A neo-sex chromosome in the Monarch butterfly, *Danaus plexippus*

James R. Walters* and Andrew Mongue

Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, USA

*Author for correspondence: James R Walters, Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, USA
phone: 785-864-6341
email: jrwalters@ku.edu

Abstract

We report the discovery of a neo-sex chromosome in Monarch butterfly, *Danaus plexippus*, and several of its close relatives. Z-linked scaffolds in the *D. plexippus* genome assembly were identified via sex-specific differences in Illumina sequencing coverage. Additionally, a majority of the *D. plexippus* genome assembly was assigned to chromosomes based on counts of 1-to-1 orthologs relative to the butterfly *Melitaea cinxia* (and two other Lepidopteran species), where genome scaffolds have been robustly mapped to linkage groups. Combining sequencing-coverage based Z-linkage with homology based chromosomal assignments provided strong evidence for a Z-autosome fusion in the *Danaus* lineage, involving the autosome homologous to chromosome 21 in *M. cinxia*. Coverage analysis also identified three scaffolds containing notable assembly errors resulting in chimeric Z-autosome fusions. The timing of this Z-autosome fusion event currently remains ambiguous due to incomplete sampling of karyotypes in the Danaini tribe of butterflies. The discovery of a neo-Z and the provisional assignment of chromosome linkage for >90% of *D. plexippus* genes lay the foundation for novel insights concerning sex chromosome evolution in this increasingly prominent female-heterogametic model species for functional and evolutionary genomics.

Introduction

Major rearrangements of karyotype and chromosome structure often have substantial evolutionary impacts on both the organisms carrying such mutations and the genes linked to such genomic reorganization (Lynch 2007; Pamela Soltis & Douglas E Soltis 2012) . Additionally, such large-scale chromosomal mutations often present novel opportunities to investigate molecular evolutionary and functional genetic processes. One prominent example of this is the evolution of neo-sex chromosomes, which can arise from the fusion of an autosome with an existing and well-differentiated allosome. This effectively instantaneous transformation of a formerly autosomal set of genes into sex-linked loci is fertile ground for comparative analyses aimed at understanding the distinct set of evolutionary forces acting on sex chromosomes relative to autosomes (Bachtrog et al. 2009; Pala, Hasselquist, et al. 2012; Bachtrog 2013; Šíchová et al. 2013). Furthermore, when the relevant taxa also happen to be tractable genetic model systems, there is opportunity to explore the functional and mechanistic changes associated with sex chromosome evolution. The congruence of neo-sex chromosomes existing in a model system is relatively rare, although there are some notable examples.

Numerous independent origins of neo-sex chromosomes are known in *Drosophila* fruit flies, where recent work has revealed much about the evolutionary and functional dynamics of these unusual sequences (Bachtrog et al. 2009; Counterman et al. 2004; Flores et al. 2008; Zhou et al. 2013; Emily J Brown & Bachtrog 2014; Nozawa et al. 2014). Substantial insights have also come from stickleback fish, where neo-sex chromosomes appear to play an important role in reproductive isolation between incipient species (Kitano et al. 2009; White et al. 2015; Yoshida et al. 2014). Looking beyond these established model systems, the rapid expansion of genomic technologies has allowed extensive analyses of gene content, sex-biased gene expression,

dosage compensation, and sequence divergence for recently evolved sex chromosomes among a very diverse set of organisms. This includes, for example, several lineages insects [Teleopsid flies, a grasshopper, and Strepsiptera (Baker & Wilkinson 2010; Mahajan & Bachtrog 2015; Palacios-Gimenez et al. 2015)], vertebrates [mammals and birds (Zhou et al. 2008; Murata et al. 2015; Pala, Hasselquist, et al. 2012) , and plants [*Silene* and *Rumex* genera (Papadopulos et al. 2015; Deborah Charlesworth 2015; Hough et al. 2014)]. A clear consensus emerges from this research that the lack of recombination associated with sex chromosomes catalyzes a cascade of evolutionary changes involving the degeneration of one allosome, the accumulation of genes with sex-biased expression, increased evolutionary rates, and (often, but not always) the acquisition of dosage compensation. Yet many of the details in this process remain elusive and unresolved, including the rate of allosome divergence, the role of positive selection versus drift, the importance sex-specific selection, and the mechanisms underlying dosage compensation (or the reasons for its absence). It is therefore important to continually identify new opportunities for novel insight into the evolution of sex chromosomes.

Overwhelmingly, research on sex chromosomes occurs in male-heterogametic (XY) species (Ellegren 2011; Parsch & Ellegren 2013; Bachtrog 2013; Vicoso & Brian Charlesworth 2006). This appears to be particularly true for neo-sex chromosomes, where contemporary genomic analyses of neo-Z or neo-W chromosomes are currently lacking [with one notable exception for birds (Pala, Hasselquist, et al. 2012)]. This imbalance is unfortunate, as ZW sex determination replaces male-specific selection with female-specific selection during the evolution of heterogamety, offering a novel framework for elucidating sex chromosome evolution. What prospects are there for improving this situation? Birds are the most prominent vertebrate taxon that is female-heterogametic, but it appears that avian neo-sex chromosomes are quite rare, and absent from prominent model species (*e.g.,* chicken, zebra finch) (Pala, Naurin, et al. 2012; Nanda et al. 2008). Fishes and squamates appear to be far more labile in sex-chromosome constitution, with numerous independent transitions between male and female-heterogamety and relatively frequent sex-autosome fusions (Pennell et al. 2015), thus there are potentially great opportunities in these taxa. However, no obviously tractable ZW model system with neo-sex chromosomes is yet apparent for these lineages.

For many reasons, Lepidoptera (moths and butterflies) may be the most promising female-heterogametic taxon for studying neo-sex chromosomes. Synteny is unusually well-conserved in Lepidoptera (Ahola et al. 2014; Heliconius Genome Consortium 2012; Pringle et al. 2007), yet there are also numerous known examples of independently evolved neo-Z and neo-W chromosomes, several of which have been well-characterized cytogenetically (Nguyen et al. 2013; Šíchová et al. 2013; Yoshido et al. 2010; Traut et al. 2007). Furthermore, comparative genomic resources in this insect order are substantial and growing quickly (www.lepbase.org).

In this context, we report the fortuitous discovery of a neo-Z chromosome in the monarch butterfly, *Danaus plexippus*, and closely related species. Monarch butterflies, renowned for their annual migration across North America, already have a strong precedent as a model system in ecology (Oberhauser & Solensky 2004). Recently monarchs have emerged as a model system for genome biology, with a well-assembled reference genome, extensive population resequencing data, and a precedent for genome engineering (Zhan et al. 2011; Merlin et al. 2013; Zhan et al. 2014). The discovery of a neo-Z chromosome substantially enriches the value of this species as a research model in genome biology and lays the

foundation for extensive future insights into the evolution and functional diversity of sex chromosomes.

Results

Identifying Z-linked scaffolds in *D. plexippus*

We identified Z-linked scaffolds in the *D. plexippus* genome assembly (Zhan & Reppert 2012; Zhan et al. 2011) by comparing sequencing coverage from male and female samples. Z chromosome DNA content in males should be twice that in females, while autosomes should have equal DNA content between sexes. Thus a corresponding two-fold difference in sequencing coverage is expected between sexes for the Z chromosome, but not autosomes, and can be used to identify Z-linked scaffolds (Mahajan & Bachtrog 2015; Vicoso et al. 2013; Martin et al. 2013). A histogram of male:female ratios of median coverage clearly identifies two clusters of scaffolds (Fig. 1). One large cluster is centered around equal coverage between sexes (Log2 M:F = 0) and a second, smaller cluster is centered around two-fold greater coverage in males (Log2 M:F =1). We can thus clearly distinguish the Z-linked scaffolds as those with Log2(M:F) > 0.5, with the remainder of the scaffolds presumed to be autosomal.

One scaffold, DPSCF300028, appeared to have an intermediate coverage ratio, falling at Log2 M:F ≈ 0.7. One likely explanation for such intermediate values is that the scaffold is a chimera of Z-linked and autosomal sequence arising from an error in genome assembly (Martin et al. 2013). In this scenario, only a portion of the scaffold is Z-linked and gives a two-fold difference in coverage between sexes; the remaining autosomal fraction of the scaffold yields equal coverages. The resulting estimate of average coverage for the entire scaffold then falls at an intermediate value between expectations for Z or autosomal scaffolds. This is clearly true for DPSCF300028, as revealed by examining basepair-level sequencing coverage across the scaffold (Fig. 2A). While average male coverage is consistent across the entire length of the scaffold, female coverage exhibits a clear transition between coverage equal to males (the autosomal portion) and coverage one half that of males (the Z-linked portion). Indeed, there are two such transitions in scaffold DPSCF300028, which we estimate to occur at 0.76 Mbp and 1.805 Mbp, creating a “sandwich” of one Z segment flanked by autosomal segments.

Ortholog counts link scaffolds to chromosomes.

