Title: Genomic analyses in *Drosophila* do not support the classic allopatric model of speciation.

Authors: Leeban H. Yusuf[1], Dominik R. Laetsch[2], Konrad Lohse[‡,2], Michael G. Ritchie[‡,1]

Affiliations:[1]Centre for Biological Diversity, School of Biology, University of St Andrews, St Andrews KY16 9TH, Scotland. [2]Institute of Ecology and Evolution, University of Edinburgh, Edinburgh, Scotland.

Corresponding author: ly36@st-andrews.ac.uk

[‡] These authors contributed equally.

Abstract: The allopatric model of speciation has dominated our understanding of speciation biology and biogeography since the Modern Synthesis. It is uncontroversial because the cessation of gene flow during allopatry will allow reproductive isolation to readily emerge as a by-product of evolutionary divergence. However, whether allopatric speciation is common has rarely been systematically tested across a continuum of closely-related species. Here, we fit a range of demographic models of evolutionary divergence to whole-genome sequence data from 93 pairs of *Drosophila* species to infer speciation histories and levels of post-divergence gene flow. We find that speciation with gene flow is common, even between currently allopatric pairs of species. Estimates of historical gene flow are not predicted by current range overlap, nor is there convincing support that secondary contact commonly leads to the cessation of gene flow as predicted by models of reinforcement. Instead, our analyses suggest that most speciation processes involve some long-term gene flow perhaps due to repeated cycles of allopatry and contact, without requiring an extensive allopatric phase.

Main text: The idea that speciation processes can be classified into distinct geographic modes (allopatric, parapatric and sympatric) emerged during the *Modern Synthesis* and continues to be a central theme of evolutionary research (*1–3*). Allopatric speciation is thought to be the most common and is often considered a null model of speciation (*2*, *4*). Allopatric speciation appears uncontroversial both empirically given the widespread prevalence of "geographic" varieties and species (which influenced Darwin and Mayr), and conceptually due to the ease with which cessation of gene flow during allopatry facilitates the build-up of reproductive isolation due to genetic drift, natural or sexual selection. Much of the debate over other modes of speciation revolves around the appearance of different forms of reproductive isolation in the face of potentially homogenising gene flow. If and when species come into secondary contact after an allopatric phase, natural selection to reduce competition (but see (*5*) or to prevent maladaptive hybridization may lead to complete isolation or the formation of stable hybrid zones (*6*).

Comparisons between currently sympatric and allopatric species pairs have provided axiomatic evidence for reinforcement. For example, analyses that contrast pre and postzygotic estimates of reproductive isolation between currently allopatric and sympatric species pairs of *Drosophila* demonstrated convincingly that mate discrimination evolves more rapidly in sympatry (*7*, *8*). Alongside other comparative surveys (*9–13*), these analyses were interpreted as evidence that incompatibilities which evolved during an allopatric phase lead to reinforcement after secondary contact (*14*). However, classifying closely-related species based on their current geographic ranges assumes that these in part reflect the geographical context during the key phases of their speciation history (*15–17*). It has been acknowledged that allopatry and sympatry are ends of a spectrum (*18*, *19*), and an increasing number of genomic studies suggest that historical and contemporary gene flow is more prevalent than the architects of the modern synthesis envisaged (*20–24*). Additionally, gene flow may diminish gradually without the need for extended periods of allopatry during divergence (*4*). It is, therefore, unclear whether most speciation processes really require an allopatric phase, and how common allopatric speciation versus speciation with gene flow is.

Here we take advantage of the availability of genomic data in the best studied group of species, *Drosophila.* Work on *Drosophila* has contributed extensively to the popularity of geographic and secondary contact models of speciation, but has also motivated the development of likelihood based inference of speciation histories that involve a continuous rate of migration (*25–27*). We compare the support for, extent, and timing of gene flow, in currently allopatric and sympatric species pairs across a range of genomic divergence (~1-5%). We analyse putatively neutral, short intronic blocks across a subset of 93 species pairs originally studied by Coyne and Orr (*8*, *28*). While our choice of taxon pairs was determined by the availability of whole-genome sequence data, we note that the main result of Coyne and Orr (*8*, *28*) (faster evolution of pre-zygotic isolation for sympatric pairs) holds for this subset (Supplementary Figure 1). In fact, both premating isolation and postzygotic isolation are higher in sympatric pairs compared to allopatric pairs (two-sample Wilcoxon tests: $W = 638.5$, $p = 0.004$ and $W = 94.5$, $p = 0.009$). We fitted alternative demographic models of strict allopatric divergence versus divergence with gene flow and secondary contact (*29*, *30*) (Figure 1) to ask the following questions: (1) Is support for (historical or contemporary) gene flow different between currently sympatric or allopatric species? (2) Do currently sympatric and allopatric pairs differ in levels of genome-wide divergence? (3) Do

currently sympatric and allopatric species differ in their average divergence time or ancestral population size? (4) Finally, is there greater support for secondary contact (particularly in sympatric pairs) compared to older, historical gene flow at the onset of species divergence? We conclude that speciation involving a low rate of long-term gene flow is ubiquitous in *Drosophila*, and therefore that there is little support for the classic allopatric model of speciation.

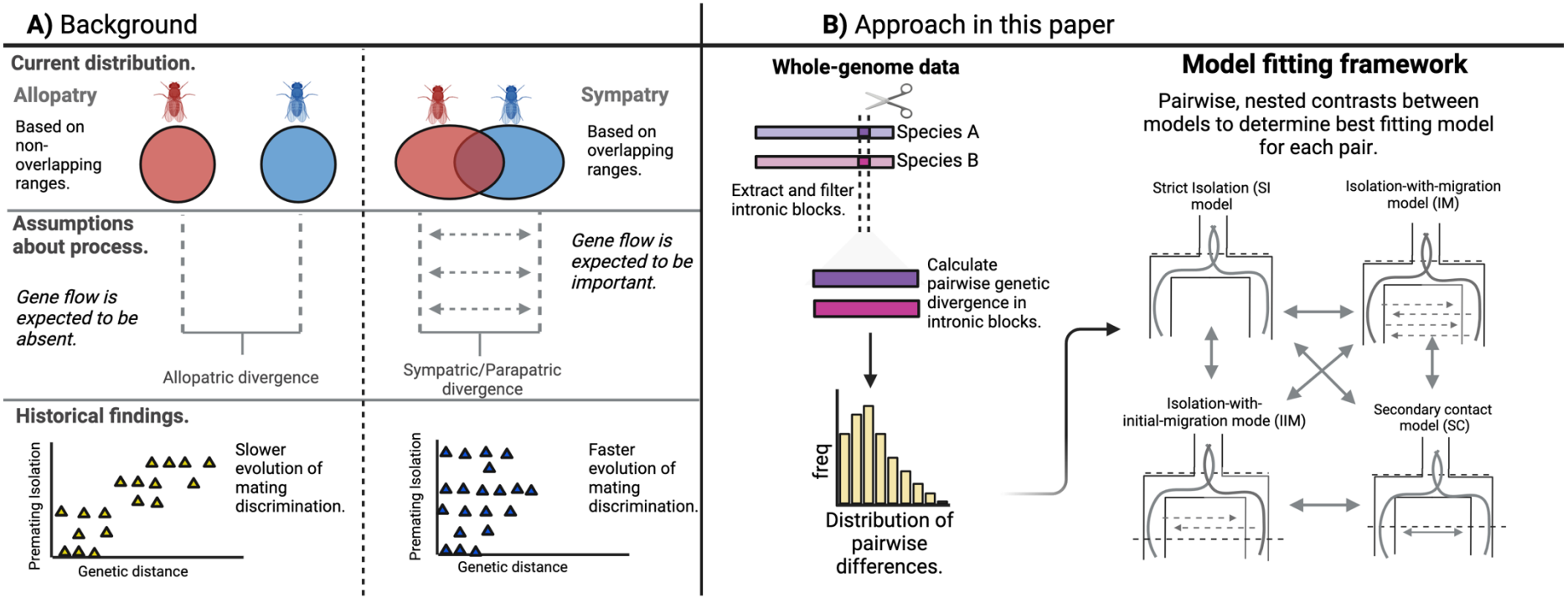


Figure 1: A) Background and **B)** our approach to understanding biogeographic modes of speciation. **A)** Previous comparative surveys on the evolution of reproductive isolation categorized species pairs by geography. This assumes that current range overlap is informative about the underlying speciation processes e.g. sympatric pairs are more likely to have exchanged migrants than allopatric taxa. **B)** Our approach uses whole-genome data to test to what extent levels of long-term gene flow differ between taxon pairs with and without current range overlap. We summarized genome-wide divergence in short, intronic blocks in terms of the distributions of pairwise differences. We fitted demographic models of speciation representing allopatric speciation and different models of speciation with gene flow, to each pair and assessed their relative support in a likelihood framework.

Speciation with gene flow is common across *Drosophila* pairs.

If current range is indicative of the mode of speciation, then we would expect that allopatric pairs best fit a history of strict isolation (SI) and that models that assume gene flow do not fit significantly better. In contrast, we may expect currently sympatric pairs to fit an isolation-with-migration (IM) model better than an SI model. We find that 81 out of 93 *Drosophila* pairs, regardless of their current distribution, fit an IM model significantly better than an SI model (Figure 1). The 12 pairs that show no support for gene flow are almost entirely from the *Drosophila nasuta* group, including the species pair *Drosophila albomicans* and *Drosophila nasuta*, which in previous analyses also showed no evidence for post-divergence gene flow (*31*). Contrary to our prediction, we find no difference in the level of support (as measured by ΔlnL) for an IM model over an SI model between allopatric and sympatric pairs (two-sample Wilcoxon test: $W = 876$, $p = 0.187$) (Figure 2). While our results suggests that the great majority of *Drosophila* pairs best fit a model of isolation with gene flow, 12 pairs that best fit an SI over an IM model include more allopatric than sympatric pairs (one-tailed Fisher exact test: $p = 0.052$). A power analysis based on data simulated under plausible parameters for *Drosophila* shows that these results are unlikely to be influenced by biases associated with multilocus inference of IM models (see Supplementary text). Thus, there is a pervasive signal of post-divergence gene flow in *Drosophila* pairs, even between currently allopatric pairs.

