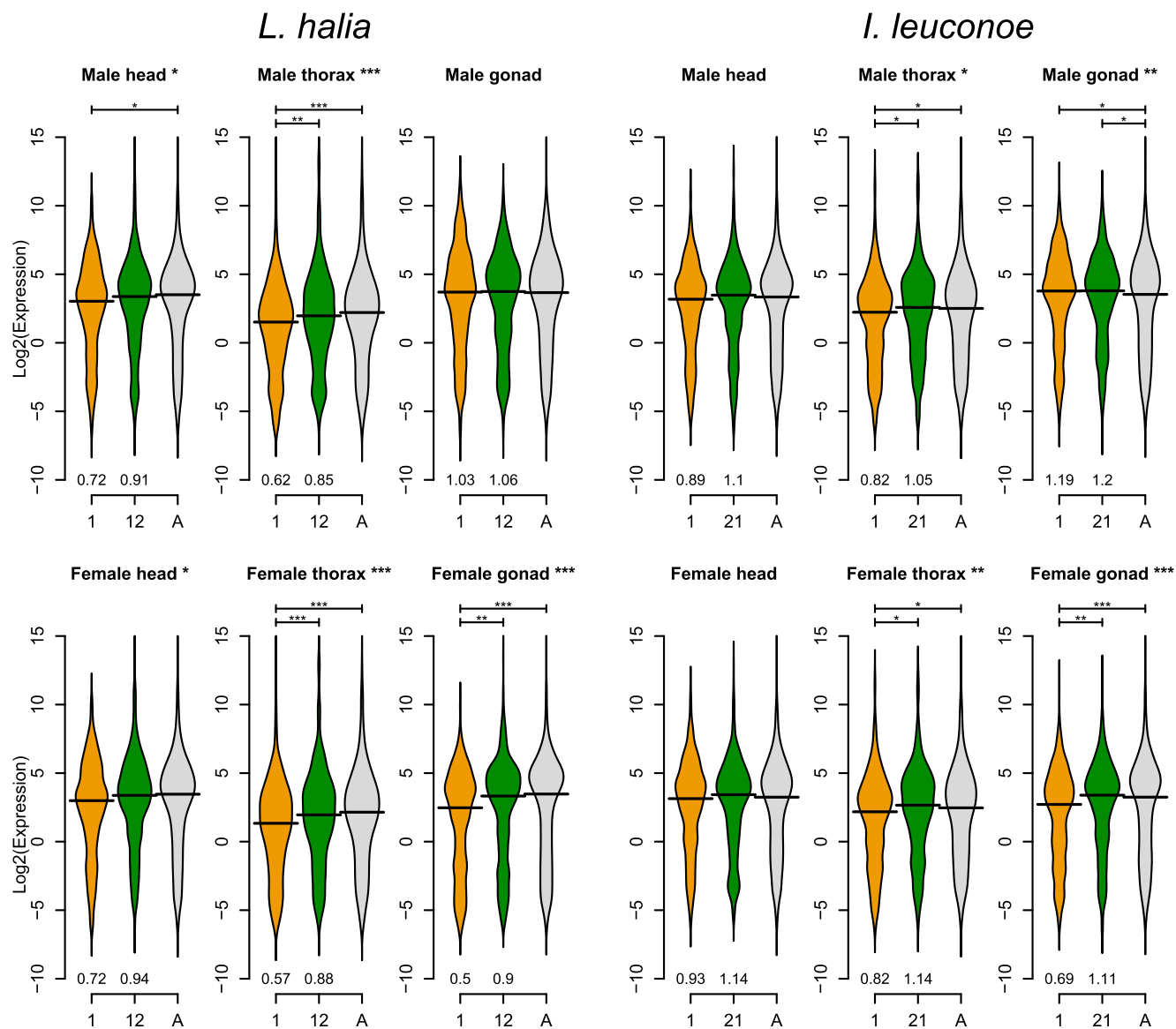


FIGURE 5 Patterns of dosage compensation in two Danaini butterflies. Distributions of Log_2 transformed FPKM expression values are given for ancestral Z (McChr 1; orange), new Z segment (McChr 12 or 21; green) and autosomes (A; grey) for each sex and three tissues in *Lycorea halia* and *Idea leuconoe*. Horizontal black lines represent median values. Z:A median ratios of untransformed FPKM are given above the X-axis. Statistical significance of Kruskal–Wallis tests applied to each data set is indicated by stars in plot titles. Significance of pairwise post-hoc Dunn's tests is indicated by bars above the distributions. * $p < .05$; ** $p < .01$; *** $p < .001$.

autosomes in *L. halia*, but slightly higher in *I. leuconoe*. In contrast, the general pattern for the ancestral Z (Zanc) is incomplete compensation. Zanc expression was significantly reduced relative to autosomes in all somatic tissues except head in *I. leuconoe*, which was also reduced but not significantly so (Figure 5; Tables S3–S5). For *L. halia* female head, the Kruskal–Wallis test was significant, with a post-hoc p -value of .056. The ancestral Z:A ratio typically fell in the range of 0.6–0.9.

We also assayed whether any dosage effects could be detected on male:female expression ratios (i.e. dosage balance c.f. Gu & Walters, 2017). We detected significant differences among the chromosomal partitions in both somatic tissues for both species

(Figure 6; Tables S6 and S7). In all cases, the Z-linked partitions exhibited a male bias. However, the magnitude of this male bias was quite small, in the range of 1–8% greater male expression compared to autosomes. This indicates that dosage balance is nearly complete for both segments of the Z chromosome in these species, with only a very modest dosage effect being detectable between sexes.

In gonads, patterns of Z chromosome expression were categorically different from somatic tissues (Figure 5; Table S3). As has been reported previously for lepidopteran gonads (Gu et al., 2017; Huylmans et al., 2017; Walters & Hardcastle, 2011), the ancestral Z exhibits a strong male bias in average expression relative to other tissues. This produces a pattern where the average expression on