The Lepidoptera show a very high level of conserved synteny across substantial evolutionary divergences (Ahola et al. 2014; Heliconius Genome Consortium 2012; Pringle et al. 2007). This means it is possible to use counts of orthologous genes to assign *D. plexippus* scaffolds to linkage groups (i.e. chromosomes) delineated in other moth or butterfly species. We generated predicted orthologs between *D. plexippus* and three other reference species where genetic linkage mapping has been used to assign genomic scaffolds to chromosomes: *Melitaea cinxia* (N=31), *Heliconius melpomene* (N=21), and *Bombyx mori* (N=28) (Heliconius Genome Consortium 2012; Ahola et al. 2014; International Silkworm Genome Consortium

2008). *M. cinxia* and *H. melpomene* are both nymphalid butterflies equally diverged from *D. plexippus*, while the silkmoth, *B. mori*, is distinctly more divergent (Wahlberg et al. 2009; Kawahara & Breinholt 2014).

To assign *D. plexippus* scaffolds to chromosome, we tabulated per scaffold the counts of one-to-one reference species orthologs per reference species chromosome. *D. plexippus* scaffolds were then assigned to the reference chromosome with the maximum count of orthologs. For a few scaffolds, a tie occurred in maximum ortholog count per reference chromosome, in which case the scaffold was removed from further analysis; at most this occurred for only 14 scaffolds per reference species and usually involved small scaffolds harboring fewer than 5 orthologs. Typically this method yielded a clear "best" reference chromosomal assignment for each *D. plexippus* scaffold.

This method of ortholog-count chromosomal "lift-over" resulted in putative chromosomal assignments for >90% of *D. plexippus* genes relative to each reference species (Table 1, Supplementary Table S2). Also, at least 4500 orthologous genes were co-localized to chromosome between *D. plexippus* and each reference species. Having several thousand orthologs mapped to chromosome in *D. plexippus* and a reference species presents the opportunity to examine the extent of chromosomal rearrangements and gene movement between the two species. Here we primarily report the comparison with *M. cinxia* because this species is believed to retain the ancestral lepidopteran karyotype of 31 chromosomes (Ahola et al. 2014). Furthermore, this count of chromosomes is closest to that reported for *Danaus* butterflies (N=30), indicating it is likely the most similar karyotype to *D. plexippus* (Keith S Brown Jr et al. 2004). *H. melpomene* and *B. mori* are known to have more derived karytoypes involving several chromosomal fusions relative to *M. cinxia*; details of comparisons to these two species are reported in the supplementary content and provide comparable support for the primary findings reported here.

Figure 3 summarizes the cross-tabulation of chromosomal linkage for >4500 orthologs between *M. cinxia* and *D. plexippus*. The overwhelming majority of orthologs fall on the diagonal, indicating substantial conservation of chromosomal linkage and relatively little gene shuffling, as has been reported elsewhere for Lepidoptera (Ahola et al. 2014; Heliconius Genome Consortium 2012; Pringle et al. 2007). The two most notable exceptions to this pattern both involve the Z chromosome (Chr1). In one case [McChr9, DpChr1] we could anticipate this because of the previously identified chimeric scaffold, DPSCF300028. This scaffold harbors 34 orthologs assigned to McChr1 and 23 orthologs assigned to McChr9, consistent with the chimeric nature of the chromosome revealed from male:female coverage ratios (Fig 2A).

The second case [McChr1, DpChr21] appeared to arise entirely from a single scaffold, DPSCF300001, the largest scaffold in the *D. plexippus* v3 assembly. This scaffold carried 107 orthologs assigned to McChr21, 28 orthologs assigned to McChr1, 13 orthologs assigned to McChr 23, and a few other orthologs assigned to other autosomes. Notably, despite the large number of apparently autosomal orthologs, the average male:female coverage ratio for DPSCF300001 was consistent with it being Z-linked [Log2(M:F coverage) = 0.92]. Nonetheless, we plotted coverage across the chromosome and detected a ~1 Mbp portion at the 3' end of the scaffold with coverage patterns consistent with being an autosome (Fig 2C). The *M. cinxia* orthologs in this autosomal portion were linked exclusively to McChr23. There was not an obvious shift in sequencing coverage between sexes to indicate a misassembled Z-autosome

chimera involving McChr21. Rather, it appeared that nearly the entirety of scaffold DPSCF300001 had twice the coverage in males than in females, consistent with the entire scaffold being Z-linked, both for regions apparently homologous to Mc1(Z) and McChr21.

A neo-Z chromosome in *D. plexippus*

The observation that a substantial portion of scaffold DPSCF300001 was Z-linked and homologous to McChr21, while another large section of the same scaffold was homologous to McChr1 (i.e., McChrZ) suggested the hypothesis that a single Z-autosome fusion could explain the karyotypic differences between *D. plexippus* (N=30) and *M. cinxia* (N=31). To further investigate this hypothesis of a major evolutionary transition in sex chromosome composition in the *Danaus* lineage, we examined the chromosomal assignments for all scaffolds identified as Z-linked via sequencing coverage ratios (Z-cov scaffolds). Specifically, we identified the unique set of reference chromosomes to which Z-cov scaffolds were assigned, and then examined the male:female coverage ratio for all scaffolds assigned to those reference chromosomes. In the case of *M. cinxia*, all Z-cov scaffolds were assigned either to McChr1 or McChr21 (Fig. 4; comparable results were obtained for *H. melpomene* and *B. mori*, Supplementary figures S1-S2). This result provides further evidence that the Z in *D. plexippus* is a neo-sex chromosome reflecting the fusion of the ancestral Z chromosome with an autosome homologous to McChr21.

This analysis intersecting Z-cov scaffolds with homology to *M. cinxia* revealed two scaffolds that did not fit with the expected pattern of sequencing coverage (Fig. 4). First, scaffold DPSCAF300044 was assigned to McChr1(Z) but had Log2 M:F ≈ 0.25, much more like other autosomes than other Z-linked chromosomes. This scaffold had seven Z-linked orthologs and 4 autosomal, potentially suggesting another chimeric scaffold. Indeed, examining coverage across the scaffold revealed a clear transition in coverage as previously observed for DPSCF300001 and DPSCF300028 (Fig 2B). Thus the low male:female coverage ratio for this scaffold is the artifact of an assembly error. Again we were able to partition the scaffold into two sections, one autosomal and one Z-linked, with a breakpoint estimated at 0.29 Mbp.

DPSCF300403 was the other scaffold where the M:F ratio of median coverage was inconsistent with the hypothesis of a neo-Z chromosome. This scaffold was assigned to McChr21 but had an autosomal coverage ratio. Coverage along the chromosome was consistent with it being entirely autosomal (Supplementary Figure S3). In this case the scaffold carried only a single one-to-one orthologous gene (and only 5 protein-coding genes total), so the assignment to McChr21 is tenuous and likely inaccurate. This scaffold also had a single one-to-one ortholog found in *B. mori*, and none identified in *H. melpomene*. We therefore consider this scaffold largely uninformative concerning the presence of a neo-Z in *D. plexippus*.

The neo-Z chromosome exists in the Monarch's close relatives

The population genomic data set of Zha *et al.* (2014) also contained male and female resequencing samples for four closely related congeners: *D. gilippus, D. chrysippus, D. erippus,* and *D. eresimus*. This presents the opportunity to assess whether this neo-Z exists in these species in addition to Monarch. Published reports of an N=30 karyotype in some of these

species leads to the strong prediction that they all also carry the same neo-Z chromosomal arrangement (Keith S Brown Jr et al. 2004). As expected, male versus female sequence coverage analysis does clearly show the same scaffolds homologous to both McChr1 and McChr21 as having sequencing coverage consistent with a neo-Z (Fig. 5). Thus it appears that the origin of this neo-Z predates the diversification of the genus *Danaus*.

Annotating chromosomal linkage

The combination of sequencing coverage analysis and comparative "liftover" allowed us to provisionally assign genes to chromosomes in *D. plexippus*. Genes falling on Z-cov scaffolds, or within the portion assessed as Z-linked for noted chimeric scaffolds, have been assigned to the Z chromosome. Otherwise genes and scaffolds are assigned to chromosomes based directly on the results of the "lift-over" relative to *M. cinxia*. Table 2 gives a tabulated summary of results, while results for every scaffold and protein coding gene are provided in Supplementary Tables S2 & S3, respectively.

Discussion

This discovery of a neo-Z chromosome in *Danaus* butterflies and our discrimination of genes falling on the ancestral versus recently autosomal portions are fundamental observations that provide the foundation for a host of future inferences. These results create novel opportunities to address rates of molecular evolution, the evolution of dosage compensation, the pattern of allosome divergence, and many other important questions in sex chromosome biology, all in an emerging genetic and genomic female-heterogametic model system.