We then quantified the relative support for the IM model over a more complex history of isolation-with-initial-migration (IIM) which assumes that gene flow is limited to a period after divergence, after which species become completely reproductively isolated or spatially separated, with no contemporary gene flow between them. We find that for most pairs, we could not reject the IM model in favor of the IIM model (90 out of 93 pairs) (Figure 3A). We also find no difference in the proportion of allopatric and sympatric pairs that fit an IIM model better than an IM model (Fisher exact test; $p = 0.267$). Finally, there is no difference in the average support (ΔlnL between IM and IIM models) between allopatric and sympatric pairs, that is, the IIM model which assumes that gene flow has ceased is not better supported for allopatric pairs (two-sample Wilcoxon test: $W = 1023$, $p = 0.866$). We also compared support for the IIM model against a secondary contact (SC) model. The latter assumes an instantaneous burst of contemporary gene flow following a period of allopatry. We find that currently sympatric species have greater support for the secondary contact (SC) model (two-sample Wilcoxon test: $W = 1389$, $p = 0.007$) (Figure 3B). Additionally, a higher proportion of sympatric (30%) compared to allopatric pairs (11%) fit a secondary contact model better compared to an IIM model, (one-tailed Fisher exact test; $p = 0.050$). These results imply greater support for a secondary contact scenario in currently sympatric compared to allopatric pairs.

Our hierarchical approach of contrasting explicit, demographic models in a pairwise manner allows us to assign a best-fitting model to every pair in our dataset. We find that 77 out of 93 *Drosophila* pairs in our dataset (83%) best fit an IM model. Of the pairs that fit an IM model best, 30 (32%) were allopatric and 47 (52%) pairs were sympatric. In comparison, 12 out of 93 (13%) *Drosophila* pairs did not fit any models of gene flow significantly better than a strict isolation model (Supplementary Figure 8). Only three species pairs best fit a secondary contact model, all of which share some degree of current geographic overlap and incomplete premating isolation.

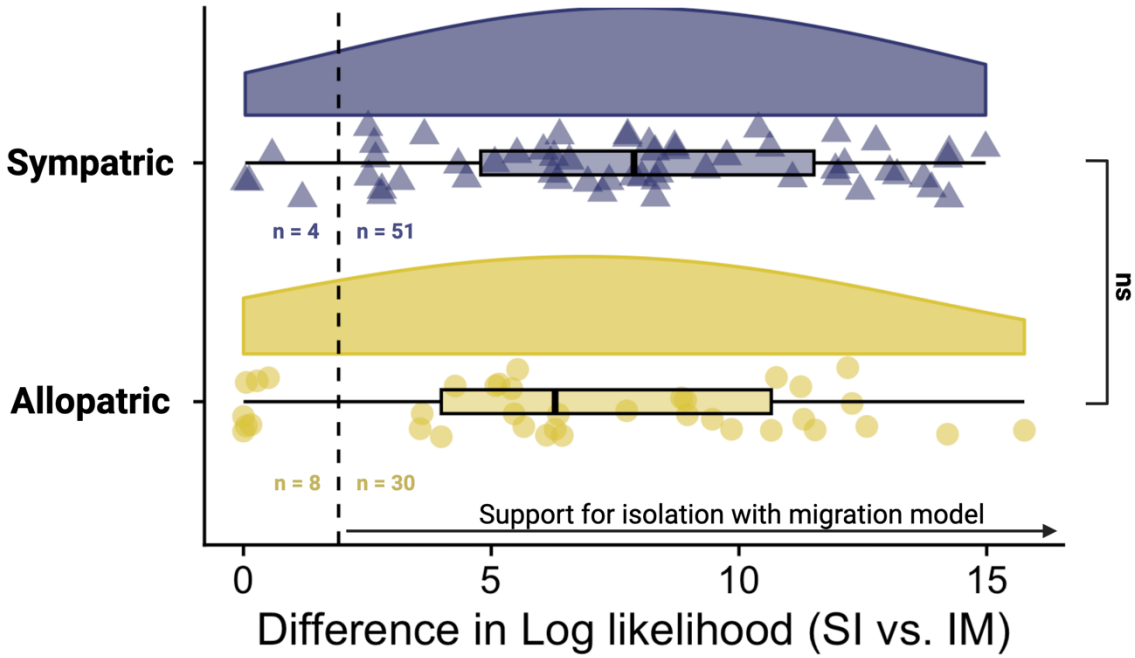


Figure 2: Consistent evidence for gene flow across allopatric and sympatric *Drosophila* pairs. The difference in log-likelihood between a strict isolation (SI) and an isolation-with-migration (IM) model for currently sympatric and allopatric pairs of *Drosophila* (represented by triangles and circles, respectively). The vertical dashed line indicates the critical value ($p < 0.05$). The numbers of species pairs that fit or do not fit an IM model significantly better than an SI model are shown on either side of the vertical line. 'ns' indicates no significant difference in log-likelihood between SI vs. IM between allopatric and sympatric pairs.

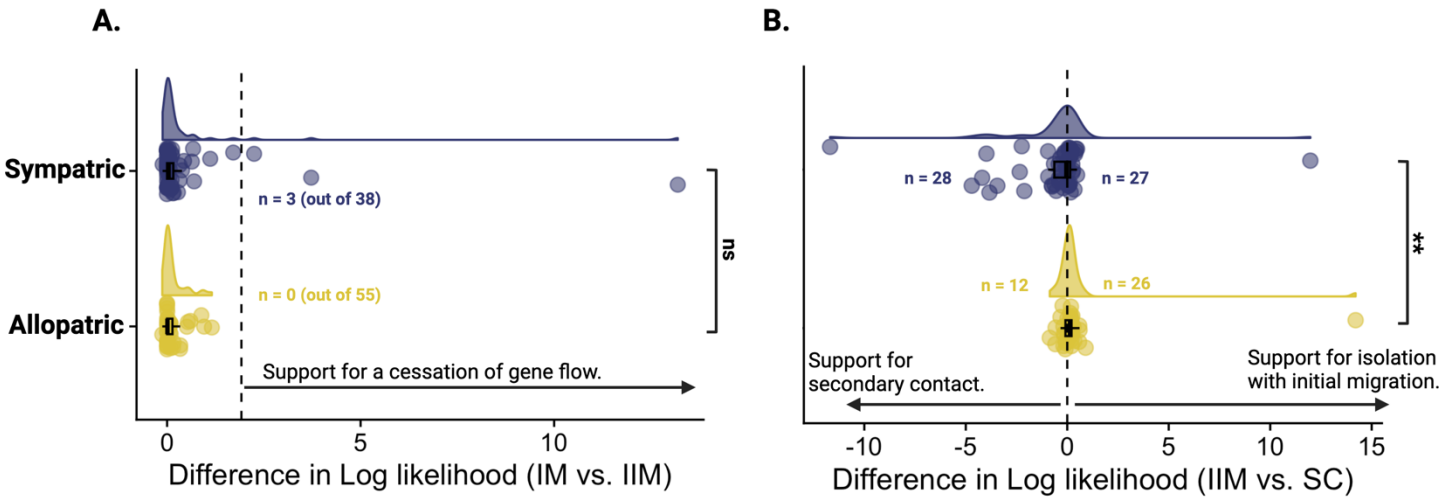


Figure 3: **Comparison of model support for isolation-with-initial divergence and secondary contact scenarios.** Difference in log-likelihood between **A.** an IM model and an isolation-with-initial-migration (IIM) model and **B.** an IIM model and a SC model. In **A**, the dashed line indicates critical value threshold ($p = 0.05$), with points (representing pairs) to the left of the dashed line showing support for the simpler demographic model. In **B.**, since the IIM and SC models share the same number of parameters, the dashed line simply represents 0, with better support for either comparison represented by the arrows shown in figure **B.** In **A.,** support for the complex model is indicated by the horizontal arrow. In **A.,** the number of pairs supporting the IIM model are given past the dashed horizontal line. In **B.**, the number of pairs that better fit the IIM and SC model are given on either side of the horizontal line. Stars indicate level of significance: <0.001 = ***, 0.01 = **, 0.05 = *, 'ns' = non-significant (*Mann-Whitney U)*.

Current range overlap poorly predicts historical demography.

To test whether historical gene flow differs between allopatric and sympatric pairs, we converted estimates of the scaled effective rate of gene flow (*M*) under the IM and IIM models, into a per lineage probability of gene flow (using the estimated duration of gene flow). This conversion allows for direct comparison between taxon pairs and models involving continuous (IM and IIM) and discrete (SC) gene flow. We find a lower long-term probability of gene flow for allopatric pairs (mean probability of migration: 66%) compared to sympatric pairs (mean probability of migration: 79%). When assuming the best fitting gene flow model for each pair, this difference is not significant (two-sample Wilcoxon test: $W = 802$, $p = 0.076$)(Figure 4A). However, if we assume an IM model for all pairs the probability of long-term gene flow is significantly lower in allopatric compared to sympatric pairs (two-sample Wilcoxon test: $W = 761$, $p = 0.026$).

In contrast, we find that currently allopatric and sympatric pairs differ neither in the onset of divergence (measured in generations) nor in ancestral effective population size. While sympatric pairs are younger (~6.03 MYA) on average than allopatric pairs (~6.2 MYA), this difference is not significant (two-sample Wilcoxon test: $W = 1154$, $p = 0.312$) (Figure 4B). Similarly, there is no difference in the estimates of the scaled ancestral population size estimates between allopatric (average ancestral N_e: 749,000) and sympatric pairs (average ancestral N_e: 904,000) (two-sample Wilcoxon test: $W = 801$, $p = 0.075$) (Figure 4C).

Secondary contact pairs do not have greater pre-mating isolation.

Reinforcement is frequently invoked to explain cases of enhanced mating discrimination in sympatry relative to allopatry (*2*, *32*). We asked whether pairs that fit a secondary contact (SC model) better than an isolation-with-initial-migration model ($\Delta\ln L < 0$; n = 40) show greater premating isolation, but not postzygotic isolation, than pairs that better fit an IIM model ($\Delta\ln L > 0$; n = 53). Contrary to this prediction, we find that pairs that better fit an SC model over an IIM model have lower premating isolation (mean pre-mating isolation: 0.81) than pairs that better fit an IIM model (mean pre-mating isolation: 0.88). However, this trend is not statistically significant (two-sample Wilcoxon test: $W = 1213$, $p = 0.122$). We observed an analogous trend for postzygotic isolation (SC mean postzygotic isolation: 0.37; IIM mean postzygotic isolation: 0.54).