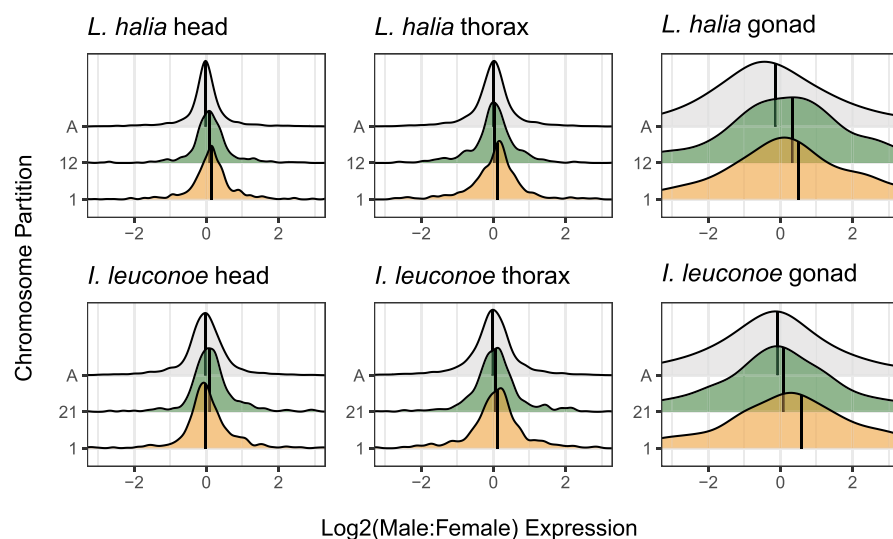


FIGURE 6 Patterns of dosage balance in two Danaini butterflies. Distributions of Log_2 transformed ratios of male:female FPKM expression values are given for ancestral Z (McChr 1; orange), new Z segment (McChr 12 or 21; green) and autosomes (A; grey) for three tissues in *Lycorea halia* and *Idea leuconoe*. Vertical black lines represent median values. All six comparisons yielded highly significant ($p < .001$) Kruskal–Wallis.

the male Z is equal to or greater than that in autosomes, while the female Z remains reduced, similar to levels in female somatic tissues. Interpreting this pattern in relation to dosage compensation remains difficult, as it could possibly reflect a range of phenomena: the lack of dosage compensating mechanisms, female meiotic silencing of the Z, sexual antagonism or some combination thereof (Gu et al., 2017; Huylmans et al., 2017).

3.7 | Identification of the W sex chromosome by GISH

We were not able to determine chromosome numbers in species under study due to shortages of material. However, different populations of all the species under study have been karyotyped before (Brown et al., 2004; Maeki & Ae, 1969). While *Idea leuconoe* and *L. halia* males were reported to have chromosome number $n=30$, *Tirumala septentrionis* and *Ideopsis similis* were reported to have derived chromosome numbers of $n=32$ and $n=13$ respectively. However, variation in chromosome numbers caused by additional 1–2 small supernumerary chromosomes was observed in both *T. septentrionis* and *I. similis* (Brown et al., 2004; Maeki & Ae, 1969).

We used GISH in three species, namely, *I. leuconoe*, *L. halia* and *T. septentrionis* to identify their W sex chromosomes. GISH experiments involve a competitive hybridization between labelled female genomic DNA and an excess of unlabelled male genomic DNA, so that highlighted regions correspond to female specific and enriched sequences. The WZ bivalent was easily detected in female pachytene nuclei of *I. leuconoe*, where GISH strongly labelled a small, nearly dot-like, W sex chromosome. The larger Z chromosome is typically wrapped around the W in most nuclei (Figure S11a). In few nuclei, the W was spread into more rod-like shape, however, still much smaller than the Z chromosome wrapped around it (not shown). In *T. septentrionis*, the female genomic probe strongly hybridized to several autosomal bivalents. The three most strongly labelled autosomes were significantly smaller than the other chromosomes in the

karyotype. Small signals were detected also on chromosome ends of many bivalents. These results point to the presence of highly repetitive sequences in subtelomeric regions and their amplification on some chromosomes. The signal labelling the W sex chromosome was rather weak (Figure S11b). Only one half of the W chromosome was marked by the probe (Figure S11b,c). In *L. halia*, the female genomic probe highlighted almost all the W sex chromosome which twisted around its Z chromosome counterpart (Figure S11d,e), which is rather typical synaptic adjustment of non-homologous lepidopteran chromosomes (Marec, 1996).

4 | DISCUSSION

4.1 | Sex chromosome–autosome fusions in the tribe Danaini

In the present study, we investigated the evolutionary origin of the *Danaus* neo-sex chromosomes within the tribe Danaini. We sequenced, de novo assembled and annotated genomes of four Danaini species, namely, *Tirumala septentrionis* (Danaina), *Ideopsis similis* (Amaurina), *Idea leuconoe* (Euploeina) and *Lycorea halia* (Itunina).

Analysing sex-specific sequencing coverage successfully identified sex-linked sequences in all four genomes. We then combined these results with orthology-based assignment to $n=31$ chromosomes of *M. cinxia* (Ahola et al., 2014; Smolander et al., 2022), which represent ancestral synteny blocks referred to as Merian elements by Wright et al. (2023). This confirmed the evolutionary conservation of the ancestral lepidopteran Z chromosome, as the corresponding McChr1 remains sex linked across Danaini (Figure 3). The neo-sex chromosome formed by fusion between the ancestral Z and the synteny block corresponding to the McChr21 [F(1;21)] in a common ancestor of *Danaus* spp. (Mongue et al., 2017; Singh et al., 2022) was recovered in *T. septentrionis* (Danaina) as well as in *I. similis* (Amaurina) and *I. leuconoe* (Euploeina). The tightly packed W chromosome of *I. leuconoe* prevents us from determining its

molecular composition. However, GISH signal differentiated only a part of the W chromosome in *T. septentrionis*, which is similar to the bipartite structure of the W chromosome observed in *D. plexippus* (cf. Mongue et al., 2017) and agrees with the presence of the neo-sex chromosomes (Figure S11b,c). Available dating implies the fusion occurred in a common ancestor of these three subtribes at least 26 Mya (Chazot et al., 2019). The F(1;21) is, thus, much older than previously assumed (Mongue et al., 2017).

Our results indicate that *I. similis* has an even more complex sex chromosome system. In addition to F(1;21), it involves another originally autosomal synteny block corresponding to McChr31. Hypothetically, this could have occurred by an additional Z-autosome fusion (i.e. Z chromosome was formed by F(1;21;31)). However, since our current assembly gives no indication of a fusion for McChr31, we favour the hypothesis that a W-autosome fusion [F(W;31)] occurred in this lineage, producing two distinct Z chromosomes. The system, thus, may resemble the multiple sex chromosomes observed in *D. chrysippus* (Martin et al., 2020; Smith et al., 2016).

Remarkably, we discovered an independently evolved neo-Z chromosome in *L. halia* (Figure 3). The coverage analysis identified an autosome homeologous to McChr12 as Z-linked in *L. halia*, which represents the early diverging subtribe Itunina. Similar to the ancestral *L. halia* Z chromosome, the chromosome homeologous to McChr12 has the coverage $\text{Log}_2(\text{M:F})$ close to 1, which suggests that there is essentially no similarity between this new portion of the neo-Z chromosome and its maternally inherited counterpart, the neo-W, if present (cf. Mongue et al., 2017). This is in agreement with the GISH results, which show nearly all the W chromosome to be well differentiated (Figure S11d,e). Thus, it is reasonable to assume that the *L. halia* neo-sex chromosome results from an independent fusion between the ancestral Danaini Z chromosome and McChr12 [F(1;12)]. The F(1;12) is likely a synapomorphy for the genus *Lycorea* as $n=30$ was reported in two species, while ancestral chromosome number $n=31$ was observed in the sister genus *Anetia* (Brown et al., 2004).