It seems evident from the results presented here that if there remains a neo-W chromosome (*i.e.,* a degraded homolog of the neo-Z segment), it must be substantially diverged from the neo-Z. We infer this from the very consistent 2:1 coverage ratio observed on scaffold regions corresponding to McChr21. If the neo-W retained substantial homology to the neo-Z, we would expect many sequencing reads emanating from the neo-W to align to the neo-Z, and shift this ratio towards equality. This evidently does not occur, strongly indicating substantial divergence between the neo-Z and any neo-W sequence that is retained. Indeed, it is not even clear at this point whether there is any neo-W chromosome at all. This is an obvious point for immediate investigation, perhaps best approached using cytogenetic techniques (Šíchová et al. 2013; Nguyen et al. 2013).

Brown *et al.* (2004) report chromosome counts from male butterflies of several species from three genera in the Danaini butterfly tribe: *Danaus* (N=30), *Anetia* (N=31), and *Lycorea* (N=30). The most recent phylogenetic study of these species reports *Anetia* within the most basally diversifying lineage in this group (Brower et al. 2010). So it is tempting to speculate that the Z-autosome fusion reported here for *Danaus* occurred within the Danaini, after the divergence from *Anetia*, which has 31 chromosomes, presumably reflecting a shared ancestral karyotype with Melitaea. However, the same phylogenetic study reports *Anetia* as sister to *Lycorea* (N = 30), within the same basally splitting lineage. Because no other chromosome counts are known for the numerous species at intermediate divergences between *Danaus* and the (*Anetia*,*Lycorea*) lineage, we are left with two plausible scenarios, assuming the reported

phylogenetic relationships are accurate. In one case, *Anetia* indeed retains the ancestral karyotype while fusions independently occurred in *Lycorea* and also in the lineage leading to *Danaus*. The alternative case is that a Z-autosome fusion predates the origins of all Danaini and that *Anetia* carries a chromosomal fission, producing one extra chromosome for N=31, a chromosome count that is convergent but not homologous to *Melitaea*. Resolving these two possibilities will likely require a comparative analysis of Z-chromosome homology within the Danaini.

In analyzing patterns of chromosomal fusion in *H. melpomene* and *B. mori* relative to *M. cinxia*, Ahola *et al.* (2014) report a significant tendency for a limited set of ancestral chromosomes to be involved in chromosomal fusion events. However, neither the ancestral Z nor McChr21 are among these frequently fused chromosomes. Thus the chromosomal fusion reported here appears to be an exception to the trends arising in these other lepidopteran lineages.

Conclusion

We have used a combination of genome sequencing coverage and comparative genomic analysis to demonstrate that *Danaus* butterflies harbor a neo-Z chromosome resulting from the fusion of the ancestral Z chromosome and an autosome homologous to Chr21 in *M. cinxia*. Our analysis also identified and resolved several Z-autosome chimeric scaffolds in the most recent assembly of the *D. plexippus* genome. This discovery and provisional assignment of chromosome linkage for >90% of *D. plexippus* genes paves the way for myriad and diverse investigations into sex chromosome evolution, which are likely to be of distinct importance given the increasing prominence of monarch butterfly as a female-heterogametic model species for functional and evolutionary genomics.

Methods

Sequencing coverage analysis

Illumina shotgun genomic DNA sequencing data for three male and three female *D. plexippus* individuals were selected for analysis from samples sequenced by Zhan *et al.* (2014) (Zhan et al. 2014). Male-female pairs were selected on the basis of approximately equal sequencing coverage. Samples were aligned to the *D. plexippus* version 3 genome assembly with bowtie2 (v2.1.0), using the “very sensitive local” alignment option (Zhan & Reppert 2012; Langmead & Salzberg 2012). The resulting alignments were parsed with the *genomecov* and *groupby* utilities in the BedTools software suite to obtain a per-base median coverage depth statistic for each scaffold (Quinlan & Hall 2010). Genomic sequencing data from other *Danaus* species, also generated by Zhan *et al.* 2014, were aligned to the same assembly using Stampy (v1.0.22) (default parameters, except for *substitutionrate=0.1*) (Lunter & Goodson 2011). Details of all sample identity, including GenBank SRA accessions, are given in Supplementary Table S1.

Coverage analyses comparing males and females were limited to scaffolds of lengths equal to or greater than the N90 scaffold (160,499 bp) (Zhan & Reppert 2012). Also, incomplete cases were excluded (i.e., scaffolds with no reads from one or more samples). For each sample, each scaffold median coverage was divided by the mean across all scaffold median coverages, thereby normalizing for differences in overall sequencing depth between samples. Samples were grouped by sex and the per-scaffold mean of normalized coverage depth was compared between sexes, formulated as the log2 of the male:female coverage ratio. Autosomal scaffolds are expected to exhibit equal coverage between sexes, yielding a log2 ratio of zero. Z linked scaffolds should have a ratio of one, due to the two-fold greater representation in males. Manipulation, analysis, and visualization of coverage data was performed using custom R scripts (R Development Core Team 2014).

For select scaffolds with intermediate median coverage ratios, we used *genomecov* to calculate per-base coverage, in order to identify potential assembly errors producing Z-Autosomal chimeric scaffolds. For each sample, coverage per base was divided by the mean of all scaffold median coverages, thus normalizing for overall sequencing depth. The normalized coverage per base was averaged (mean) within sex and visualized along the length of the scaffold by using the median of a 5kbp sliding window, shifted by 1kbp steps.

Point estimates for Z-autosomal break points in chimeric scaffolds were generated using a sliding window analysis of male:female coverage ratios. Putative break points were obtained as the maximum of the absolute difference between adjacent non-overlapping windows. A window of 150 Kbp with 10 kbp steps was used for DPSCF30001 and the 5' break point of DPSCF30028. A window of 10 kbp with 1 kbp steps was used for DPSCF30044 and in a second, localized analysis between 1.5 Mb and the 3' terminus of DPSCF30028 to localize the second, 3' break point.

Orthology-based chromosomal assignments for *D. plexippus* scaffolds.

Putative chromosomal linkage was predicted for *D. plexippus* scaffolds relative to the genome assemblies of three reference species, *M. cinxia, B. mori,* and *H. melpomene*, based on counts of orthologous genes (Heliconius Genome Consortium 2012; Ahola et al. 2014; International Silkworm Genome Consortium 2008). Orthologous proteins were predicted between *D. plexippus* and each reference species using the *Proteinortho* pipeline (Lechner et al. 2011). Using only 1-to-1 orthologs, we tabulated per *D. plexipppus* scaffold the number of genes mapped to each chromosome in the reference species. Each *D. plexippus* scaffold was tentatively assigned to the chromosome with the highest count of orthologs in the reference species. Scaffolds were excluded from analysis when maximum ortholog count was tied between two or more scaffolds, though this situation was rare and always involved scaffolds with low counts of (orthologous) genes.

Acknowledgements

This manuscript is dedicated to Chip Taylor, Ann Ryan, and the many hard-working members of

MonarchWatch.org. This research was supported by NSF-DEB 1457758 (to J.R.W.). The computing for this project was performed on the Community Cluster at the Center for Research Computing at the University of Kansas.

Works Cited

Ahola V et al. 2014. The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. Nature Communications. 5:4737. doi: 10.1038/ncomms5737.

Bachtrog D. 2013. Y-chromosome evolution:emerging insights into processesof Y-chromosome degeneration. Nat Rev Genet. 14:113–124. doi: 10.1038/nrg3366.

Bachtrog D, Jensen JD, Zhang Z. 2009. Accelerated adaptive evolution on a newly formed X chromosome. PLoS Biol. 7:e82. doi: 10.1371/journal.pbio.1000082.

Baker RH, Wilkinson GS. 2010. Comparative Genomic Hybridization (CGH) Reveals a Neo-X Chromosome and Biased Gene Movement in Stalk-Eyed Flies (Genus Teleopsis). 1–14.

Brower AVZ, Wahlberg N, Ogawa JR, Boppré M, Vane-Wright RI. 2010. Phylogenetic relationships among genera of danaine butterflies (Lepidoptera: Nymphalidae) as implied by morphology and DNA sequences. Systematics and Biodiversity. 8:75–89. doi: 10.1080/14772001003626814.

Brown EJ, Bachtrog D. 2014. The chromatin landscape of Drosophila: comparisons between species, sexes, and chromosomes. Genome Res. 24:1125–1137. doi: 10.1101/gr.172155.114.

Brown KS Jr, Schoultz Von B, Suomalainen E. 2004. Chromosome evolution in Neotropical Danainae and Ithomiinae (Lepidoptera). Hereditas. 141:216–236. doi: 10.1111/j.1601-5223.2004.01868.x.

Charlesworth D. 2015. Plant contributions to our understanding of sex chromosome evolution. New Phytol. 208:52–65. doi: 10.1111/nph.13497.

Counterman BA, Ortíz-Barrientos D, Noor MAF. 2004. Using comparative genomic data to test for fast-X evolution. Evolution. 58:656–660.

Ellegren H. 2011. Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. Nat Rev Genet. 12:157–166. doi: 10.1038/nrg2948.