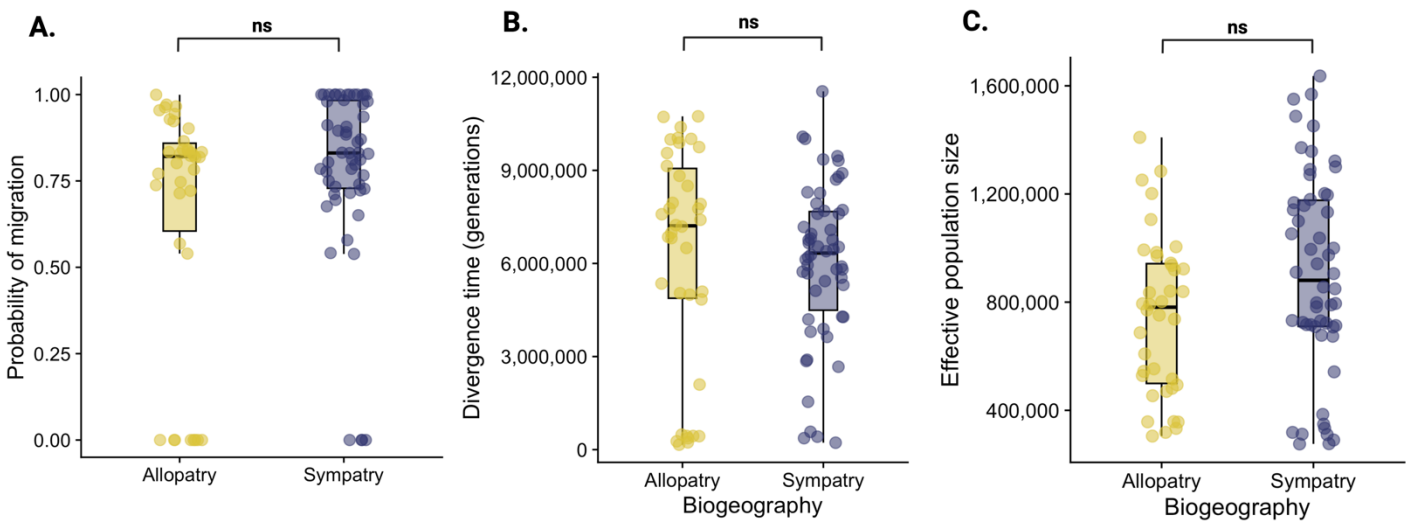


Figure 4: Sympatric and allopatric pairs show similar divergence times, effective population size and levels of long-term gene flow. Current range overlap does not reflect inferred historical demography. **A**) We find no difference in long-term rates of gene flow between allopatric and sympatric pairs. Gene flow was calculated for each pair using parameter estimates from the best-fitting model. **B** and **C**) We find no difference in scaled divergence time or effective population size (the best-fitting model for each pair) between allopatric and sympatric pairs. *'ns'* indicates no significant difference and '*' indicates $p < 0.05$ via *Mann-Whitney U* test.

Discussion

Conventional opinion amongst evolutionary biologists is that allopatric speciation is the most common mode of speciation because the cessation of gene flow during a period of allopatry, perhaps with subsequent reinforcement, facilitates evolutionary divergence (*2*). Using 93 *Drosophila* species pairs and a hierarchical modelling framework, we find that most pairs show evidence of post-divergence gene flow. Perhaps surprisingly, while we find some evidence for more gene flow between currently sympatric pairs, the overwhelming trend is for gene flow to be common, even between most currently allopatric pairs. There is little or no difference in statistical support for gene flow between pairs that are currently sympatric or allopatric. Sympatric pairs do seem to show more evidence for secondary contact histories and very recent gene flow, but historical rates of gene flow are similar to allopatric pairs. Our analysis suggests that most diverging species experience some level of long-term gene flow due, probably due to repeated cycles of allopatry and secondary contact (*16*, *33*). Speciation does not require an extensive allopatric phase, as predicted by the allopatric model of speciation (*13*, *12*).

These results have implications for interpreting classic comparative surveys in *Drosophila* and other taxonomic groups. Mating discrimination has been shown to evolve more rapidly in currently sympatric *Drosophila* species pairs relative to allopatric pairs, and these observations have been interpreted as support for reinforcement (*2*, *14*, *32*, *35*). Likewise, our finding of significantly greater support for the secondary contact model in sympatric compared to allopatric species pairs also suggests that reinforcement plays an important role in speciation. However, we find little evidence that overall levels of historic gene flow differs between currently allopatric and sympatric *Drosophila* pairs, which is more surprising. One explanation for this finding is the limited resolution of demographic inference: the modelling framework we use here can, by definition, only detect rather low rates of long-term gene flow. Moreover, there is an inherent limit to demographic inferences, because periods of high gene flow (or low N_e) erase the genomic footprints of older demographic events. In particular, our estimates of the scaled rate of migration M range between 0 and 10 (with a median = 2.91) and histories involving M >> 5 migrants per generation, i.e., the regime assumed by models of reinforcement might be hard to distinguish from panmixia. Thus, whilst our results do not contradict reinforcement as an explanation for the faster evolution of premating barriers in sympatry, they call into question the assumption that gene flow has largely been absent in currently allopatric taxa.

Our results strongly imply that speciation does not require an extended allopatric phase to allow the build-up of reproductive barriers. Some level of gene flow at all stages of speciation appears to be relatively common in *Drosophila* pairs. Supported by other genomic analysis across a broad range of taxa, current geographical patterns do not seem to predict historical levels of gene flow (*24*, *36*, *37*), so speciation with gene flow is common and the traditional classification into distinct geographic modes of speciation is outdated (*18*, *19*). We stress that demographic analyses of larger genomic datasets that make use of intraspecific variation and allow more realistic models of divergence may help to clarify the connection between our findings and previous comparative surveys. There is also a need for similar comparative analyses of speciation histories for taxa that have richer geographic range and life history information than *Drosophila* to

contextualise the evolution of reproductive barriers and the extent, timing and levels of gene flow in different geographical contexts.

Acknowledgments: We thank Sam Ebdon for bioinformatic advice. Roger K. Butlin for providing helpful feedback that improved the manuscript and Cher Chow for assistance with figures.

Funding: LHY was supported by a University of St Andrews studentship. KL was supported by a fellowship from the Natural Environment Research Council (NERC, NE/L011522/1) and a European Research Council starting grant (ModelGenomLand 757648) which also supported DRL.

Author contributions: Conceptualisation: LHY, KL, MGR. Analysis: LHY, DRL, KL. Writing (original draft): LHY, MGR. Writing (review and editing): LHY, DRL, KL, MGR.

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: Data availability detailed in the Supplementary material. Mathematica notebook, and bash and R scripts will be deposited in https://github.com/LeebanY/Speciation-continuum-Drosophila.

References

1. E. Mayr, *Systematics and the Origin of Species* (New York, NY : Columbia University Press, 1942).

2. J. A. Coyne, H. A. Orr, *Speciation* (Sinauer Associaties, Inc., 2004).

3. E. Mayr, *Animal Species and Evolution* (Harvard University Press; London: Oxford University Press, 1963).

4. P. Nosil, *Ecological Speciation* (Oxford University Press, 2012).

5. S. A. S. Anderson, J. T. Weir, The role of divergent ecological adaptation during allopatric speciation in vertebrates. *Science* **378**, 1214–1218 (2022).

6. R. Abbott, D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, R. K. Butlin, U. Dieckmann, F. Eroukhmanoff, A. Grill, S. H. Cahan, J. S. Hermansen, G. Hewitt, A. G. Hudson, C. Jiggins, J. Jones, B. Keller, T. Marczewski, J. Mallet, P. Martinez-Rodriguez, M. Möst, S. Mullen, R. Nichols, A. W. Nolte, C. Parisod, K. Pfennig, A. M. Rice, M. G. Ritchie, B. Seifert, C. M. Smadja, R. Stelkens, J. M. Szymura, R. Väinölä, J. B. W. Wolf, D. Zinner, Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013).

7. J. A. Coyne, H. A. Orr, Patterns of Speciation in *Drosophila*. *Evolution*, doi: 10.2307/2409213 (1989).

8. J. A. Coyne, H. A. Orr, "Patterns of speciation in *Drospohila* " Revisited. **51**, 295–303 (1997).

9. R. Yukilevich, Asymmetrical patterns of speciation uniquely support reinforcement in *Drosophila*. *Evolution* **66** (2012).

10. R. Yukilevich, The rate test of speciation: Estimating the likelihood of non-allopatric speciation from reproductive isolation rates in *Drosophila*. *Evolution* **68**, 1150–1162 (2014).

11. P. Nosil, Degree of sympatry affects reinforcement in *Drosophila*. *Evolution* **67** (2013).

12. M. Turelli, J. R. Lipkowitz, Y. Brandvain, On the Coyne and Orr-igin of species: Effects of intrinsic postzygotic isolation, ecological differentiation, X chromosome size, and sympatry on drosophila speciation. *Evolution* **68** (2014).

13. D. J. Funk, P. Nosil, W. J. Etges, Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3209–3213 (2006).

14. D. R. Matute, B. S. Cooper, Comparative studies on speciation: 30 years since Coyne and Orr. *Evolution* **75**, 764–778 (2021).

15. J. B. Losos, R. E. Glor, Phylogenetic comparative methods and the geography of speciation. *Trends Ecol. Evol.* **18**, 220–227 (2003).

16. G. M. Hewitt, Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* **58**, 247–276 (1996).

17. T. G. Barraclough, A. P. Vogler, Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.* **155**, 419–434 (2000).

18. B. M. Fitzpatrick, J. A. Fordyce, S. Gavrilets, Pattern, process and geographic modes of speciation. *J. Evol. Biol.* **22**, 2342–2347 (2009).

19. R. K. Butlin, J. Galindo, J. W. Grahame, Sympatric, parapatric or allopatric: The most important way to classify speciation? *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 2997–3007 (2008).

20. A. Suvorov, B. Y. Kim, J. Wang, E. E. Armstrong, D. Peede, E. R. R. D'Agostino, D. K. Price, P. Waddell, M. Lang, V. Courtier-Orgogozo, J. R. David, D. Petrov, D. R. Matute, D. R. Schrider, A. A. Comeault, Widespread introgression across a phylogeny of 155 *Drosophila* genomes. *Curr. Biol.* **32** (2022).