Sex chromosome turnover observed in Danaini supports the notion that sex chromosome-autosome fusions are common in Lepidoptera, despite them being a female heterogametic group (Carabajal Paladino et al., 2019). Indeed, a recent comparison of over 200 chromosome-level lepidopteran genomes confirmed that the Z chromosome was involved in the highest number of fusion events (Wright et al., 2023). This is unexpected as sex chromosome-autosome fusions should be in theory relatively rare in taxa with high chromosome numbers such as Lepidoptera, assuming they are random (Anderson et al., 2020).

4.2 | Drivers of sex chromosome turnover in Lepidoptera

It was shown that small and repeat-rich chromosomes are predominantly involved in chromosome fusions in moths and butterflies (Ahola et al., 2014; Wright et al., 2023). However, the Z chromosome,

which is the most frequently fused chromosome, is usually the largest element in lepidopteran karyotype, albeit with repeat content higher than predicted by its length (Wright et al., 2023). High repeat content could make sex chromosomes more prone to fusions through ectopic recombination (Ahola et al., 2014; Lucek et al., 2023). However, it seems unlikely this influences the spread of the resulting neo-sex chromosomes in lepidopteran populations. While small chromosomes are indeed disproportionately involved in autosomal fusions, there is notable variation among which ones fuse with sex chromosomes (Wright et al., 2023). Whether this pattern can be sufficiently explained by genetic drift alone is questionable (cf. Nguyen & Carabajal Paladino, 2016), and highlights the role that selection potentially plays in fixing neo-sex chromosomes, especially the possibility that the fate of a fusion may depend on particular features of the fused autosomes, such as their gene content.

For example, failure to compensate dosage-sensitive genes can constrain sex chromosome evolution (Adolfsson & Ellegren, 2013) and the presence of these genes could have prevented fixation of fusions between sex chromosomes and small Merian elements (ME) 27 and 28 (Wright et al., 2023). Similarly, it was proposed that gene content can predispose certain synteny blocks for a role in sex determination (Graves & Peichel, 2010; O'Meally et al., 2012; Pan et al., 2021). An excess of genes experiencing sexually antagonistic selection might also increase the probability of fixation for a neo-sex chromosome because sex linkage offers an advantage over being autosomal for such genes (Charlesworth & Charlesworth, 1980; Rice, 1984). For instance, ME25 carries over half of all ovary-specific genes in a silkworm *Bombyx mori* (Suetsugu et al., 2013). Provaznikova et al. (2023) hypothesized becoming sex-linked would offer a selective advantage for this cluster of female-specific genes due to sexual antagonism, and that this underlies the repeated involvement of ME25 in lepidopteran neo-sex chromosomes (Hill et al., 2019; Provaznikova et al., 2023; Steward et al., 2021; Wright et al., 2023).

To investigate whether sexual antagonism could be promoting the fixation of sex chromosome-autosome fusions in Danaini, we identified sex-biased genes in *I. leuconoe* and *L. halia* and assessed their distribution across chromosomes (Figure 4). Notably, McChr24 was significantly depleted in male-biased genes in both species. This could reflect the fact that this synteny block was likely part of an X chromosome in ancestral insect male heterogametic system (Toups & Vicoso, 2023). Thus, for more than 150 My, McChr24 would have spent two thirds of its evolutionary time in females and got feminized. Our results also showed that both the ancestral Z chromosome and the autosome homeologous to *M. cinxia* chromosome 31 (corresponding to ME31) are significantly depleted in female-biased genes, although this depletion is significant only for the ancestral Z chromosome in *L. halia*. The same chromosomes are also enriched in male-biased genes in both species. ME31 becoming Z-linked could be advantageous due to the excess of male-beneficial genes it carries. This may have actually driven the fixation of F(W;31) in *I. similis*, which is an intriguing possibility. Nonetheless, if the male bias for ME31 in Danaini is representative of other Lepidoptera, it offers an explanation as to why ME31 was the most common autosome

observed in fusions with lepidopteran sex chromosomes (Wright et al., 2023).

Similarly, the autosome corresponding to McChr21 was significantly enriched in female-biased genes in *L. halia*. It is tempting to speculate that this enrichment represents the ancestral state, which could have driven the fixation of the F(1;21) in a common ancestor of the *Danaina*, *Amaurina* and *Euploeina* lineages. In this case, the fusion would be advantageous to females because it produces linkage between genes important for female function and female sex. However, such fusion would be an evolutionary trap due to differentiation of the resulting neo-sex chromosomes. The non-recombining neo-W chromosome would have inevitably degenerated and the female-biased genes survived only on the neo-Z chromosome. However, hemizygous expression of unbiased or female-biased genes in females could create another round of strong selection favouring sex chromosome-autosome fusion (Mongue et al., 2022), despite the loss of neo-W-linked alleles. Yet, female-biased genes would spend two thirds of their time in males and the resulting sexual conflict could reduce their content as observed on McChr21 in *I. leuconoe* (compared to *L. halia*) (Figure 4).

Interestingly, in both *L. halia* and *I. leuconoe*, the homeologues of McChr28 (corresponding to ME25) are relatively (though not significantly) depauperate for female-biased genes; this contrasts with *B. mori*, where this synteny block carries over half of all ovary-specific genes (Suetsugu et al., 2013). The discrepancy could be ascribed to iteroparity of the studied species, which emerge with few maturing eggs (in contrast to *B. mori*) and thus do not yet express some ovary-specific genes such as chorion genes in the sampled tissue. However, we cannot exclude that it reflects the relatively rapid evolution of sex-biased genes (Catalan et al., 2018; Dapper & Wade, 2020; Grath & Parsch, 2016).

It seems that while chances of any fusion are affected by chromosome repeat content (Ahola et al., 2014; Wright et al., 2023), enrichment in sex-biased genes could promote the fixation of sex chromosome-autosome fusions. This is in agreement with sexually antagonistic selection driving spread of neo-sex chromosome (Charlesworth & Charlesworth, 1980; Matsumoto et al., 2016) as sex-biased genes have been used as a proxy for sexual antagonism (Grath & Parsch, 2016; Mank, 2017).

4.3 | Dosage compensation

While sexual antagonism may underlie evolution of sex-biased genes, these are not necessarily sexually antagonistic (Mank, 2017; Parsch & Ellegren, 2013). Interpretation of sex-biased gene expression is complicated for sex chromosomes and gonads due to specific regulatory mechanisms such as meiotic sex chromosome inactivation (MSCI) or dosage compensation (Mank & Ellegren, 2009; Parsch & Ellegren, 2013). Little is known about the MSCI in Lepidoptera (Traut et al., 2019) and in *D. plexippus*, it was shown that ancestral and new parts of neo-sex chromosomes are compensated by two distinct mechanisms (Gu et al., 2019).