Flores SV, Evans AL, McAllister BF. 2008. Independent origins of new sex-linked chromosomes in the melanica and robusta species groups of Drosophila. BMC Evol Biol. 8:33. doi: 10.1186/1471-2148-8-33.

Heliconius Genome Consortium. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature. 487:94–98. doi: 10.1038/nature11041.

Hough J, Hollister JD, Wang W, Barrett SCH, Wright SI. 2014. Genetic degeneration of old and young Y chromosomes in the flowering plant Rumex hastatulus. Proceedings of the National Academy of Sciences. 111:7713–7718. doi: 10.1073/pnas.1319227111.

International Silkworm Genome Consortium. 2008. The genome of a lepidopteran model insect, the silkworm Bombyx mori. Insect Biochem Mol Biol. 38:1036–1045. doi: 10.1016/j.ibmb.2008.11.004.

Kawahara AY, Breinholt JW. 2014. Phylogenomics provides strong evidence for relationships of butterflies and moths. Proceedings of the Royal Society B: Biological Sciences. 281:20140970–20140970. doi: 10.1126/science.1174096.

Kitano J et al. 2009. A role for a neo-sex chromosome in stickleback speciation. Nature. 461:1079–1083. doi: 10.1038/nature08441.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9:357–359. doi: 10.1038/nmeth.1923.

Lechner M et al. 2011. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. BMC Bioinformatics. 12:124. doi: 10.1186/1471-2105-12-124.

Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res. 21:936–939. doi: 10.1101/gr.111120.110.

Lynch M. 2007. *The Origins of Genome Architecture*. Sinauer Associates Incorporated: Sunderland, MA.

Mahajan S, Bachtrog D. 2015. Partial Dosage Compensation in Strepsiptera, a Sister Group of Beetles. Genome Biology and Evolution. 7:591–600. doi: 10.1093/gbe/evv008.

Martin SH et al. 2013. Genome-wide evidence for speciation with gene flow in Heliconius butterflies. Genome Res. 23:1817–1828. doi: 10.1101/gr.159426.113.

Merlin C, Beaver LE, Taylor OR, Wolfe SA, Reppert SM. 2013. Efficient targeted mutagenesis in the monarch butterfly using zinc-finger nucleases. Genome Res. 23:159–168. doi: 10.1101/gr.145599.112.

Murata C et al. 2015. Initiation of recombination suppression and PAR formation during the early stages of neo-sex chromosome differentiation in the Okinawa spiny rat, Tokudaia muenninki. BMC Evol Biol. 15:234. doi: 10.1186/s12862-015-0514-y.

Nanda I, Schlegelmilch K, Haaf T, Schartl M, Schmid M. 2008. Synteny conservation of the Z chromosome in 14 avian species (11 families) supports a role for Z dosage in avian sex determination. Cytogenet Genome Res. 122:150–156. doi: 10.1159/000163092.

Nguyen P et al. 2013. Neo-sex chromosomes and adaptive potential in tortricid pests. Proceedings of the National Academy of Sciences. 110:6931–6936. doi: 10.1073/pnas.1220372110.

Nozawa M, Fukuda N, Ikeo K, Gojobori T. 2014. Tissue- and Stage-Dependent Dosage Compensation on the Neo-X Chromosome in Drosophila pseudoobscura. Molecular Biology and Evolution. 31:614–624. doi: 10.1093/molbev/mst239.

Oberhauser KS, Solensky MJ. 2004. *Monarch Butterfly Biology & Conservation*. Cornell University Press: Ithaca, NY.

Pala I, Hasselquist D, Bensch S, Hansson B. 2012. Patterns of molecular evolution of an avian neo-sex chromosome. Molecular Biology and Evolution. 29:3741–3754. doi: 10.1093/molbev/mss177.

Pala I, Naurin S, et al. 2012. Evidence of a neo-sex chromosome in birds. Heredity. 108:264–272. doi: 10.1038/hdy.2011.70.

Palacios-Gimenez OM, Marti DA, Cabral-de-Mello DC. 2015. Neo-sex chromosomes of Ronderosia bergi: insight into the evolution of sex chromosomes in grasshoppers. Chromosoma. 124:353–365. doi: 10.1007/s00412-015-0505-1.

Papadopulos AST, Chester M, Ridout K, Filatov DA. 2015. Rapid Y degeneration and dosage compensation in plant sex chromosomes. Proceedings of the National Academy of Sciences. 112:13021–13026. doi: 10.1073/pnas.1508454112.

Parsch J, Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. Nat Rev Genet. 14:83–87. doi: 10.1038/nrg3376.

Pennell MW et al. 2015. Y fuse? Sex chromosome fusions in fishes and reptiles. PLoS Genet. 11:e1005237. doi: 10.1371/journal.pgen.1005237.

Pringle EG et al. 2007. Synteny and chromosome evolution in the lepidoptera: evidence from mapping in Heliconius melpomene. Genetics. 177:417–426. doi: 10.1534/genetics.107.073122.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841–842. doi: 10.1093/bioinformatics/btq033.

R Development Core Team. 2014. R: A Language and Environment for Statistical Computing. http://www.R-project.org/.

Soltis P, Soltis DE. 2012. *Polyploidy and Genome Evolution*. Springer Science & Business Media: New York, NY.

Šíchová J, Nguyen P, Dalíková M, Marec F. 2013. Chromosomal Evolution in Tortricid Moths: Conserved Karyotypes with Diverged Features Doucet, D, editor. PLoS ONE. 8:e64520. doi: 10.1371/journal.pone.0064520.t001.

Traut W, Sahara K, Marec F. 2007. Sex chromosomes and sex determination in Lepidoptera. Sex Dev. 1:332–346. doi: 10.1159/000111765.

Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet. 7:645–653. doi: 10.1038/nrg1914.

Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D. 2013. Comparative Sex Chromosome Genomics in Snakes: Differentiation, Evolutionary Strata, and Lack of Global Dosage Compensation Barton, NH, editor. PLoS Biol. 11:e1001643. doi: 10.1371/journal.pbio.1001643.s024.

Wahlberg N et al. 2009. Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary. Proceedings of the Royal Society B: Biological Sciences. 276:4295–4302. doi: 10.1098/rspb.2009.1303.

White MA, Kitano J, Peichel CL. 2015. Purifying Selection Maintains Dosage-Sensitive Genes during Degeneration of the Threespine Stickleback Y Chromosome. Molecular Biology and Evolution. 32:1981–1995. doi: 10.1093/molbev/msv078.

Yoshida K et al. 2014. Sex Chromosome Turnover Contributes to Genomic Divergence between Incipient Stickleback Species Zhang, J, editor. PLoS Genet. 10:e1004223. doi: 10.1371/journal.pgen.1004223.s007.

Yoshido A, Sahara K, Marec F, Matsuda Y. 2010. Step-by-step evolution of neo-sex chromosomes in geographical populations of wild silkmoths, Samia cynthia ssp. Heredity. 106:614–624. doi: 10.1038/hdy.2010.94.

Zhan S et al. 2014. The genetics of monarch butterfly migration and warning colouration. Nature. 514:317–321. doi: 10.1038/nature13812.

Zhan S, Merlin C, Boore JL, Reppert SM. 2011. The Monarch Butterfly Genome Yields Insights into Long-Distance Migration. Cell. 147:1171–1185. doi: 10.1016/j.cell.2011.09.052.

Zhan S, Reppert SM. 2012. MonarchBase: the monarch butterfly genome database. Nucleic Acids Research. 41:D758–D763. doi: 10.1093/nar/gks1057.

Zhou Q et al. 2008. Neo-sex chromosomes in the black muntjac recapitulate incipient evolution of mammalian sex

chromosomes. Genome Biol. 9:R98. doi: 10.1186/gb-2008-9-6-r98.

Zhou Q et al. 2013. The Epigenome of Evolving Drosophila Neo-Sex Chromosomes: Dosage Compensation and Heterochromatin Formation Becker, PB, editor. PLoS Biol. 11:e1001711. doi: 10.1371/journal.pbio.1001711.s019.

Table 1. Summary of assigning *D. plexippus* genes and scaffolds to chromosomes via orthology "liftover" relative to *M. cinxia*.

1:1 orthologs identified	6,740
1:1 orthologs assigned to *M. cinxia* chromosome	4,607 (68.4%)
Protein coding genes in *D. plexippus*	15,130
D. plexippus protein coding genes assigned to chromosome	14,129 (93.4%)
D. plexippus scaffolds putatively assigned to chromosomes	454

Table 2. Summary of provisional chromosomal linkage for *D. plexippus* protein coding genes, with chromosomal identity reflecting homology to *M. cinxia*

Chromosome	Number of Genes
1	1101
2	704
3	758
4	582
5	494
6	689
7	483
8	535
9	647
10	452
11	574
12	576
13	501
14	429
15	493
16	524
17	604
18	561
19	414
20	399
22	302
23	318
24	185
25	329
26	274
27	250
28	284
29	294
30	151
31	168
Not Assigned	1055

Figure 1. Distribution of median normalized male:female genomic sequencing coverage ratios for *D. plexippus* version 3 assembly scaffolds. Only scaffolds of length equal to or greater than the N90 scaffold are shown. The dotted line at 0.5 represents the value used to partition scaffolds as autosomal (grey) or Z-linked (red).