21. N. B. Edelman, P. B. Frandsen, M. Miyagi, B. Clavijo, J. Davey, R. B. Dikow, G. García-accinelli, S. M. V. Belleghem, N. Patterson, Genomic architecture and introgression shape a butterfly radiation. *Science* **599**, 594–599 (2019).

22. M. C. Fontaine, J. B. Pease, A. Steele, R. M. Waterhouse, D. E. Neafsey, I. V. Sharakhov, X. Jiang, A. B. Hall, F. Catteruccia, E. Kakani, S. N. Mitchell, Y. Wu, H. A. Smith, R. R. Love, M. K. Lawniczak, M. A. Slotman, S. J. Emrich, M. W. Hahn, N. J. Besansky, Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science* **347**, 1258524 (2015).

23. C. Roux, C. Fraïsse, J. Romiguier, Y. Anciaux, N. Galtier, N. Bierne, Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence. *PLoS Biol.* **14**, 1–22 (2016).

24. A. De Jode, A. Le Moan, K. Johannesson, R. Faria, S. Stankowski, A. M. Westram, R. K. Butlin, M. Rafajlović, C. Fraïsse, Ten years of demographic modelling of divergence and speciation in the sea. *Evol. Appl.* **16**, 542–559 (2023).

25. R. M. Kliman, P. Andolfatto, J. A. Coyne, F. Depaulis, M. Kreitman, A. J. Berry, J. McCarter, J. Wakeley, J. Hey, The Population Genetics of the Origin and Divergence of the *Drosophila simulans* Complex Species. *Genetics* **156**, 1913–1931 (2000).

26. C. A. Machado, R. M. Kliman, J. A. Markert, J. Hey, Inferring the history of speciation from multilocus DNA sequence data: The case of *Drosophila pseudoobscura* and close relatives. *Mol. Biol. Evol.* **19** (2002).

27. J. Hey, R. Nielsen, Multilocus Methods for Estimating Population Sizes, Migration Rates and Divergence Time, With Applications to the Divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747–760 (2004).

28. J. A. Coyne, H. A. Orr, Patterns of Speciation in *Drosophila*. *Evolution*, doi: 10.2307/2409213 (1989).

29. H. M. Wilkinson-Herbots, The distribution of the coalescence time and the number of pairwise nucleotide differences in the “isolation with migration” model. *Theor. Popul. Biol.*, doi: 10.1016/j.tpb.2007.11.001 (2008).

30. H. M. Wilkinson-Herbots, The distribution of the coalescence time and the number of pairwise nucleotide differences in a model of population divergence or speciation with an initial period of gene flow. *Theor. Popul. Biol.* **82**, 92–108 (2012).

31. D. Bachtrog, The speciation history of the *Drosophila nasuta* complex. *Genet. Res.* **88**, 13–26 (2006).

32. M. A. F. Noor, Reinforcement and other consequences of sympatry. *Heredity* **83**, 503–508 (1999).

33. G. M. Hewitt, Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences* (2004)vol. 359.

34. R. Butlin, Speciation by reinforcement. *Trends Ecol. Evol.*, doi: 10.1016/0169-5347(87)90193-5 (1987).

35. S. Singhal, G. E. Derryberry, G. A. Bravo, E. P. Derryberry, R. T. Brumfield, M. G. Harvey, The dynamics of introgression across an avian radiation. *Evol. Lett.* **5**, 568–581 (2021).

36. J. V. Peñalba, L. Joseph, C. Moritz, Current geography masks dynamic history of gene flow during speciation in northern Australian birds. *Mol. Ecol.* **28**, 630–643 (2019).

Supplementary - Materials and methods

1.1 Data sampling and quality control.

To infer the demographic histories of species pairs of *Drosophila*, we collated data on reproductive isolation, range overlap and genetic distance for *Drosophila* species pairs originally collated by Coyne and Orr (*1*, *2*), and expanded by Yukilevich (*3*, *4*), representing the largest collection of reproductive isolation measurements in *Drosophila*. As in (*1*, *2*), and other comparative surveys of *Drosophila* (*3–7*), the term "species pairs" throughout these analyses refers to closely-related but non-independent pairs of species rather than independent sister taxa. Our goal was to infer speciation histories for as many species pairs found in (*1*, *2*) as possible to ensure results were consistent with previous comparative analyses. Briefly, premating isolation from previous comparative surveys was calculated for each species pair as the percentage of successful copulations for homotypic and heterotypic matings during no-choice, single-choice, and multiple-choice mating experiments. Postzygotic isolation was calculated by considering the percentage of F1 hybrid males and females that were sterile or inviable. For each reciprocal mating, measures ranged from 0 (F1 hybrids are viable and fertile) to 1 (all hybrids produced are either sterile or inviable). The overall postzygotic isolation index is the average of both reciprocal mating interactions. For detailed descriptions of reproductive isolation measurements and range determinations, see references (*3*, *4*).

We sought to combine reproductive isolation data and range overlap information with demographic analyses of genomic data. We obtained genome assemblies from publicly available datasets via NCBI for species with measures of reproductive isolation, range overlap and genetic distances, including large-scale efforts to produce highly contiguous *Drosophila* genome assemblies (*8*) (see Supplementary Table 1-3). To standardize quality and quantity of

whole-genome sequencing datasets, we filtered whole-genome sequencing datasets for each species using sample and project metadata (see Supplementary Table 1 for a list of samples used). Specifically, we restricted our analyses to high coverage datasets (i.e. total bases between 1 GB and 50 GB) and removed pool-sequenced samples or samples that had been experimentally manipulated before sequencing was performed. We utilized datasets with wild-caught individuals where possible. Finally, we randomly sampled a single whole-genome sequencing dataset for each species in the dataset, to minimize bias in selecting any given dataset for a particular species.

1.2 Genome annotation, mapping and variant calling

For species pairs with multiple genome assemblies, we assessed genome quality (using BUSCO v5.1.2 single-copy completeness percentages) and selected the genome with the highest BUSCO score (>90% single-copy completeness) (*9*) as the reference for each pair. Annotations for each reference genome were performed in the same way to minimize differences in gene prediction. Annotation was performed for all genome assemblies using BRAKER2 v2.1.6 (*10*). We used only *D. melanogaster* (Flybase v.6.39) proteins as evidence for all annotations. Subsequently, we extracted intronic regions from each genome annotation using BEDOPS v2.4.39 and bedtools v2.30.0 (*11*, *12*).

For each species, a single dataset was randomly-sampled using a custom R script which selected a single sample from a list (filtering described above) of publicly-available samples. Each randomly sampled whole-genome sequencing dataset (representing a single sample) was downloaded for each species in every pair using the sratools functions pre-fetch v2.11.0 and fasta-dump v2.11.0. All reads were trimmed using fastp v.0.20.1. For every pair, a single genome assembly was indexed (Supplementary Table 3) using bwa-mem2 for species with

paired-end reads, and bwa-mem for species with single-end reads (*13*, *14*). Species-specific paired-end or single-end reads were then mapped to pair-specific genome assemblies using bwa-mem2 (v. 2.2.1) and bwa-mem (v. 0.7.17) (*13*, *14*). Where a representative genome for mapping and variant calling did not exist for either species in a given pair, we used a genome belonging to an outgroup species not within the species pair but closely-related; this was only done for a handful of species pairs. Reads were sorted, duplicates mapped and read groups were added using sambamba v. 0.8.1 and picard v2.26.3(*15*, *16*). Trimming and mapping were implemented in the workflow engine Snakemake v. 6.5.1 (*17*). Variants were called for reads in each pair using Freebayes v1.3.5 and subsequently filtered for missing genotypes, read mapping bias (RPL<1 | RPR<1 | SAF<1 | SAR<1), and a minimum read depth of 2, using bcftools v1.11 and gIMble v.0.6.1 (*18–20*). We refer to pair-specific datasets which have been filtered as above as 'callable and filtered' datasets. The workflow for the filtering process is summarized in Supplementary Figure 13.

1.3 Generating intronic blocks datasets across species pairs.

To generate comparable datasets for demographic modelling across all pairs, we sampled intronic blocks from available pair-specific, callable and filtered variants. Our rationale for sampling intronic blocks, as opposed to coding and/or intergenic regions, was two-fold (Supplementary Figure 13): first, functional constraint in coding regions make them less suitable for demographic inference based on samples of contiguous blocks that are assumed to evolve neutrally. Second, although intergenic regions show, on average, similar levels of sequence conservation to introns in *Drosophila*, correctly annotating and distinguishing conserved from selectively-neutral intergenic regions is challenging (*21–23*). Instead, we decided to utilize introns since annotation of genic regions is comparatively easier and more

reliable across pairs than filtering non-genic regions interspersed by poorly-annotated repetitive regions. Since functional constraint varies across introns in *Drosophila*, we applied additional filters to the filtered callable datasets for all pairs (*21, 23*). Specifically, first introns were removed because they evolve more slowly than non-first introns (*21*), likely owing to their higher abundance of regulatory elements. We also removed the first and last 10 bases from all remaining introns in our analyses using bedtools v2.30.0, and a custom python script.

We intersected filtered introns with callable positions using bedtools intersect v2.30.0. To minimise the effect of linkage between blocks on our analysis, each gene could only contribute a single intronic interval to each pair-specific dataset. To confirm that our dataset of filtered introns consisted of intronic intervals that were more selectively neutral than the background, we compared absolute genetic divergence d_{XY} calculated across all callable sites across the whole-genome to d_{XY} calculated across only filtered introns. As expected, (given the near absence of functional constraint in introns), d_{XY} in filtered introns was higher than the d_{XY} estimate across the whole genome in all pairs (Supplementary Figure 14).

For each pair, we partitioned filtered introns into a dataset of short intronic blocks using gIMble v.0.6.1. Additionally, for every pair, the block length was chosen to give an average of three pairwise differences (pair-specific block length = $3/d_{XY}$) between species per block. Pairs with fewer than 250 blocks were removed from the analysis. Additionally, pairs with higher within-species heterozygosity compared to fixed pairwise differences between species were also removed from the analysis. These filters removed 9 species pairs from the analysis. For the remaining 93 species pairs, block sizes ranged from 60 – 341, and the number of blocks ranged from 367 – 6390 (Supplementary Table 2). The final datasets for each species consisting of pair-specific, filtered, short intronic blocks (hereinafter referred to as "short intronic blocks" or

SIBs) were converted into blockwise site-frequency spectra (bSFS), containing counts of all possible mutational configurations, using gIMble (v.0.6.1).