Patterns of dosage compensation in *L. halia* and *I. leuconoe* are grossly consistent with those reported for *D. plexippus* and also other lepidopterans generally. As reported previously, the ancestral portion of the Z chromosome appears to be reduced similarly in both sexes relative to autosomes and the neo-Z (Catalan et al., 2018; Huylmans et al., 2017; Walters et al., 2015; Walters & Hardcastle, 2011). However, the magnitude (and thus statistical significance) of this pattern varies by tissue and species. For instance, previous analyses of *D. plexippus* reported that in heads, the ancestral Z:autosome median ratios were about 0.55 (Gu et al., 2019; Ranz et al., 2021). However, in the two species assayed here, the ratios in heads ranged from 0.7 to 0.9 and showed significant differences only in *L. halia*. Meanwhile, in thorax, the ancestral Z was significantly reduced in all samples, while *D. plexippus* thorax yielded a notable dosage effect between sexes (Ranz et al., 2021). Patterns on the new Z segments were much more consistent across species and tissues, where the Znew segments appear near-fully compensated, without any reduction in average expression relative to autosomes. Directly comparing male:female expression ratios is particularly useful for detecting subtle dosage effects on Z-linked expression (Gu & Walters, 2017). In both species studied here, somatic tissues yielded a very modest dosage effect that manifested as a slight male bias on Z chromosome for both ancestral and neo segments. This slight dosage effect was not reported previously in *D. plexippus*, but is similar to observations made in *Heliconius* butterflies (Catalan et al., 2018; Gu et al., 2019; Ranz et al., 2021; Walters et al., 2015).

It is noteworthy that the new Z segment in *L. halia* exhibits nearly complete dosage compensation. This appears to represent the independent evolution of a dichotomous pattern of dosage compensation between the new versus ancestral Z segments. This pattern is comparable to that reported in *D. plexippus* but with a completely different autosome fused with the ancestral Z (Gu et al., 2019). Essentially complete compensation for neo-Z chromosomes in *Danaini* contrasts with the lack of such compensation reported for the neo-Z in *Cydia pomonella*, which was balanced between sexes, but reduced relative to autosomes, similar to the ancestral Z (Gu et al., 2017). Thus, it appears that considerable variation exists among lepidopteran lineages concerning the trajectory of dosage compensation following Z-autosome fusion. Assessing dosage compensation in *I. similis*, which carries yet another independently evolved neo-Z segment, will be quite interesting, as will further examining to what extent shared molecular mechanisms underlie these patterns.

Dosage compensation analysis of neo-Z chromosomes in gonads also lends some insight into a lingering debate concerning whether or not such molecular mechanisms function in reproductive tissues. As observed here, male-versus-female comparisons of lepidopteran gonadal gene expression often show a pattern consistent with the absence of dosage compensation for the ancestral Z (Catalan et al., 2018; Gu et al., 2017; Gu & Walters, 2017; Huylmans et al., 2017). However, this pattern may also be explained by an excess of highly expressed male-specific genes on the Z chromosome,

as predicted by sexual antagonism theory and often empirically observed (Catalan et al., 2018; Huylmans et al., 2017; Mongue & Walters, 2018; Walters & Hardcastle, 2011). This latter explanation is supported by the observation that the new Z segments are not reduced relative to autosomes in females, as would be expected in the absence of dosage compensating mechanisms. However, this argument also assumes that such mechanisms are shared and similarly controlled between the ancestral and new segments of the Z chromosome, which might not be the case.

4.4 | Conclusion

To conclude, sex chromosome evolution in the tribe Danaini involves independent sex chromosome-autosome rearrangements and further underlines high incidence of neo-sex chromosomes in Lepidoptera (cf. Pennell et al., 2015; Wright et al., 2023). Sex-biased gene content could facilitate spread and fixation of neo-sex chromosomes in Lepidoptera, suggesting a role of sexual antagonism in their sex chromosome turnover. Moreover, genomic resources generated in this study will advance investigations into interesting biology and life history of Danaini butterflies such as their monsoon-driven migrations (Bhaumik & Kunte, 2017).

AUTHOR CONTRIBUTIONS

PN and JRW designed the study. PM, MH, AV, PK, JŠ, MD and JRW performed research. PM, MH, AV, PK, JŠ, MD and JRW analysed data. PM, MH, JRW and PN wrote the manuscript.

ACKNOWLEDGEMENTS

We would like to thank Daniel Linke for his kind help with pinning and spreading the butterflies under study and their photodocumentation and the editor and three anonymous reviewers for their stimulating feedback. This research was funded by the Czech-American scientific cooperation grant (Reg. No. LTAUSA19086) under the INTER-ACTION program of the Ministry of Education, Youth and Sports of the Czech Republic. Pablo Mora thanks the University of Jaén for his "Convocatoria de Recualificación del Sistema Universitario Español-Margarita Salas" postdoctoral grant under the "Plan de Recuperación Transformación" program funded by the Spanish Ministry of Universities with European Union's NextGenerationEU funds (grant no. UJAR10MS). Computational resources were supplied by the project "e-Infraestructura CZ" (e-INFRA CZ LM2018140) supported by the Ministry of Education, Youth and Sports of the Czech Republic and by the ELIXIR-CZ project (LM2018131), part of the international ELIXIR infrastructure. James Walters was supported by the Fulbright Foundation of Czechia and US National Science Foundation grants NSF-ABI-166145 and DBI-2213824.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The experimental data supporting the results of this study are available in this article or in the [Supporting Information](#). The raw data have been deposited in NCBI under the bioproject accession no. PRJNA1005810. The genome assemblies and annotation files were deposited in the Dryad repository: <https://doi.org/10.5061/dryad.hmgqnk9p2>.