Figure 2. Normalized male and female coverage along the length of chimeric scaffolds, for (A) DPSCF300028, (B) DPSCF300044, and (C) DPSCF300001. Coverages are plotted as sliding windows (width = 5Kbp, step = 1 Kbp) of median basepair values. The associated male:female ratio of coverage for each window is plotted as a red line below the pair of sex-specific plots. Asterisks indicate the estimated break point between Z linked and autosomal segments of each scaffold, as determined by the maximum difference in adjacent, non-overlapping windows of male:female ratio (see methods for details).

Figure 3. Chromosomal co-linkage between *D. plexippus* and *M. cinxia* for predicted orthologous proteins.

Figure 4. Ratios of male:female median normalized genomic sequencing coverage plotted by scaffold length. Scaffolds homologous via "liftover" procedure to *M. cinxia* chromosomes 1/Z (blue) and 21 (green) are plotted in distinct colors. Dotted lines indicate expected values for Z-linked (red) and autosomal (black) scaffolds.

Figure 5. Ratios of male:female median normalized genomic sequencing coverage plotted by scaffold length for four species of *Danaus* butterflies.

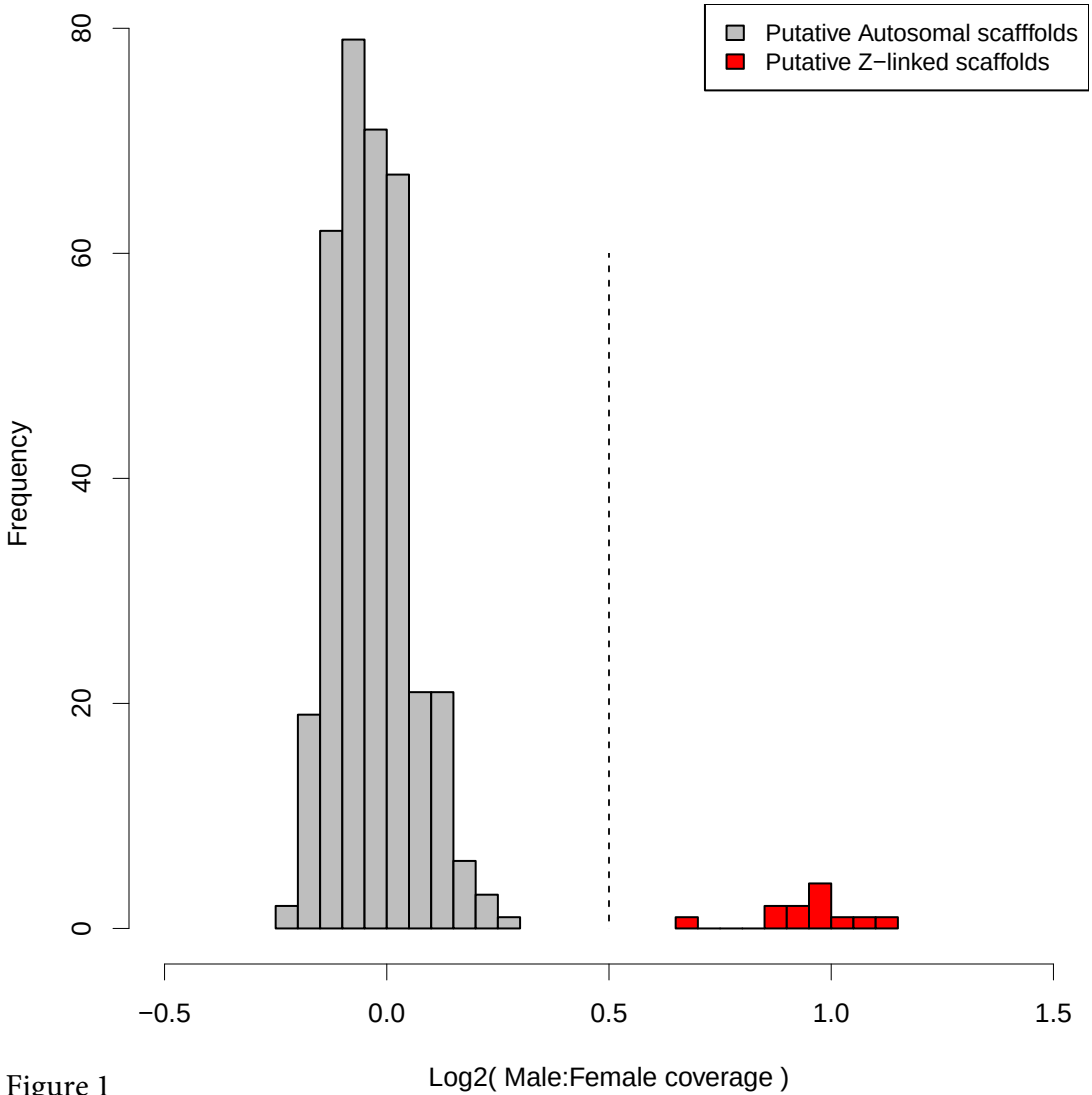

Putative Autosomal scafffolds
Putative Z−linked scaffolds
Frequency
0
20
40
60
80
−0.5
0.0
0.5
1.0
1.5
Log2(Male:Female coverage)

Figure 1

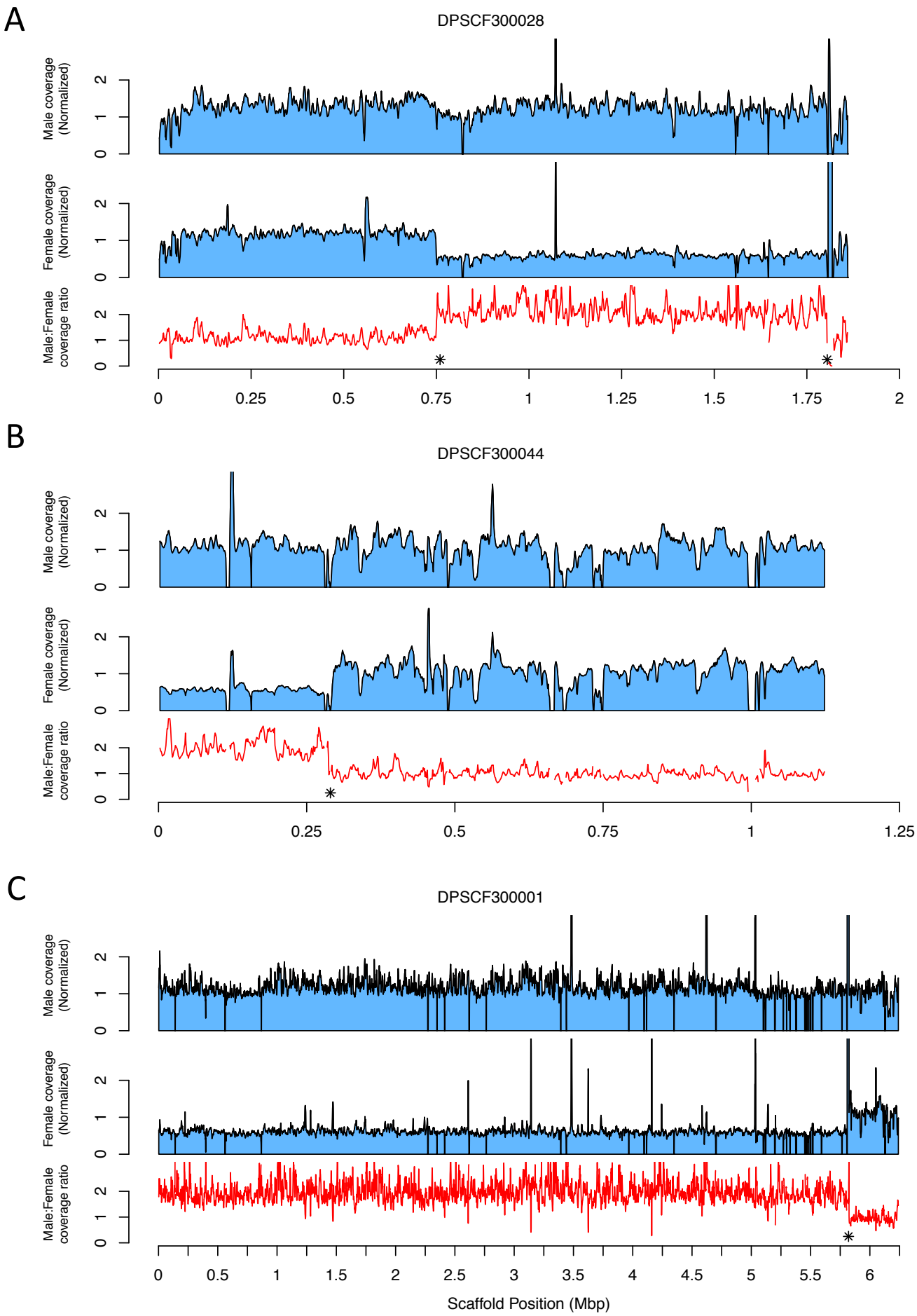

A
DPSCF300028
Male coverage (Normalized)
Female coverage (Normalized)
Male:Female coverage ratio
0
0.25
0.5
0.75
1
1.25
1.5
1.75
2
B
DPSCF300044
Male coverage (Normalized)
Female coverage (Normalized)
Male:Female coverage ratio
0
0.25
0.5
0.75
1
1.25
C
DPSCF300001
Male coverage (Normalized)
Female coverage (Normalized)
Male:Female coverage ratio
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
Scaffold Position (Mbp)