1.4 Modelling gene flow and divergence.

For each pair, we computed average genetic divergence (d_{XY}) for filtered, callable sites across the whole genome and in SIBs. To understand demographic histories and the degree of gene flow for species pairs, we fitted a range of models of demographic history to the distribution of pairwise differences between species: (a) **a strict isolation (SI) model** (most consistent with *strict* allopatric speciation) characterised as an instantaneous split of an ancestral population at T_0 without gene flow, (b**) isolation with migration (IM)**, where an ancestral population diverges with symmetric migration (M = $4N_e$m migrants per generation, where N_e is the ancestral population) between the time of divergence and the present, (c) an **isolation with initial migration model (IIM)**, where an ancestral population diverges with an initial period of symmetric migration and gene flow ceases at T_1 and (d) **a secondary contact (SC) model**, where an ancestral population diverges in allopatry, and an instantaneous recent, pulse of gene flow, i.e. a total proportion (*f*) of the population is introgressed at time T_1. In all cases, we assumed a single N_e parameter that is shared between the ancestral population and the daughter species. Note that, we are only considering the interspecific distribution of pairwise differences (S) which is little affected by variation in intraspecific N_e. See Supplementary Figure 7 for visualisations of each model.

To test for the ability to identify an initial divergence time, we also fitted an isolation with **initial migration limit** (**IIM Limit**) and a **migration-only (MIG) model** both of which also do not have an identifiable initial divergence time, i.e. T_0 → infinity (Supplementary Figure 8). The **MIG model** assumes an infinite period of continuous gene flow (M) and an effective

population size parameter. All other models contain an effective population size parameter ($\theta = 4N_e \mu$), and an initial divergence time parameter (T_0) (Supplementary Figure 8). The strict divergence model contains only these two parameters, and the IM contains an additional symmetric migration rate parameter ($M = 4N_e m$ migrants per generation). The isolation-with-initial-migration and the secondary contact model are nested within the IM model and assume that gene flow ceased at time T_1 (1).

Extracting genomic DNA for a single sample for whole-genome sequencing of *Drosophila* typically requires pooling multiple individuals from isofemale lines. Pooling multiple individuals may be problematic for demographic analysis by complicating inference of local genomic phase. Specifically, it is difficult to determine the phase of shared and sample-specific heterozygous sites using minimal whole-genome sampling. To circumvent this, we implemented a binomial sampling procedure to consider all ways of randomly phasing heterozygous sites in SIBs (implemented in Mathematica v12.3).

We considered the S distribution in filtered, intronic data to fit the models described above to each species pair. For a single sequence per species, and assuming an infinite sites mutation model and no recombination within blocks, the expected distribution of pairwise differences has been analytically derived for the models considered (*24*, *25*). For each pair, we obtained maximum-likelihood parameter estimates under all five models and since most models are nested, we used likelihood ratio tests for model comparison. Maximum likelihood parameter estimation was performed in *Mathematica* v12.3. The strict divergence model is nested within the IM model, which in turn is nested within the IIM model. Critical values for nested models were determined given a χ^2 distribution at a significance level of 0.05. To assign a best-fitting model to each species pair, we compared model support in a pairwise fashion between

competing models. To scale parameter estimates, we used the spontaneous mutation rate estimate of 3.32 x 10^{-9} per base and generation for *Drosophila melanogaster* and additionally include scaled parameter estimates based on 95% confidence intervals of the mutation rate estimate of *D. melanogaster* (2.52 x 10^{-9} − 4.30 x 10^{-9})(*26*).

1.5 Simulations – examining the effect of realistic recombination rates on model fit.

To understand how our inference based on the S distribution is impacted by recombination within blocks and to test whether the pervasive signal of long term gene flow could be an artefact of ignoring recombination within blocks, we used *msprime* (*27*, *28*) to simulate blocks of the same size as our empirical dataset (Supplementary Table 2) under the best fitting SI history for each species pair. For every taxon pair we simulated data without recombination and with recombination assuming a rate of ρ = 1.03 cM/Mb which is half the average recombination in *D. melanogaster* females to account for the achiasmy in males (ρ= 2.06 cM/Mb)(*26*). Additionally, we assumed a mutation rate of 3.32 x 10^{-9} as mentioned above. For every pair, we simulated 2,000 blocks.

To quantify the bias in parameter estimates and to determine the false positive rate of model selection with and without recombination, we fitted simulated S distributions for every pair to the SI and IM models and assessed the difference in log-likelihood as detailed above for the empirical data. Simulation results can be found in the Supplementary Text section.

1.6. Statistical Analyses

1.6.1 Evaluating difference in support for gene flow between pairs based on current range overlap.

Statistical analyses were performed in R using the package stats v.4.0.2, and in *Mathematica* (*29*). We hypothesised that if current range overlap is predicted by the demographic history of speciation, sympatric pairs should (1) fit an IM model better than an SI model more often than allopatric pairs, and (2) show greater support for an IM model over an SI model compared to allopatric pairs. We tested these two predictions by calculating the difference in log-likelihood (ΔlnCL) between the IM model and SI model. Pairs showed support for an IM model over an SI model if ΔlnCL was larger than the critical threshold ($p < 0.05$). We compared the number of allopatric and sympatric pairs fitting an IM model over an SI model using Fisher's exact test. To evaluate whether there was a difference in average support (ΔlnCL) for an IM model over an SI model between allopatric and sympatric pairs, we used *Mann-Whitney U* tests. This framework was used to test the difference in support between allopatric and sympatric pairs, and difference in the proportion of pairs fitting a given model over another for four other pairwise model comparisons. The other model comparisons include:

1. IM model vs. migration-only (MIG)
2. Isolation with migration (IM) vs. Isolation with initial migration (IIM).
3. Isolation with initial migration (IIM) vs. Secondary Contact (SC).
4. Isolation with initial migration (IIM) vs. Isolation-with-initial migration limit (IIM Limit) model.

We also assessed differences in the estimates of the scaled migration rate (M) between allopatric and sympatric pairs across models with gene flow (IM, IIM, SC) using *Mann-Whitney U* tests. Additionally, we computed the probability that an individual lineage is involved in gene flow (bounded from 0 – 1) as $1 - e^{-\tau 0 * M}$ under the IM and IIM model. This

allows to compare M estimates obtained under the IM and IIM model directly with the admixture fraction (*f*) which can be thought of as a per lineage probability between 0 – 1. Similarly, we asked whether contemporary allopatric and sympatric pairs differed in the average probability of migration and/or the scaled divergence times and effective population sizes using *Mann-Whitney U* tests.

1.6.2 Assigning a best-fitting model.

We assigned species pairs a best-fitting and most parsimonious model using the nested model comparisons described above. For example, the SI model was deemed the best-fitting model for pairs that did not fit an IM model significantly better. Alternatively, pairs which fit an IM model better than a SI and MIG model and did not show significant support for an IIM model over an IM model, fit an IM model best. Pairs that fit an IIM model significantly better than an IM model, and where an IM model fit significantly better than a SI and MIG model best-fit an IIM model overall. For pairs best-fitting an IIM model, but where an IIM model did not fit significantly better than an IIM limit model, these pairs fit an IIM limit model best overall. For pairs best fitting an IIM model after a comparison with the IIM limit model, a final, non-nested model comparison was performed between the IIM model and the secondary contact (SC) model. Pairs showing better support for a secondary contact model over an IIM model in this final comparison best fit a SC model overall. See Supplementary Figure 15.

Supporting Text

Simulation results.

To understand the potential impact of recombination on our inference of demographic history, and to test whether the pervasive signal for gene flow we find across Drosophila species pairs can arise as an artefact of ignoring recombination within blocks, we carried out a simulation-based sensitivity analysis. We simulated speciation histories for each taxon pair under an SI model given the divergence time and ancestral effective population size estimated from the empirical data. We simulated two pseudo-observed datasets (PODS): (1) without recombination and (2) with realistic recombination rate variation (see section 1.5). For both simulated datasets, we asked i) whether there is significant support for an IM relative to an SI model and ii) how rates of gene flow inferred from simulated data compare to empirical estimates.

When simulating blockwise data under a history of strict divergence and no recombination, none of the PODS (0 out 93 species pairs) erroneously fit an IM model over a SI model (Supplementary Figure 12, close to the expected false positive rate (FPR) given the $\chi 2$ test we performed on ΔlnL) and parameter estimates closely match the true value. However, when simulating under a SI model with recombination, 24 out of 93 PODS erroneously support an IM over an SI model which corresponds to a FPR of 25%. However, comparing the relative support (ΔlnL) between SI and IM for PODS (simulated under SI with recombination) with the empirical estimates shows that the IM model has considerably higher support in the real data compared to the PODS simulated without gene flow (two-sample Kolmogorov-Smirnov test: $D = 0.79$, $p = <0.001$).