ORCID

Petr Nguyen  <https://orcid.org/0000-0003-1395-4287>

REFERENCES

- Adolfsson, S., & Ellegren, H. (2013). Lack of dosage compensation accompanies the arrested stage of sex chromosome evolution in ostriches. *Molecular Biology and Evolution*, 30(4), 806–810. <https://doi.org/10.1093/molbev/mst009>
- Ahola, V., Lehtonen, R., Somervuo, P., Salmela, L., Koskinen, P., Rastas, P., Välimäki, N., Paulin, L., Kvist, J., Wahlberg, N., Tanskanen, J., Horneet, E. A., Ferguson, L. C., Luo, S., Cao, Z., de Jong, M. A., Duploux, A., Smolander, O. P., Vogel, H., ... Hanski, I. (2014). The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nature Communications*, 5, 4737. <https://doi.org/10.1038/ncomms5737>
- Anderson, N. W., Hjelmén, C. E., & Blackmon, H. (2020). The probability of fusions joining sex chromosomes and autosomes. *Biology Letters*, 16(11), 20200648. <https://doi.org/10.1098/rsbl.2020.0648>
- Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Bachtrog, D. (2013). Y-chromosome evolution: Emerging insights into processes of Y-chromosome degeneration. *Nature Reviews Genetics*, 14, 113–124. <https://doi.org/10.1038/nrg3366>
- Bhaumik, V., & Kunte, K. (2017). Female butterflies modulate investment in reproduction and flight in response to monsoon-driven migrations. *Oikos*, 127, 285–296. <https://doi.org/10.1111/oik.04593>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Brown, K. S., von Schoutz, B., & Suomalainen, E. (2004). Chromosome evolution in neotropical Danainae and Ithomiinae (Lepidoptera). *Hereditas*, 141(3), 216–236.
- Brüna, T., Hoff, K. J., Lomsadze, A., Stanke, M., & Borodovsky, M. (2021). BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics and Bioinformatics*, 3, lqaa108. <https://doi.org/10.1093/nargab/lqaa108>
- Buntrock, L., Marec, F., Krueger, S., & Traut, W. (2012). Organ growth without cell division: Somatic polyploidy in a moth, *Ephesia kuehniella*. *Genome*, 55(11), 755–763. <https://doi.org/10.1139/g2012-060>
- Carabajal Paladino, L. Z., Provazníková, I., Berger, M., Bass, C., Aratchige, N. S., López, S. N., Marec, F., & Nguyen, P. (2019). Sex chromosome turnover in moths of the diverse superfamily Gelechioidea. *Genome Biology and Evolution*, 11(4), 1307–1319. <https://doi.org/10.1093/gbe/evz075>
- Catalan, A., Macias-Munoz, A., & Briscoe, A. D. (2018). Evolution of sex-biased gene expression and dosage compensation in the eye and brain of *Heliconius* butterflies. *Molecular Biology and Evolution*, 35(9), 2120–2134. <https://doi.org/10.1093/molbev/msy111>

- Chan, P. P., & Lowe, T. M. (2019). tRNAscan-SE: searching for tRNA genes in genomic sequences. In M. Kollmar (Ed.), *Gene prediction. Methods in molecular biology*, vol. 1962 (pp. 1–14). Humana. https://doi.org/10.1007/978-1-4939-9173-0_1
- Charlesworth, D., & Charlesworth, B. (1980). Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genetics Research*, 35(2), 205–214. <https://doi.org/10.1017/S0016672300014051>
- Chazot, N., Willmott, K. R., Lamas, G., Freitas, A. V. L., Piron-Prunier, F., Arias, C. F., Mallet, J., de-Silva, D. L., & Elias, M. (2019). Renewed diversification following Miocene landscape turnover in a Neotropical butterfly radiation. *Global Ecology and Biogeography*, 28(8), 1118–1132. <https://doi.org/10.1111/geb.12919>
- Cheng, C. D., & Kirkpatrick, M. (2016). Sex-specific selection and sex-biased gene expression in humans and flies. *PLoS Genetics*, 12(9), e1006170. <https://doi.org/10.1371/journal.pgen.1006170>
- Dagilis, A. J., Sardell, J. M., Josephson, M. P., Su, Y. H., Kirkpatrick, M., & Peichel, C. L. (2022). Searching for signatures of sexually antagonistic selection on stickleback sex chromosomes. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 377(1856), 20210205. <https://doi.org/10.1098/rstb.2021.0205>
- Dapper, A. L., & Wade, M. J. (2020). Relaxed selection and the rapid evolution of reproductive genes. *Trends in Genetics*, 36(9), 640–649. <https://doi.org/10.1016/j.tig.2020.06.014>
- De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, 34(15), 2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
- Delhomme, N., Padiou, I., Furlong, E. E., & Steinmetz, L. M. (2012). easyRNASeq: A bioconductor package for processing RNA-Seq data. *Bioinformatics*, 28(19), 2532–2533. <https://doi.org/10.1093/bioinformatics/bts477>
- Dobin, A., & Gingeras, T. R. (2015). Mapping RNA-seq reads with STAR. *Current Protocols in Bioinformatics*, 51, 11141–19. <https://doi.org/10.1002/0471250953.bi111451>
- Ellegren, H. (2011). Sex-chromosome evolution: Recent progress and the influence of male and female heterogamety. *Nature Reviews Genetics*, 12(3), 157–166. <https://doi.org/10.1038/nrg2948>
- Filatov, D. A. (2018). The two "rules of speciation" in species with young sex chromosomes. *Molecular Ecology*, 27(19), 3799–3810. <https://doi.org/10.1111/mec.14721>
- Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of National Academy of Sciences USA*, 117(17), 9451–9457. <https://doi.org/10.1073/pnas.1921046117>
- Fraïsse, C., Picard, M. A. L., & Vicoso, B. (2017). The deep conservation of the Lepidoptera Z chromosome suggests a non-canonical origin of the W. *Nature Communications*, 8(1), 1486. <https://doi.org/10.1038/s41467-017-01663-5>
- Furman, B. L. S., Metzger, D. C. H., Darolti, I., Wright, A. E., Sandkam, B. A., Almeida, P., Shu, J. J., & Mank, J. E. (2020). Sex chromosome evolution: So many exceptions to the rules. *Genome Biology and Evolution*, 12(6), 750–763. <https://doi.org/10.1093/gbe/evaa081>
- Grath, S., & Parsch, J. (2016). Sex-biased gene expression. *Annual Review of Genetics*, 50, 29–44.
- Graves, J. A. M. (2016). Did sex chromosome turnover promote divergence of the major mammal groups?: De novo sex chromosomes and drastic rearrangements may have posed reproductive barriers between monotremes, marsupials and placental mammals. *BioEssays*, 38(8), 734–743. <https://doi.org/10.1002/bies.201600019>
- Graves, J. A. M., & Peichel, C. L. (2010). Are homologies in vertebrate sex determination due to shared ancestry or to limited options? *Genome Biology*, 11(4), 205. <https://doi.org/10.1186/gb-2010-11-4-205>
- Gu, L., Reilly, P. F., Lewis, J. J., Reed, R. D., Andolfatto, P., & Walters, J. R. (2019). Dichotomy of dosage compensation along the neo Z chromosome of the monarch butterfly. *Current Biology*, 29(23), 4071–4077. <https://doi.org/10.1016/j.cub.2019.09.056>
- Gu, L. Q., & Walters, J. R. (2017). Evolution of sex chromosome dosage compensation in animals: A beautiful theory, undermined by facts and bedeviled by details. *Genome Biology and Evolution*, 9(9), 2461–2476. <https://doi.org/10.1093/gbe/evx154>
- Gu, L. Q., Walters, J. R., & Knipple, D. C. (2017). Conserved patterns of sex chromosome dosage compensation in the Lepidoptera (WZ/ZZ): Insights from a moth neo-Z chromosome. *Genome Biology and Evolution*, 9(3), 802–816. <https://doi.org/10.1093/gbe/evx039>
- Guan, D., McCarthy, S. A., Wood, J., Howe, K., Wang, Y., & Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*, 36(9), 2896–2898. <https://doi.org/10.1093/bioinformatics/btaa025>
- Herbert, P. N. D., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences*, 101(41), 14812–14817. <https://doi.org/10.1073/pnas.04061661>
- Hill, J., Rastas, P., Hornett, E. A., Neethiraj, R., Clark, N., Morehouse, N., de la Paz Celorio-Mancera, M., Cols, J. C., Dirksen, H., Meslin, C., Keehnen, N., Prüsscher, P., Sikkink, K., Vives, M., Vogel, H., Wiklund, C., Woronik, A., Boggs, C. L., Nylin, S., & Wheat, C. W. (2019). Unprecedented reorganization of holocentric chromosomes provides insights into the enigma of lepidopteran chromosome evolution. *Science Advances*, 5(6), eaau3648. <https://doi.org/10.1126/sciadv.aau3648>
- Holley, G., Beyter, D., Ingimundardottir, H., Møller, P. L., Kristmundsdottir, S., Eggertsson, H. P., & Halldorsson, B. V. (2021). Ratatosk: Hybrid error correction of long reads enables accurate variant calling and assembly. *Genome Biology*, 22(1), 28. <https://doi.org/10.1186/s13059-020-02244-4>
- Hu, J., Fan, J., Sun, Z., & Liu, S. (2020). NextPolish: A fast and efficient genome polishing tool for long-read assembly. *Bioinformatics*, 36(7), 2253–2255. <https://doi.org/10.1093/bioinformatics/btz891>
- Humann, J. L., Lee, T., Ficklin, S., & Main, D. (2019). Structural and functional annotation of eukaryotic genomes with GenSAS. In M. Kollmar (Ed.), *Gene Prediction. Methods in Molecular Biology*, vol. 1962. Humana. https://doi.org/10.1007/978-1-4939-9173-0_3
- Huylmans, A. K., Macon, A., & Vicoso, B. (2017). Global dosage compensation is ubiquitous in Lepidoptera, but counteracted by the masculinization of the Z chromosome. *Molecular Biology and Evolution*, 34(10), 2637–2649. <https://doi.org/10.1093/molbev/msx190>
- Innocenti, P., & Morrow, E. H. (2010). The sexually antagonistic genes of *Drosophila melanogaster*. *PLoS Biology*, 8(3), e1000335. <https://doi.org/10.1371/journal.pbio.1000335>
- Kirkpatrick, M., & Guerrero, R. F. (2014). Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics*, 197(2), 531. <https://doi.org/10.1534/genetics.113.156026>
- Kitano, J., & Peichel, C. L. (2012). Turnover of sex chromosomes and speciation in fishes. *Environmental Biology of Fishes*, 94(3), 549–558. <https://doi.org/10.1007/s10641-011-9853-8>
- Kitano, J., Ross, J. A., Mori, S., Kume, M., Jones, F. C., Chan, Y. F., Absher, D. M., Grimwood, J., Schmutz, J., Myers, R. M., Kingsley, D. M., & Peichel, C. L. (2009). A role for a neo-sex chromosome in stickleback speciation. *Nature*, 461(7267), 1079–1083. <https://doi.org/10.1038/nature08441>
- Kiuchi, T., Koga, H., Kawamoto, M., Shoji, K., Sakai, H., Arai, Y., Ishihara, G., Kawaoka, S., Sugano, S., Shimada, T., Suzuki, Y., Suzuki, M. G., & Katsuma, S. (2014). A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature*, 509, 633–636. <https://doi.org/10.1038/nature13315>

- Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, 37(5), 540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- Kopylova, E., Noé, L., & Touzet, H. (2012). SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, 28(24), 3211–3217. <https://doi.org/10.1093/bioinformatics/bts611>
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome Biology*, 5(2), R12. <https://doi.org/10.1186/gb-2004-5-2-r12>
- Laetsch, D. R., & Blaxter, M. L. (2017). BlobTools: Interrogation of genome assemblies. *F1000Research*, 6, 1287. <https://doi.org/10.12688/f1000research.12232.1>
- Lagesen, K., Hallin, P., Rødland, E. A., Stærfeldt, H.-H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, 35(9), 3100–3108. <https://doi.org/10.1093/nar/gkm160>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lasne, C., Sgro, C. M., & Connallon, T. (2017). The relative contributions of the X chromosome and autosomes to local adaptation. *Genetics*, 205(3), 1285–1304. <https://doi.org/10.1534/genetics.116.194670>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Lichilin, N., El Taher, A., & Bohne, A. (2021). Sex-biased gene expression and recent sex chromosome turnover. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 376(1833), 20200107. <https://doi.org/10.1098/rstb.2020.0107>
- Lomsadze, A., Burns, P. D., & Borodovsky, M. (2014). Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Research*, 42(15), e119. <https://doi.org/10.1093/nar/gku557>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Lucek, K., Giménez, M. D., Joron, M., Rafajlović, M., Searle, J. B., Walden, N., Westram, A. M., & Faria, R. (2023). The impact of chromosomal rearrangements in speciation: From micro- to macroevolution. *Cold Spring Harbor Perspectives in Biology*, 15, a041447. <https://doi.org/10.1101/cshperspect.a041447>
- Lund-Hansen, K. K., Olito, C., Morrow, E. H., & Abbott, J. K. (2021). Sexually antagonistic coevolution between the sex chromosomes of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 118(8), e2003359118. <https://doi.org/10.1073/pnas.2003359118>
- Maeki, K., & Ae, S. A. (1969). Studies of the chromosomes of Formosan Rhopalocera 4. Danaidae and Satyridae. *Kontyu*, 37(1), 99–109.
- Mank, J. E. (2017). The transcriptional architecture of phenotypic dimorphism. *Nature Ecology & Evolution*, 1(1). <https://doi.org/10.1038/s41559-016-0006>
- Mank, J. E., & Ellegren, H. (2009). Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. *Evolution*, 63(6), 1464–1472. <https://doi.org/10.1111/j.1558-5646.2009.00618.x>
- Mank, J. E., Hosken, D. J., & Wedell, N. (2014). Conflict on the sex chromosomes: Cause, effect, and complexity. *Cold Spring Harbor Perspectives in Biology*, 6(12), a017715. <https://doi.org/10.1101/cshperspect.a017715>
- Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, 38(10), 4647–4654. <https://doi.org/10.1093/molbev/msab199>
- Marçais, G., & Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, 27(6), 764–770. <https://doi.org/10.1093/bioinformatics/btr011>
- Marec, F. (1996). Synaptonemal complexes in insects. *International Journal of Insect Morphology and Embryology*, 25(3), 205–233. [https://doi.org/10.1016/0020-7322\(96\)00009-8](https://doi.org/10.1016/0020-7322(96)00009-8)
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, 17(1), 10–12.
- Martin, S. H., Singh, K. S., Gordon, I. J., Omufwoko, K. S., Collins, S., Warren, I. A., Munby, H., Brattström, O., Traut, W., Martins, D. J., Smith, D. A. S., Jiggins, C. D., Bass, C., & French-Constant, R. H. (2020). Whole-chromosome hitchhiking driven by a male-killing endosymbiont. *PLoS Biology*, 18(2), e3000610. <https://doi.org/10.1371/journal.pbio.3000610>
- Matsumoto, T., Yoshida, K., & Kitano, J. (2016). Gene flow and sex chromosome evolution. *Genes & Genetic Systems*, 91(6), 357.
- McAllister, B. F., Sheeley, S. L., Mena, P. A., Evans, A. L., & Schlotterer, C. (2008). Clinal distribution of a chromosomal rearrangement: A precursor to chromosomal speciation? *Evolution*, 62(8), 1852–1865. <https://doi.org/10.1111/j.1558-5646.2008.00435.x>
- Mediouni, J., Fukova, I., Frydrychova, R., Dhoubi, M. H., & Marec, F. (2004). Karyotype, sex chromatin and sex chromosome differentiation in the carob moth, *Ectomyelois ceratoniae* (Lepidoptera: Pyralidae). *Caryologia*, 57(2), 184–194.
- Mongue, A. J., Hansen, M. E., & Walters, J. R. (2022). Support for faster and more adaptive Z chromosome evolution in two divergent lepidopteran lineages. *Evolution*, 76(2), 332–345. <https://doi.org/10.1111/evo.14341>
- Mongue, A. J., Nguyen, P., Voleniková, A., & Walters, J. R. (2017). Neo-sex chromosomes in the monarch butterfly, *Danaus plexippus*. *G3: Genes, Genomes, Genetics*, 7(10), 3281–3294. <https://doi.org/10.1534/g3.117.300187>
- Mongue, A. J., & Walters, J. R. (2018). The Z chromosome is enriched for sperm proteins in two divergent species of Lepidoptera. *Genome*, 61(4), 248–253. <https://doi.org/10.1139/gen-2017-0068>
- Mora, P., Hospodarska, M., Volenikova, A., Koutecky, P., Stundlova, J., Dalikova, M., Walters, J. R., & Nguyen, P. (2023). Sex chromosome evolution in the tribe Danaini. *Dryad, Dataset*. <https://doi.org/10.5061/dryad.hmgqnk9p2>
- Nguyen, P., & Carabajal Paladino, L. (2016). On the neo-sex chromosomes of Lepidoptera. In P. Pontarotti (Ed.), *Evolutionary biology: Convergent evolution, evolution of complex traits, concepts and methods* (pp. 171–185). Springer.
- Nguyen, P., Sýkorová, M., Šichová, J., Kůta, V., Dalíková, M., Čapková Frydrychová, R., Neven, L. G., Sahara, K., & Marec, F. (2013). Neo-sex chromosomes and adaptive potential in tortricid pests. *Proceedings of the National Academy of Sciences*, 110(17), 6931–6936. <https://doi.org/10.1073/pnas.1220372110>
- Novák, P., Ávila Robledillo, L., Kobližková, A., Vrbová, I., Neumann, P., & Macas, J. (2017). TAREAN: A computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Research*, 45(12), e111. <https://doi.org/10.1093/nar/gkx257>
- O'Meally, D., Ezaz, T., Georges, A., Sarre, S. D., & Graves, J. A. M. (2012). Are some chromosomes particularly good at sex? Insights from amniotes. *Chromosome Research*, 20(1), 7–19. <https://doi.org/10.1007/s10577-011-9266-8>
- Oshlack, A., Robinson, M. D., & Young, M. D. (2010). From RNA-seq reads to differential expression results. *Genome Biology*, 11(12), 220. <https://doi.org/10.1186/gb-2010-11-12-220>