Figure 2

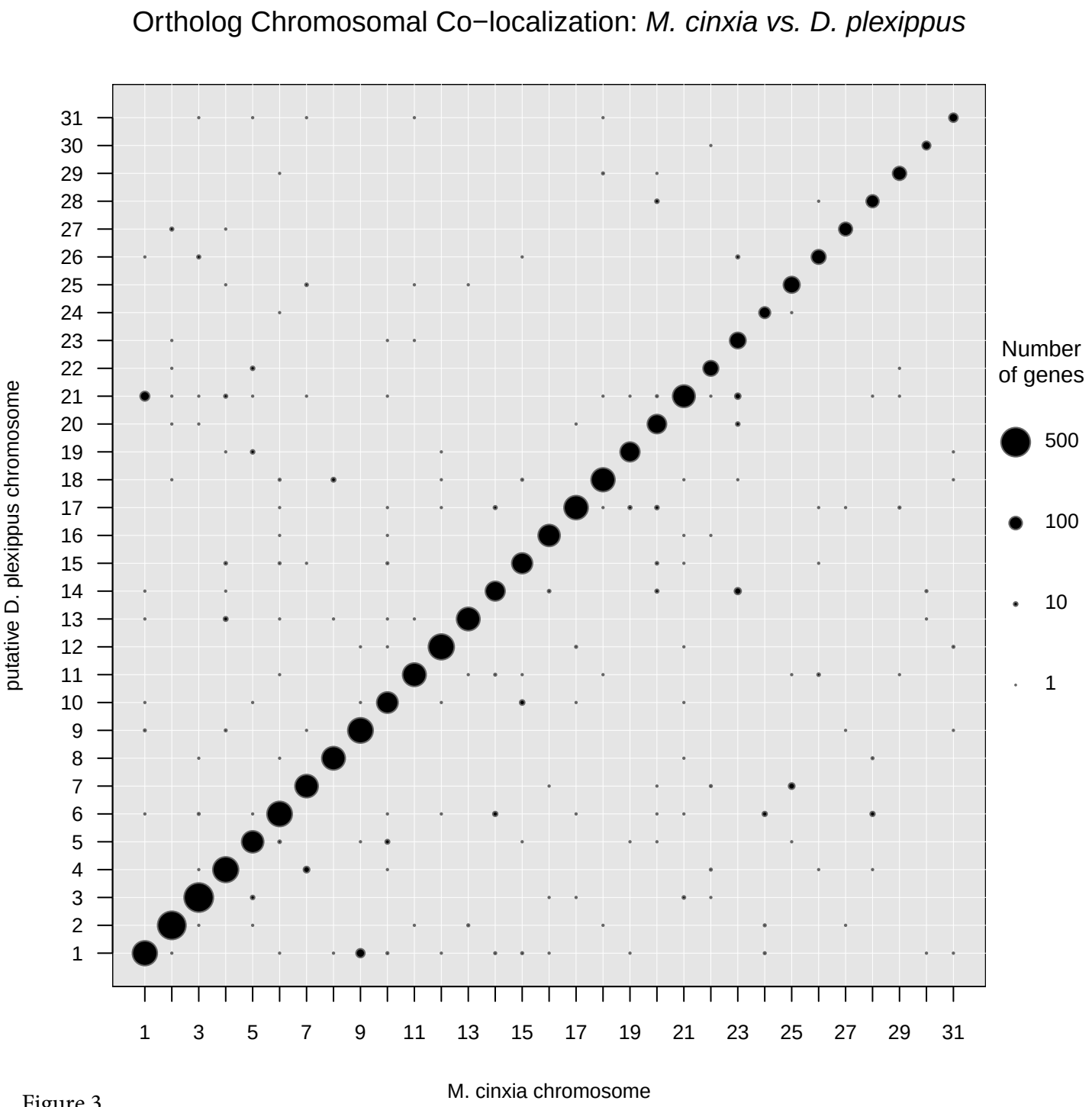

Ortholog Chromosomal Co−localization: *M. cinxia vs. D. plexippus*
putative D. plexippus chromosome
M. cinxia chromosome
Number of genes
500
100
10
1

Figure 3

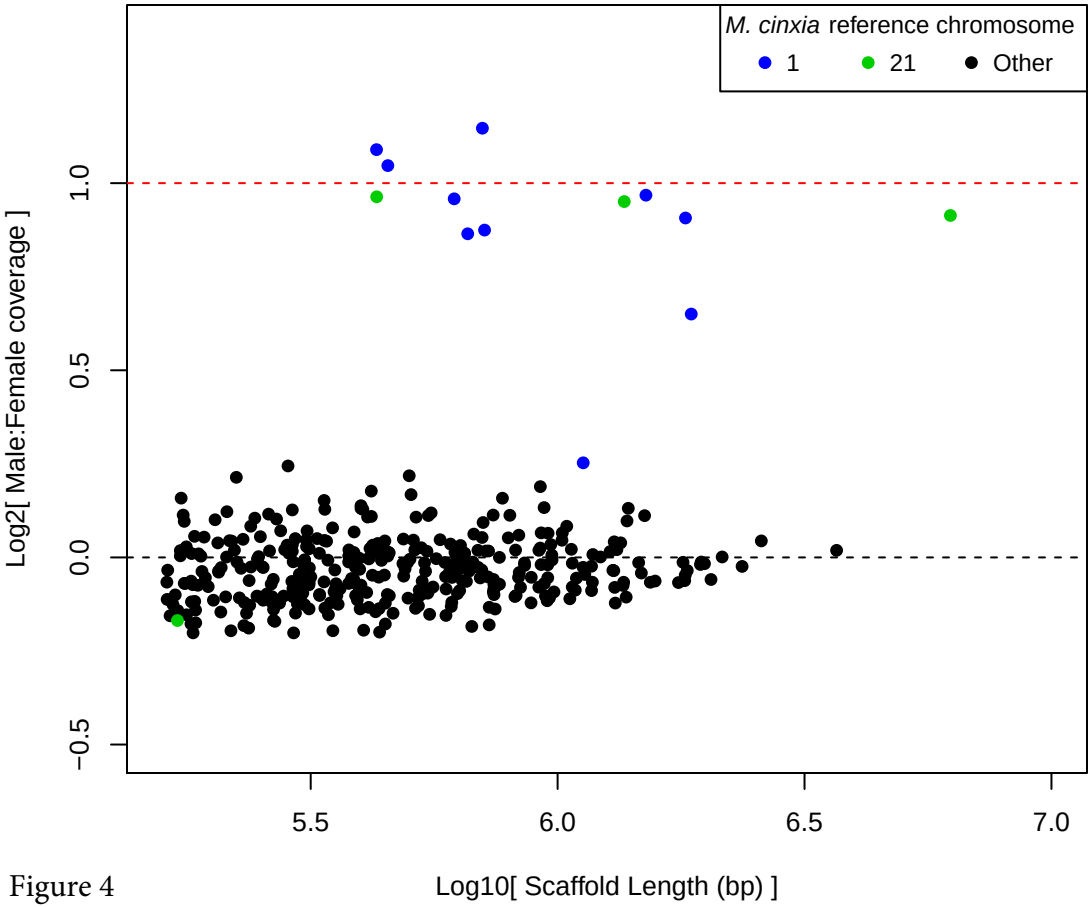

M. cinxia reference chromosome
1
21
Other
Log2[Male:Female coverage]
1.0
0.5
0.0
−0.5
5.5
6.0
6.5
7.0
Log10[Scaffold Length (bp)]