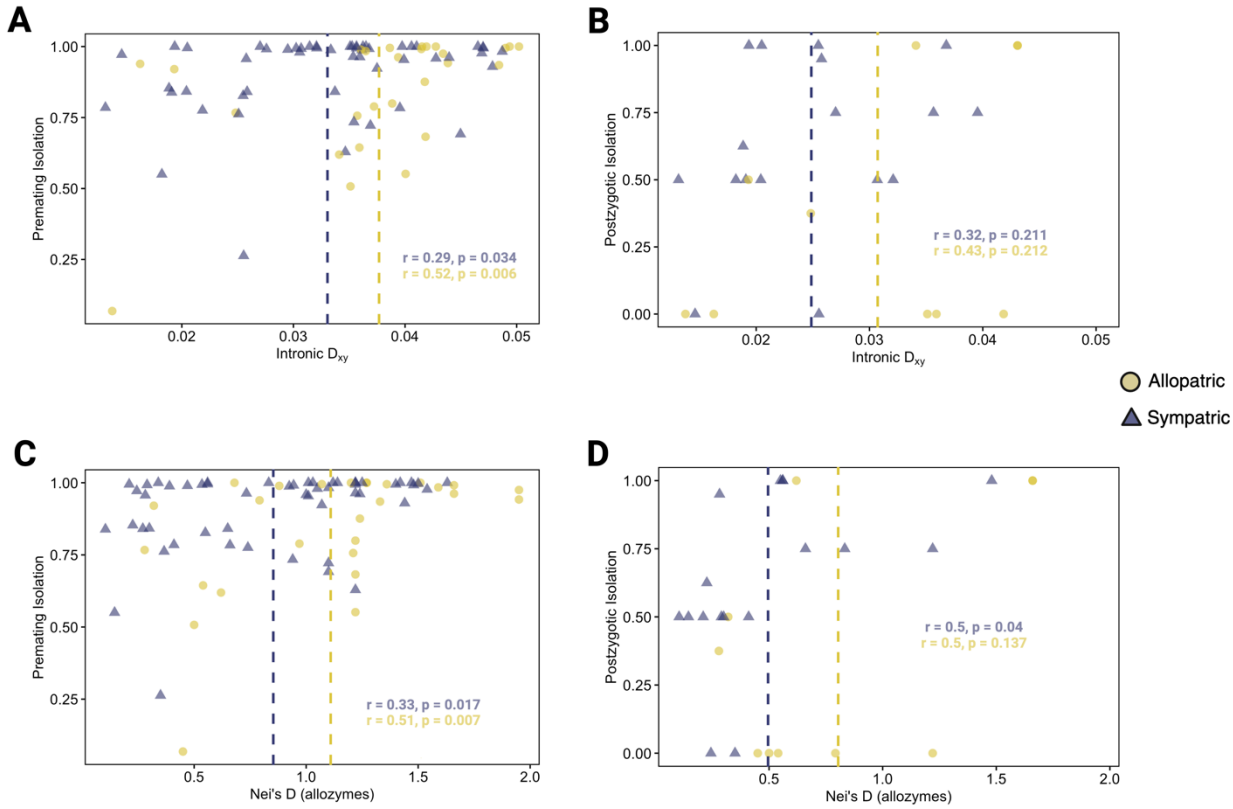
Furthermore, we investigate the impact of recombination on our inference by comparing scaled rate of gene flow (M) and divergence time (T0) estimates for PODS simulated under a SI model with recombination, with empirical parameter estimates. In the 77 species pairs that best fit an IM model in the empirical data the M estimates for the corresponding PODS (simulated with recombination under SI) are much lower than the empirical estimates (two-sample Kolmogorov-Smirnov test: D = 0.45, p= <0.001). Similarly, estimates for divergence time (two-sample Kolmogorov-Smirnov test: D = 0.58, p= <0.001) are also considerably lower in the simulated estimates compared to the empirical estimates across the 77 pairs. Thus, we conclude that while plausible rates of recombination may result in erroneous inference of gene flow, this effect of recombination on inference can neither explain the amount of gene flow we infer, nor the strength of support for speciation histories that involve gene flow that we find in the empirical data.

Little support for limit models over IM and IIM models.

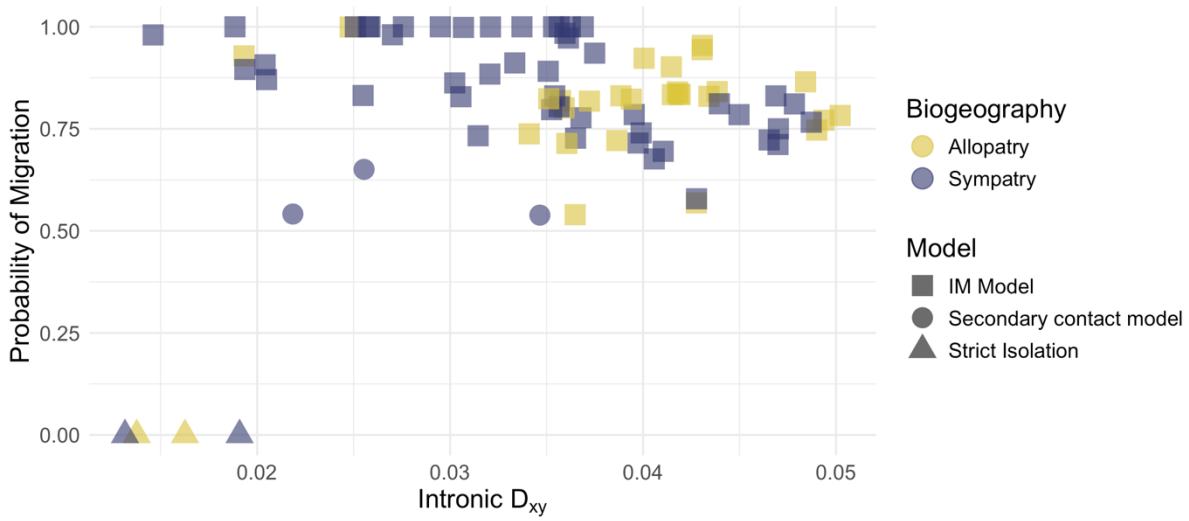
Models that contain no initial divergence time and so no definable starting point for gene flow are useful for identifying pairs where gene flow may have largely erased the information contained in the S distribution about the onset of species divergence. Comparing the MIG and IIM limit models (see Supplementary figures 10 and 11) against IM and IIM models respectively clarifies whether definable starting times for divergence are discernible across pairs. For most pairs (n = 87), we found significant support for an IM model over a MIG model ($p < 0.05$) (Supplementary Figure 11). With one exception, all pairs that did not fit an IM model significantly better than a MIG model showed greatest support for a SI model. Only *D. bipectinata – D. parabipectinata* best fit a MIG model in our modelling framework. These two species are largely sympatric, show relatively low levels of absolute genetic divergence (d_{XY} =

0.018) and based on previous phylogenetic analyses are extremely closely-related (*30*). For the IIM Limit model, only 14 out of 93 pairs did not fit an IIM model significantly better than an IIM Limit model. For 12 of these 14 pairs, the SI model was the best supported model, for one pair each the best supported model was an IM and a MIG model respectively (the same pair mentioned above). Overall, our results suggest that the S distribution is sufficient to identify the onset of divergence for almost all pairs.

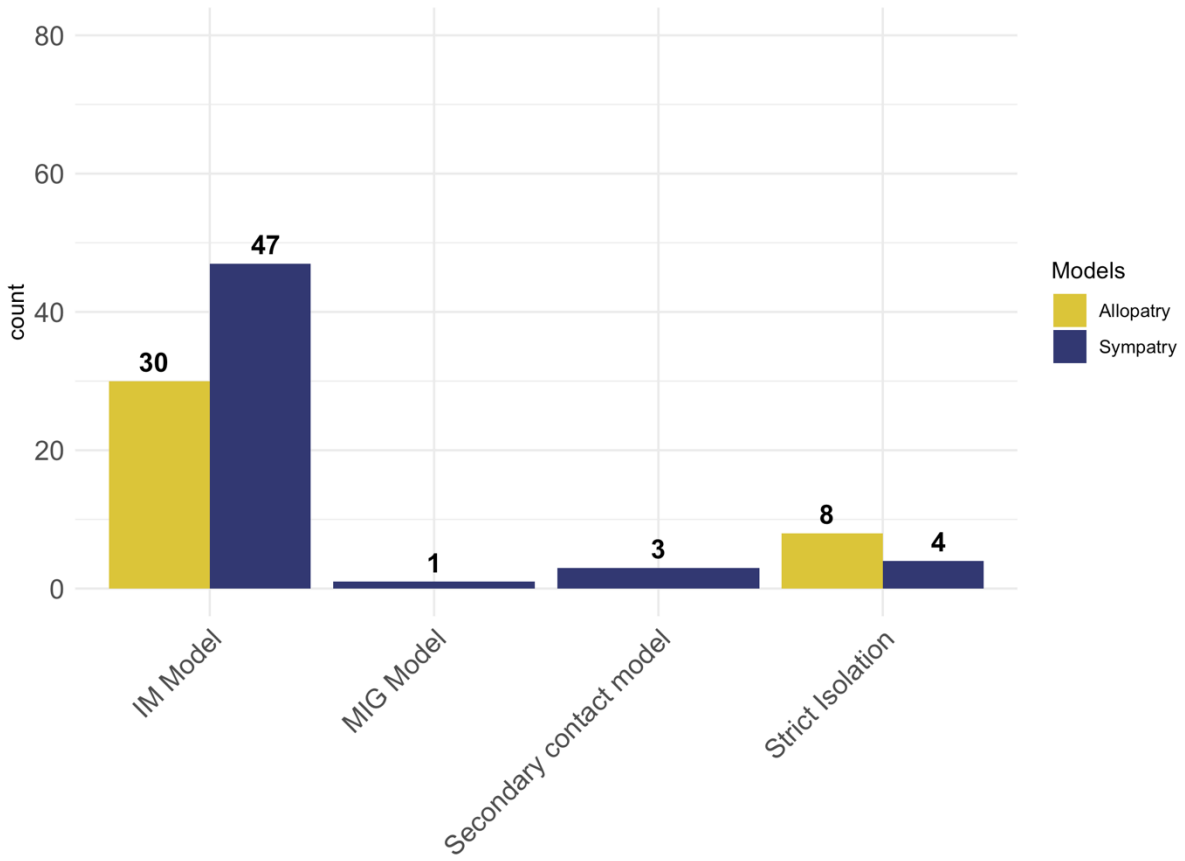
Supplementary Figures



Supplementary Figure 1: Relationship between genetic divergence and measures of reproductive isolation. Genetic divergence calculated based on short intronic blocks sampled from whole-genome data in this study (A and B), and previous estimates of genetic divergence (Nei's D) based on allozymes (C and D). Dashed lines on the top two panels indicate mean intronic d_{xy} (A and B) for allopatric and sympatric species pairs, and dashed lines in the bottom two panels indicate mean Nei's D calculated from allozymes for allopatric and sympatric pairs. A and C show estimates of premating isolation on the y-axis, and B and D show estimates of postzygotic isolation taken from previous comparative surveys of *Drosophila*. Pearson correlation coefficients and p-values for the relationship between measures of RI and measures of genetic divergence for sympatric and allopatric pairs are shown in each panel.