- Palmer, D. H., Rogers, T. F., Dean, R., & Wright, A. E. (2019). How to identify sex chromosomes and their turnover. *Molecular Ecology*, 28(21), 4709–4724. <https://doi.org/10.1111/mec.15245>
- Pan, Q. W., Kay, T., Depince, A., Adolphi, M., Scharl, M., Guiguen, Y., & Herpin, A. (2021). Evolution of master sex determiners: TGF-beta signalling pathways at regulatory crossroads. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 376(1832), 20200091. <https://doi.org/10.1098/rstb.2020.0091>
- Parsch, J., & Ellegren, H. (2013). The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews Genetics*, 14(2), 83–87. <https://doi.org/10.1038/nrg3376>
- Payseur, B. A., Presgraves, D. C., & Filatov, D. A. (2018). Introduction: Sex chromosomes and speciation. *Molecular Ecology*, 27(19), 3745–3748. <https://doi.org/10.1111/mec.14828>
- Pedersen, B. S., Collins, R. L., Talkowski, M. E., & Quinlan, A. R. (2017). Indexcov: Fast coverage quality control for whole-genome sequencing. *GigaScience*, 6(11), 1–6. <https://doi.org/10.1093/gigascience/gix090>
- Pennell, M. W., Kirkpatrick, M., Otto, S. P., Vamori, J. C., Peichel, C. L., Valenzuela, N., & Kitano, J. (2015). Y fuse? Sex chromosome fusions in fishes and reptiles. *PLoS Genetics*, 11(5), e1005237. <https://doi.org/10.1371/journal.pgen.1005237>
- Perrin, N. (2021). Sex-chromosome evolution in frogs: What role for sex-antagonistic genes? *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 376(1832), 20200094. <https://doi.org/10.1098/rstb.2020.0094>
- Pohlert, T. (2022). PMCMRplus: Calculate pairwise multiple comparisons of mean rank sums extended.
- Pokorna, M., Altmanova, M., & Kratochvil, L. (2014). Multiple sex chromosomes in the light of female meiotic drive in amniote vertebrates. *Chromosome Research*, 22(1), 35–44. <https://doi.org/10.1007/s10577-014-9403-2>
- Provaznikova, I., Dalikova, M., Volenikova, A., Roessingh, P., Sahara, K., Provaznik, J., Marec, F., & Nguyen, P. (2023). Multiple sex chromosomes of *Yponomeuta* ermine moths suggest a role of sexual antagonism in sex chromosome turnover in Lepidoptera. *bioRxiv*. <https://doi.org/10.1101/2023.06.06.543653>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Ranz, J. M., González, P. M., Clifton, B. D., Nazario-Yepiz, N. O., Hernández-Cervantes, P. L., Palma-Martínez, M. J., Valdivia, D. I., Jiménez-Kaufman, A., Lu, M. M., Markow, T. A., & Abreu-Goodger, C. (2021). A de novo transcriptional atlas in *Danaus plexippus* reveals variability in dosage compensation across tissues. *Communications Biology*, 4, 791. <https://doi.org/10.1038/s42003-021-02335-3>
- Rice, W. R. (1984). Sex-chromosomes and the evolution of sexual dimorphism. *Evolution*, 38(4), 735–742. <https://doi.org/10.2307/2408385>
- Rice, W. R. (1987). The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution*, 41(4), 911–914. <https://doi.org/10.1111/j.1558-5646.1987.tb05864.x>
- Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11(3), R25. <https://doi.org/10.1186/gb-2010-11-3-r25>
- Ross, J. A., Urton, J. R., Boland, J., Shapiro, M. D., & Peichel, C. L. (2009). Turnover of sex chromosomes in the stickleback fishes (Gasterosteidae). *PLoS Genetics*, 5(2), e1000391. <https://doi.org/10.1371/journal.pgen.1000391>
- Sahara, K., Marec, F., & Traut, W. (1999). TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Research*, 7(6), 449–460. <https://doi.org/10.1023/a:1009297729547>
- Schurich, N. J., Schofield, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., Owen-Hughes, T., Blaxter, M., & Barton, G. J. (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA*, 22(6), 839–851. <https://doi.org/10.1261/rna.053959.115>
- Šíchová, J., Nguyen, P., Dalíková, M., & Marec, F. (2013). Chromosomal evolution in tortricid moths: Conserved karyotypes with diverged features. *PLoS One*, 8(5), e64520. <https://doi.org/10.1371/journal.pone.0064520>
- 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*, 15, 89. <https://doi.org/10.1186/s12862-015-0375-4>
- Sigeman, H., Ponnikas, S., Chauhan, P., Dierickx, E., Brooke, M. D., & Hansson, B. (2019). Repeated sex chromosome evolution in vertebrates supported by expanded avian sex chromosomes. *Proceedings of the Royal Society B: Biological Sciences*, 286(1916), 20192051. <https://doi.org/10.1098/rspb.2019.2051>
- Singh, K. S., De-Kayne, R., Omufwoko, K. S., Martins, D. J., Bass, C., French-Constant, R., & Martin, S. H. (2022). Genome assembly of *Danaus chrysippus* and comparison with the monarch *Danaus plexippus*. *G3: Genes, Genomes, Genetics*, 12, jkab449. <https://doi.org/10.1093/g3journal/jkab449>
- Smit, A. F. A., Hubley, R., & Green, P. (2013–2015). RepeatMasker Open-4.0. <http://www.repeatmasker.org>
- Smith, D. A. S., Gordon, I. J., Traut, W., Herren, J., Collins, S., Martins, D. J., Saitoti, K., Ireri, P., & French-Constant, R. (2016). A neo-W chromosome in a tropical butterfly links colour pattern, male-killing, and speciation. *Proceedings of the Royal Society B: Biological Sciences*, 283(1835), 20160821. <https://doi.org/10.1098/rspb.2016.0821>
- Smolander, O.-P., Blande, D., Ahola, V., Rastas, P., Tanskanen, J., Kammonen, J. I., Oostra, V., Pellegrini, L., Ikonen, S., Dallas, T., DiLeo, M. F., Duploux, A., Duru, I. C., Halimaa, P., Kahilainen, A., Kuwar, S. S., Kärenlampi, S. O., Lafuente, E., Luo, S., ... Saastamoinen, M. (2022). Improved chromosome-level genome assembly of the Glanville fritillary butterfly (*Melitaea cinxia*) integrating Pacific biosciences long reads and a high-density linkage map. *GigaScience*, 11, giab097. <https://doi.org/10.1093/gigascience/giab097>
- Stanke, M., & Morgenstern, B. (2005). AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research*, 33, W465–W467. <https://doi.org/10.1093/nar/gki458>
- Steward, R. A., Okamura, Y., Boggs, C. L., Vogel, H., & Wheat, C. W. (2021). The genome of the margined white butterfly (*Pieris macdunnoughii*): Sex chromosome insights and the power of polishing with PoolSeq data. *Genome Biology and Evolution*, 13(4), evab053. <https://doi.org/10.1093/gbe/evab053>
- Suetsugu, Y., Futahashi, R., Kanamori, H., Kadono-Okuda, K., Sasanuma, S.-I., Narukawa, J., Ajimura, M., Jouraku, A., Namiki, N., Shimomura, M., Sezutsu, H., Osanai-Futahashi, M., Suzuki, M. G., Daimon, T., Shinoda, T., Taniai, K., Asaoka, K., Niwa, R., Kawaoka, S., ... Mita, K. (2013). Large scale full-length cDNA sequencing reveals a unique genomic landscape in a lepidopteran model insect, *Bombyx mori*. *G3: Genes, Genomes, Genetics*, 3(9), 1481–1492. <https://doi.org/10.1534/g3.113.006239>
- Toups, M. A., & Vicoso, B. (2023). The X chromosome of insects likely predates the origin of class Insecta. *Evolution*, 77(11), 2504–2511. <https://doi.org/10.1093/evolut/apad169>
- Traut, W., Schubert, V., Dalíková, M., Marec, F., & Sahara, K. (2019). Activity and inactivity of moth sex chromosomes in somatic and meiotic cells. *Chromosoma*, 128(4), 533–545. <https://doi.org/10.1007/s00412-019-00722-8>
- Veltsos, P., Rodrigues, N., Studer, T., Ma, W. J., Sermier, R., Leuenberger, J., & Perrin, N. (2020). No evidence that Y-chromosome differentiation affects male fitness in a Swiss population of common frogs. *Journal of Evolutionary Biology*, 33(4), 401–409. <https://doi.org/10.1111/jeb.13573>