Figure 4

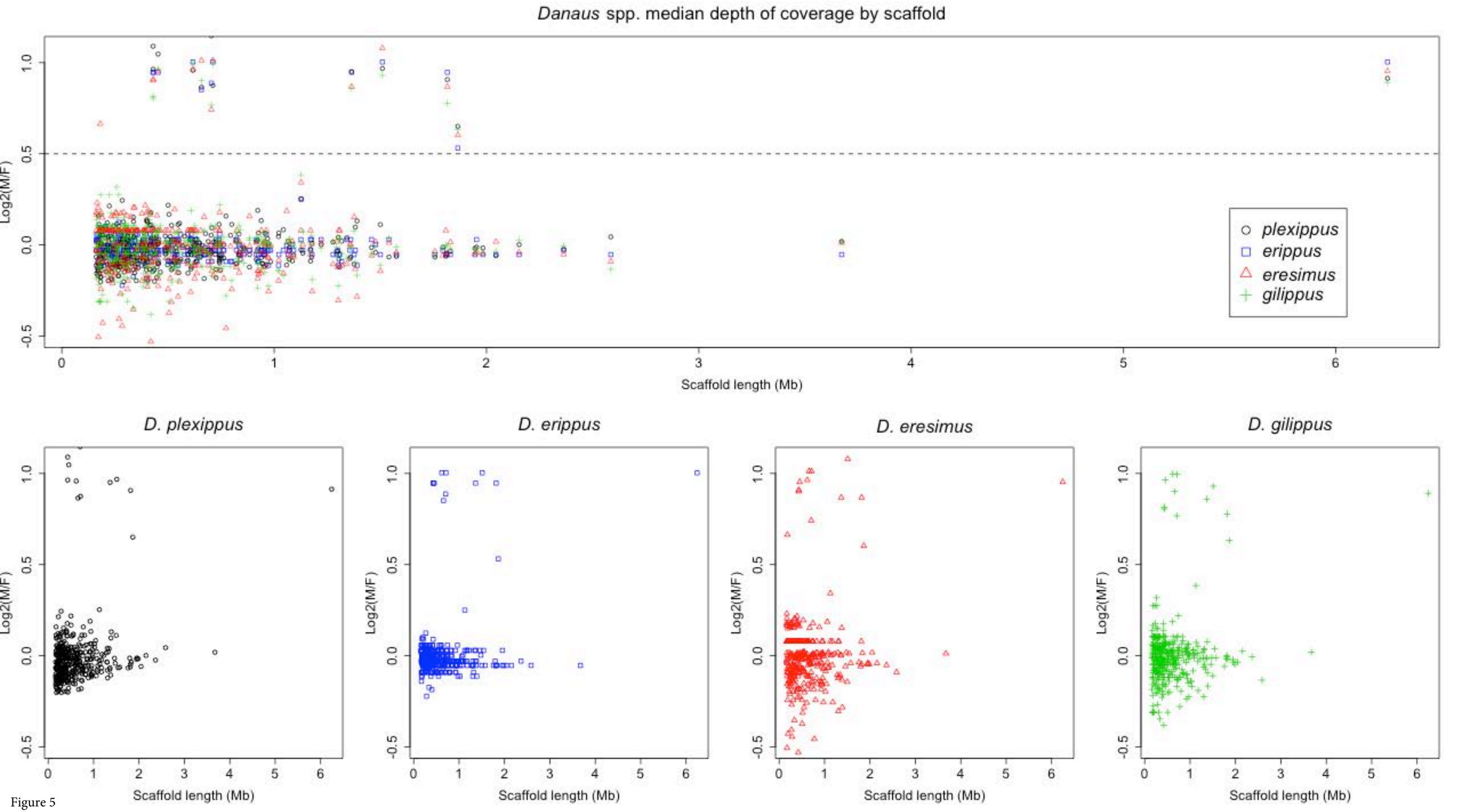
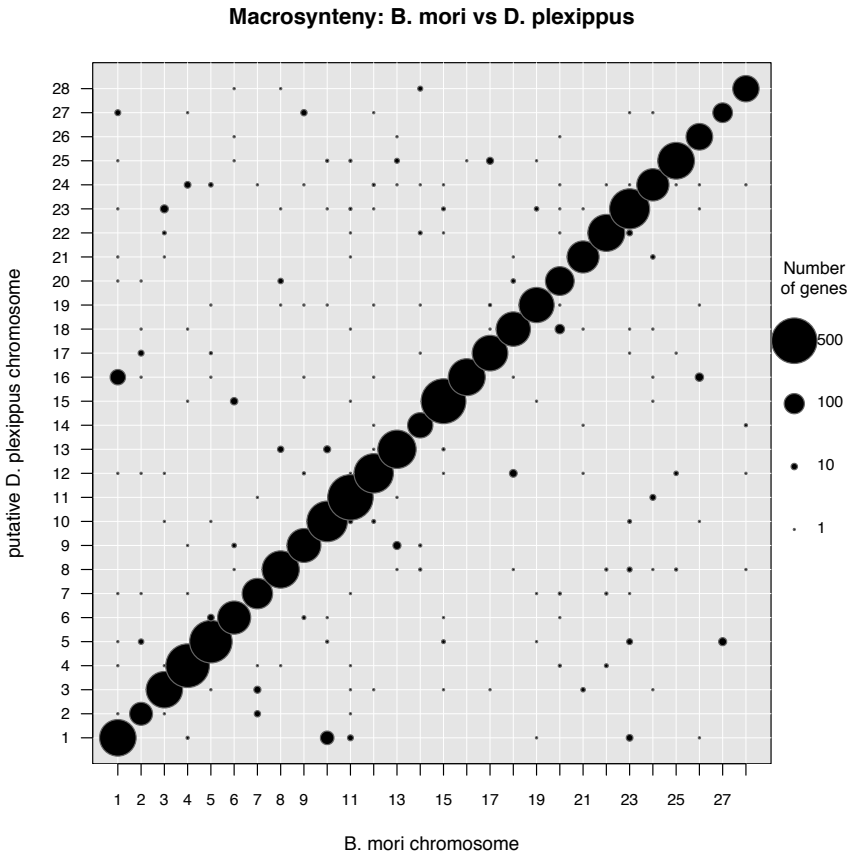

Danaus spp. median depth of coverage by scaffold
Log2(M/F)
Scaffold length (Mb)
plexippus
erippus
eresimus
gilippus
D. plexippus
D. erippus
D. eresimus
D. gilippus

Figure 5



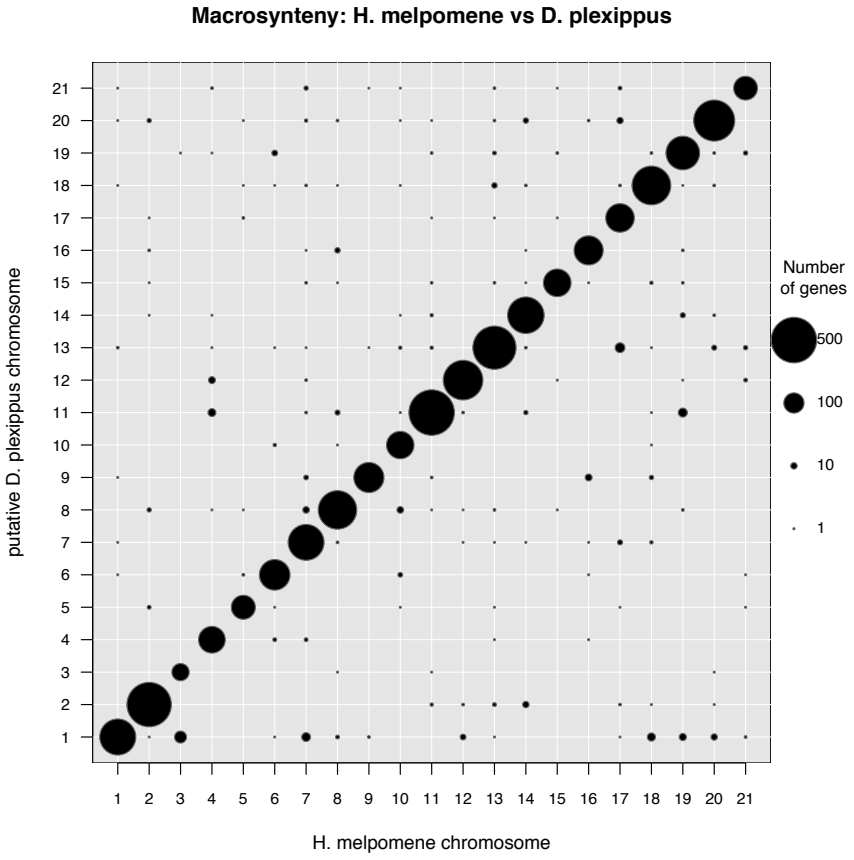
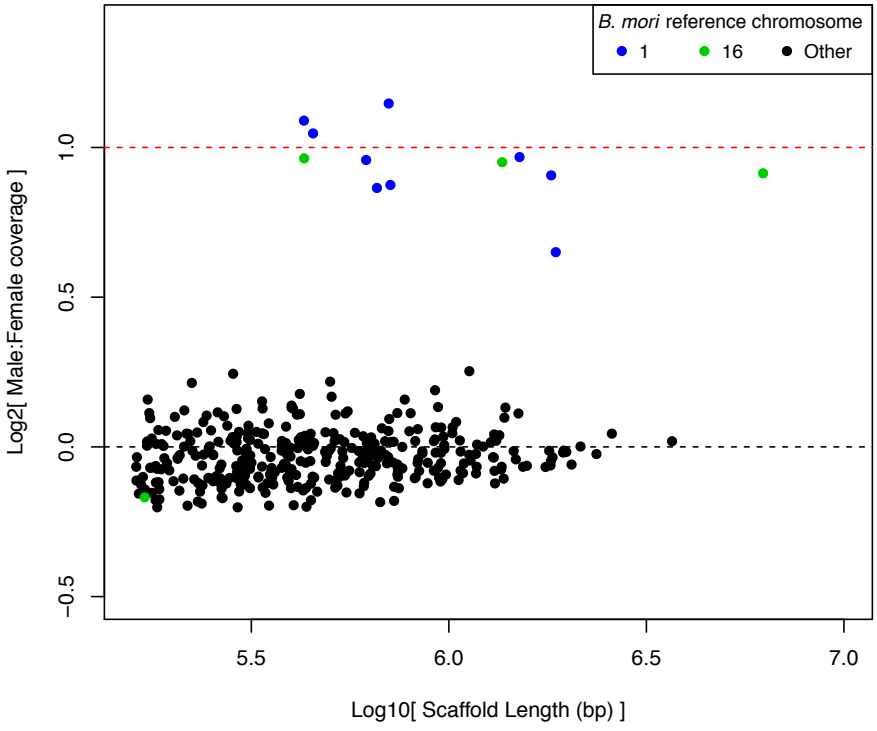