Supplementary Figure 2: Relationship between d_{xy} (absolute genetic divergence in filtered introns) and probability of migration based on the best-fitting model. Pairs either fit an IM model with continuous gene flow, a secondary contact model with a burst of gene flow or a strict isolation model best (denoted by different shapes). Biogeography (whether pairs are currently allopatric or sympatric) is denoted by colours. The MIG model contains an infinite period of continuous gene flow and no definable divergence parameter, thus it is not possible calculate a probability of long-term migration for the one pair that showed best support for the MIG model.



Supplementary Figure 3: Pairs and their best-fitting models. Pairs categorised into allopatry and sympatry, and the values on top of bars represent the number of allopatric and sympatric pairs best-fitting one of the three models. We found no pairs best supporting an IIM model, so IIM is not shown in the figure. Similarly, no allopatric pairs fit an SC or MIG model best, so these are not represented here.

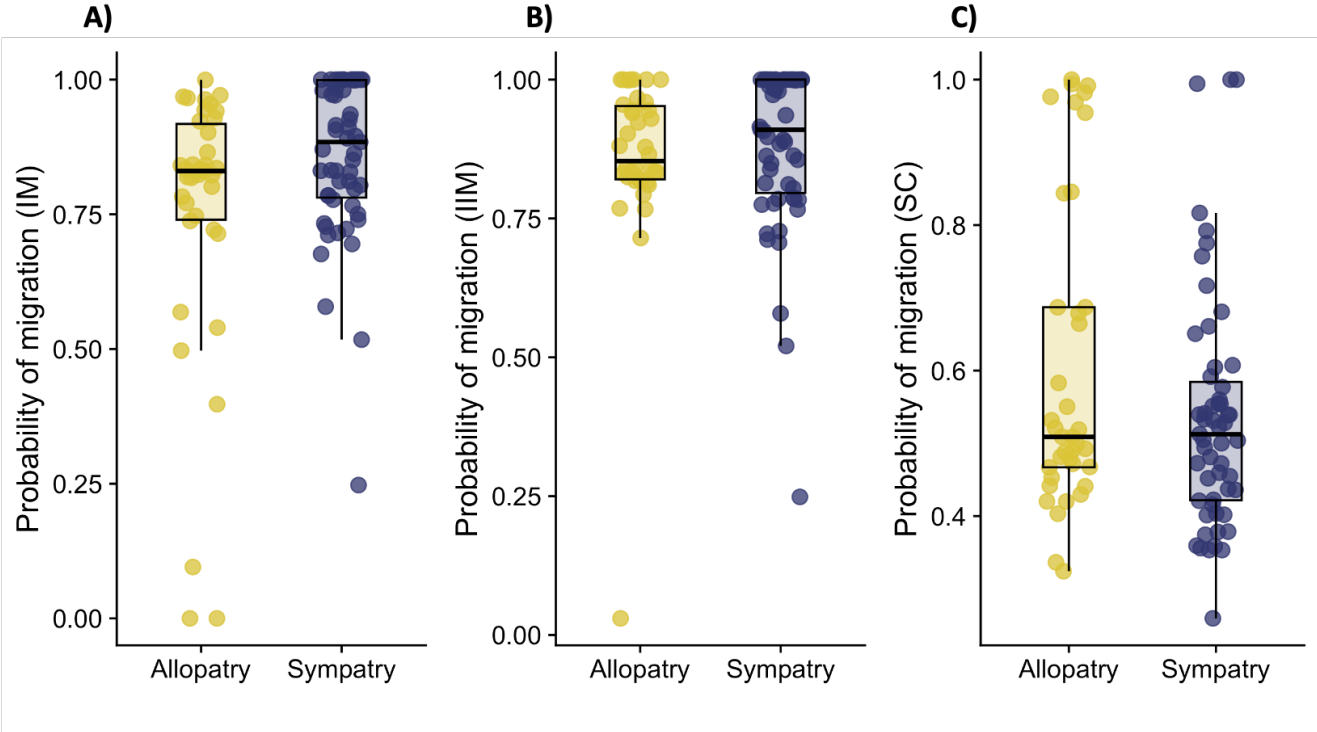
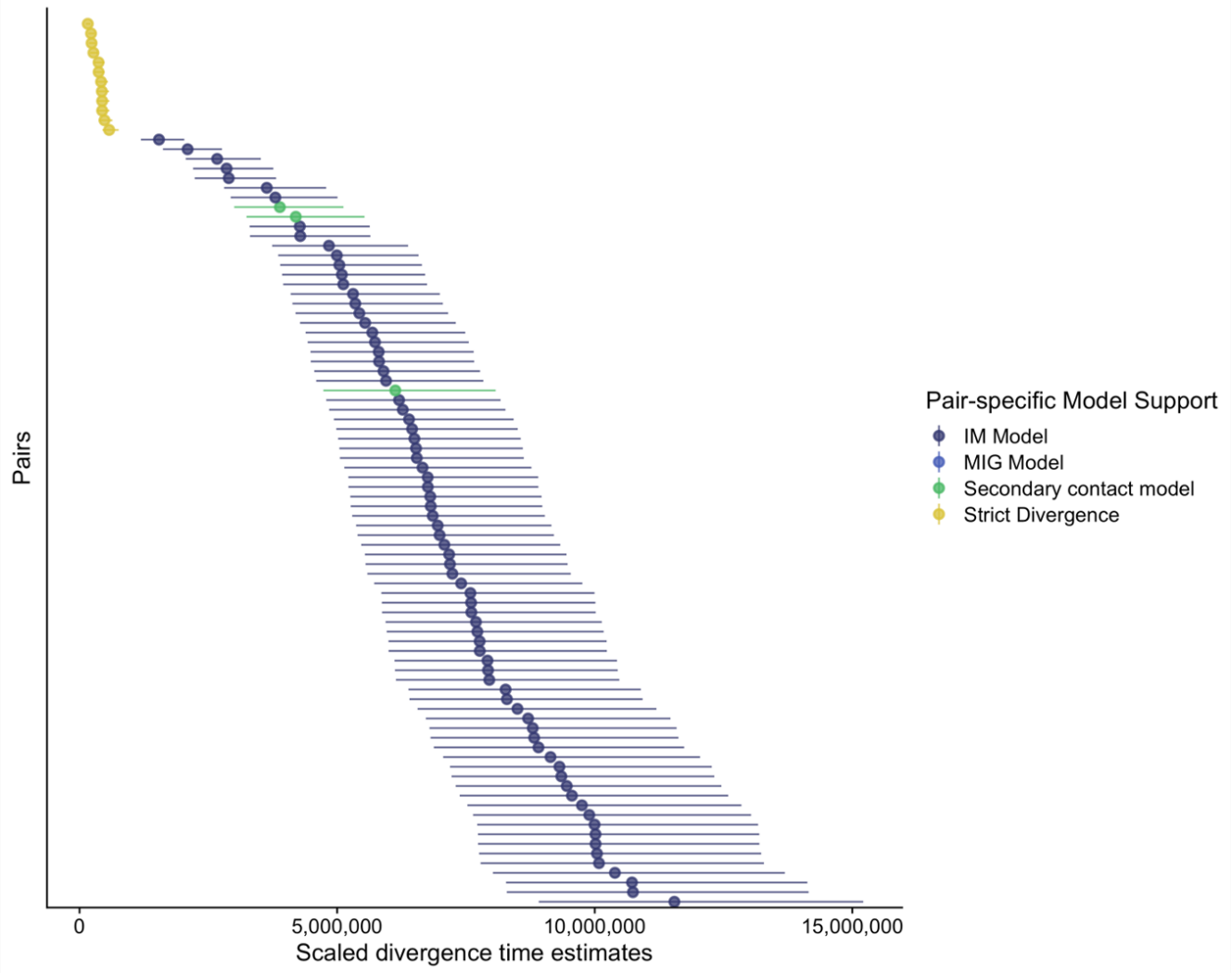
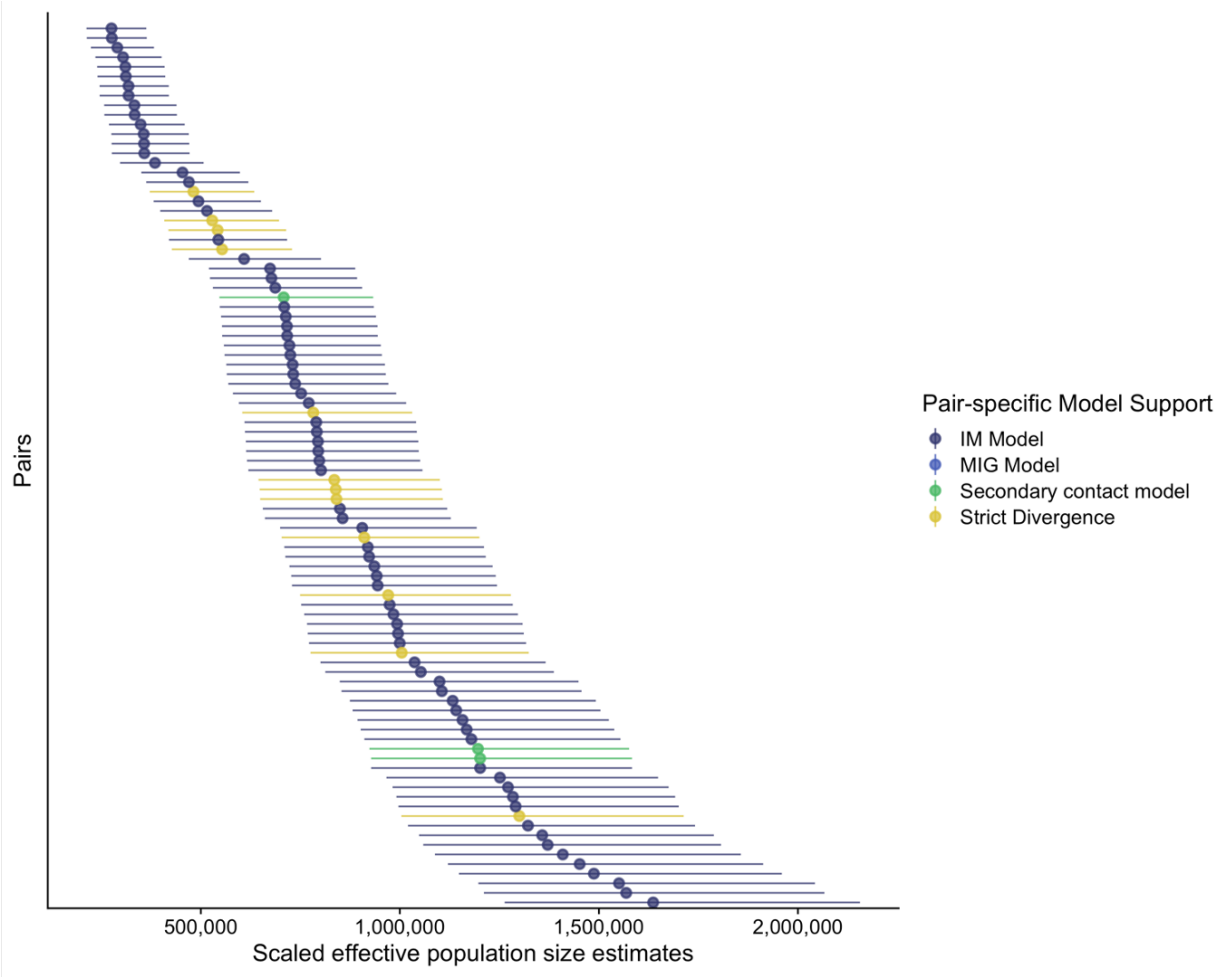