- Vicoso, B., & Charlesworth, B. (2006). Evolution on the X chromosome: Unusual patterns and processes. *Nature Reviews Genetics*, 7(8), 645–653. <https://doi.org/10.1038/nrg1914>
- Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., & Schatz, M. C. (2017). GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics*, 33(14), 2202–2204. <https://doi.org/10.1093/bioinformatics/btx153>
- Walters, J. R., & Hardcastle, T. J. (2011). Getting a full dose? Reconsidering sex chromosome dosage compensation in the silkworm, *Bombyx mori*. *Genome Biology and Evolution*, 3, 491–504. <https://doi.org/10.1093/gbe/evr036>
- Walters, J. R., Hardcastle, T. J., & Jiggins, C. D. (2015). Sex chromosome dosage compensation in *Heliconius* butterflies: Global yet still incomplete? *Genome Biology and Evolution*, 7(9), 2545–2559. <https://doi.org/10.1093/gbe/evv156>
- Wright, A. E., Dean, R., Zimmer, F., & Mank, J. E. (2016). How to make a sex chromosome. *Nature Communications*, 7, 12087. <https://doi.org/10.1038/ncomms12087>
- Wright, C. J., Stevens, L., Mackintosh, A., Lawniczak, M., & Blaxter, M. (2023). Chromosome Evolution in Lepidoptera. *ssiv*. <https://doi.org/10.1101/2023.05.12.540473>
- Yoshido, A., Šichová, 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(3), 138–154. <https://doi.org/10.1038/s41437-020-0325-9>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mora, P., Hospodářská, M., Voleníková, A. C., Koutecký, P., Štundlová, J., Dalíková, M., Walters, J. R., & Nguyen, P. (2024). Sex-biased gene content is associated with sex chromosome turnover in *Danaini* butterflies. *Molecular Ecology*, 00, e17256. <https://doi.org/10.1111/mec.17256>