Figure S2. Chromosomal co-linkage between *D. plexippus* and *B. mori* (top) or *H. melpomene* (bottom) for predicted orthologous proteins.



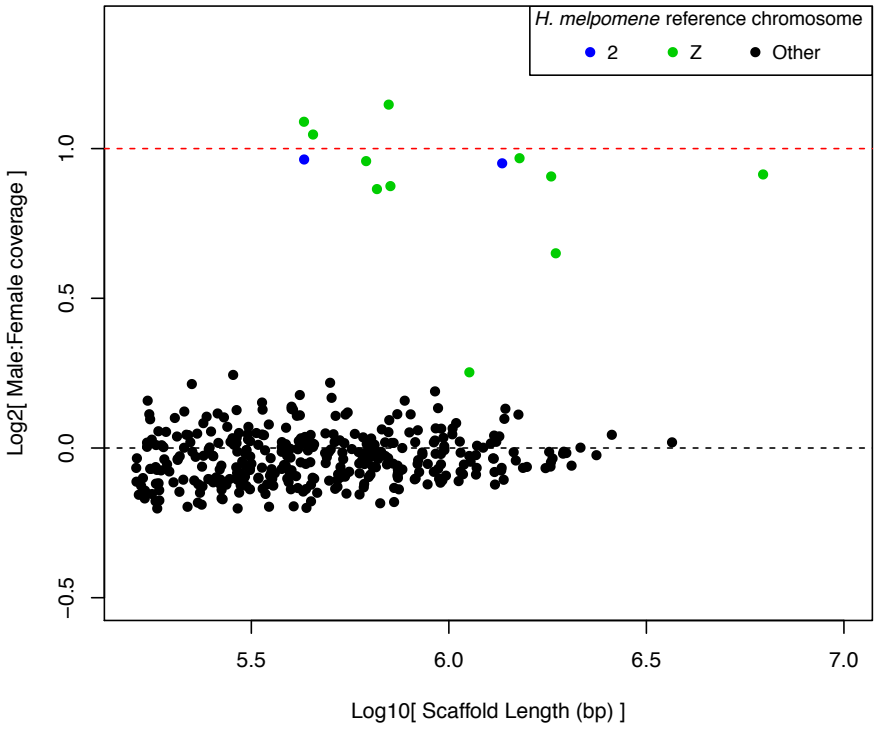


Figure S2. Ratios of male:female median normalized genomic sequencing coverage plotted by scaffold length. Scaffolds assigned to chromosomes putative homologous to the neo-Z chromosome in *D. plexippus* are plotted in distinct colors. Top, relative to *B. mori*, to chromosomes 1 (*i.e.,* Z ; blue) and 16 (green). Bottom, relative to *H. melpomene*, chromosomes 1 (*i.e.,* Z ; green) and 2 (blue).

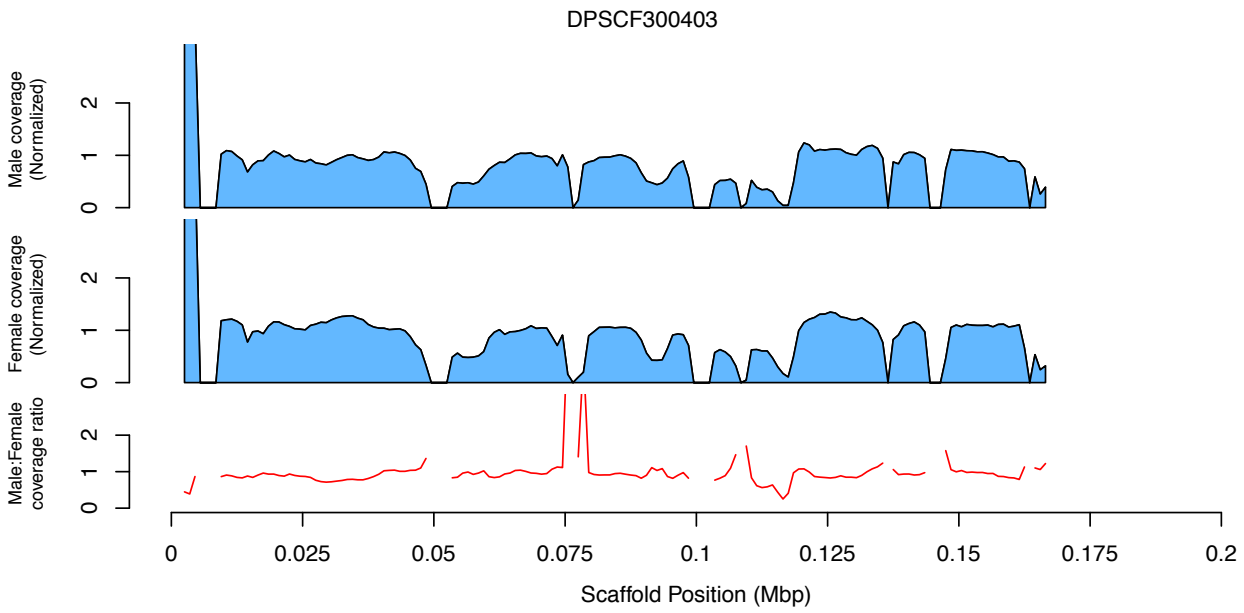


Figure S3. Normalized male and female coverage along the length DPSCF300403. Coverages are plotted as sliding windows (width = 5Kbp, step = 1 Kbp) of median basepair values. The associated male:female ratio of coverage for each window is plotted as a red line below the pair of sex-specific plots.

Supplementary Table S1. Sample identification details for sequencing data used in coverate analyses.

Region	Sample	Species	Sex	Collecting location	Date	Accession
North America	Plex_MA_HI004_M	*plexippus*	male	Massachusetts, USA	July 3 2008	SRX679269
	Plex_MA_HI035_F	*plexippus*	female	Massachusetts, USA	August 10 2009	SRX679310
	Plex_FLn_StM123_F	*plexippus*	female	St. Marks, Florida, USA	October 2008	SRX680105
	Plex_FLn_StM146_M	*plexippus*	male	St. Marks, Florida, USA	October 2009	SRX681753
	Plex_WSM_M36_M	*plexippus*	male	Samoa	June 2007	SRX680118
	Plex_WSM_M38_F	*plexippus*	female	Samoa	June 2007	SRX681528
Other *Danaus*	Erip_BRA_16005_F	*erippus*	female	Brazil	September 2010	SRX682069
	Erip_BRA_16008_M	*erippus*	male	Brazil	September 2010	SRX682070
	Eres_CRC_92_F	*eresimus*	female	Costa Rica	July 24 2010	SRX682071
	Eres_FL_27_M	*eresimus*	male	Florida, USA	July 20 2009	SRX682072
	Gili_CRC_30_M	*gilippus*	male	Costa Rica	July 24 2010	SRX682073
	Gili_TX_01_F	*gilippus*	female	Texas, USA	October 2010	SRX998564

Table S2. Summary of assigning *D. plexippus* genes and scaffolds to chromosomes via orthology "liftover" relative to three different reference assemblies.

M. cinxia	*H. melpomene*	*B. mori*	
15130	15130	15130	total number of protein coding genes in target
14129	14427	14566	number of target genes assigned to chromosome
0.934	0.954	0.963	fraction of target genes assigned to chromosome
454	514	508	number of target scaffolds assigned to chromosomes
6740	8190	7928	number of 1:1 orthologs identified
4607	7150	7534	number of 1:1 orthologs assigned to reference chromosome
0.684	0.873	0.95	fraction of 1:1 orthologs assigned to reference chromosome