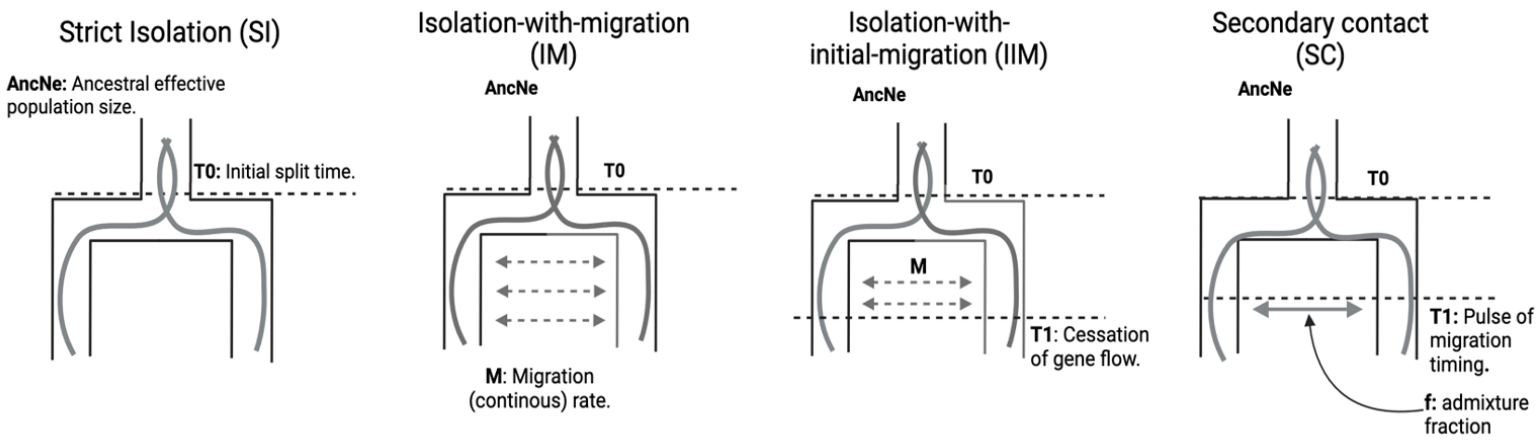
Figure Supplementary 4: Probability of migration under different demographic histories of species divergence. For A, B and C), plots compare the probability of a single lineage migrating during the migration phase. A) Sympatric pairs are inferred to have a higher probability of migration during a continuous migration phase than allopatric pairs under an IM model. B) Similar levels of inferred migration are seen under a limited phase of migration following initial divergence. C) Similar levels of inferred migration are seen during a short admixture phase under a secondary contact scenario.



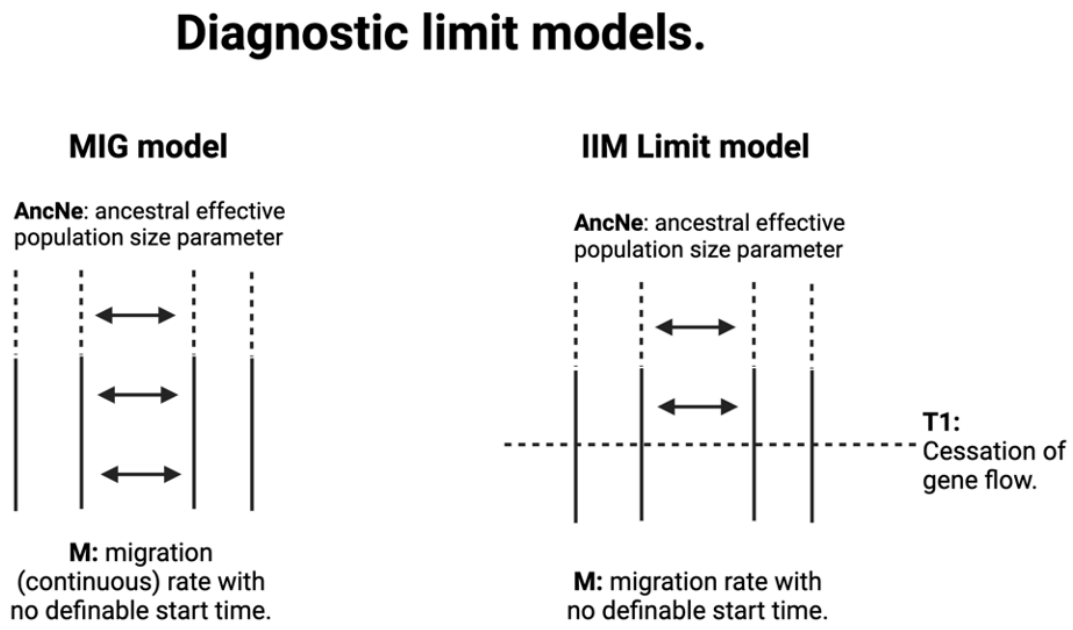
Supplementary Figure 5: Scaled initial divergence time estimates for pairs using best-fitting model parameter estimates. Parameter estimates were scaled using an across population *D. melanogaster* mutation rate of 3.32 x 10^{-9}, represented by the circles. Upper and lower bound of divergence time for each were scaled using the upper and lower bounds of the D. melanogaster mutation rate (upper: 2.52 x 10^{-9}; lower: 4.30 x 10^{-9}). Pairs are coloured by their best-fitting model.



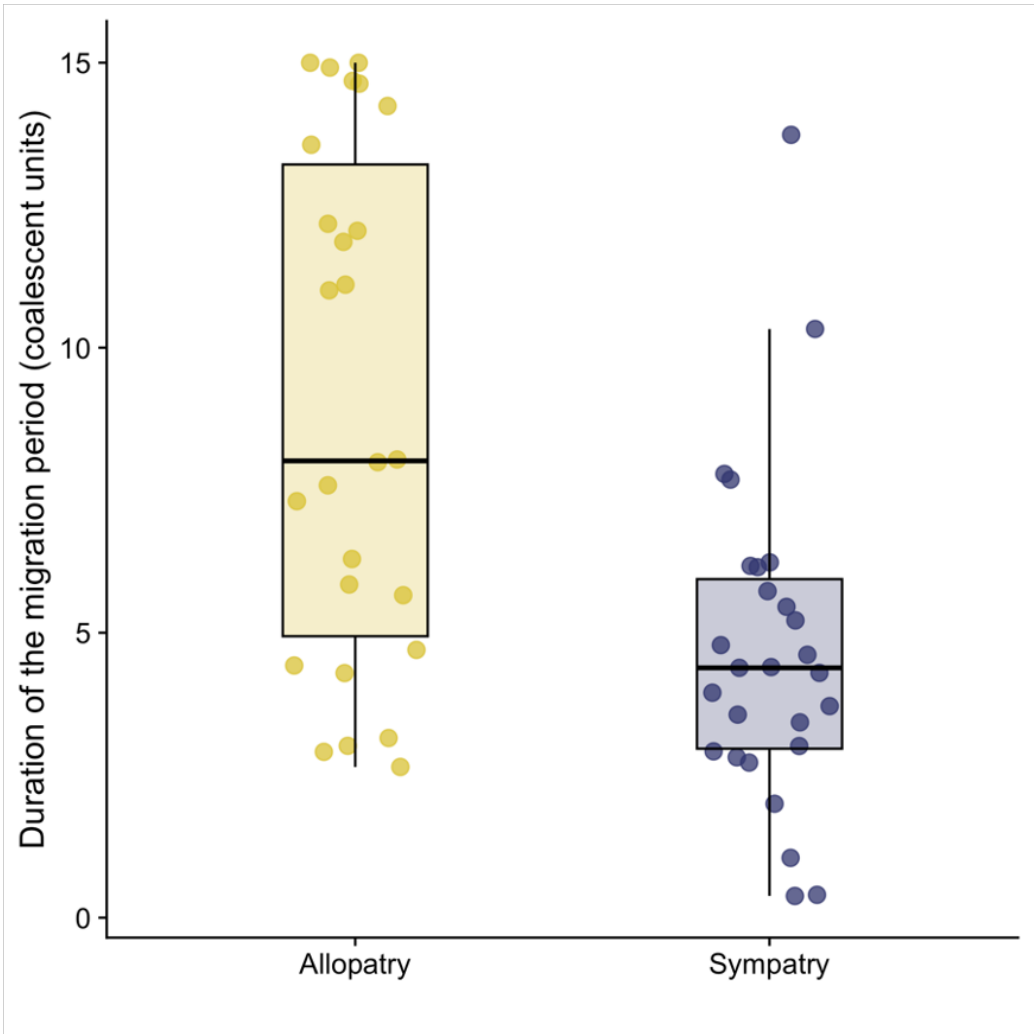
Supplementary Figure 6: Scaled effective population size estimates for pairs using best-fitting model parameter estimates. Parameter estimates were scaled using an across population *D. melanogaster* mutation rate of 3.32 x 10^{-9}, represented by the circles. Upper and lower bound of divergence time for each were scaled using the upper and lower bounds of the D. melanogaster mutation rate (upper: 2.52 x 10^{-9}; lower: 4.30 x 10^{-9}). Pairs are coloured by their best-fitting model.



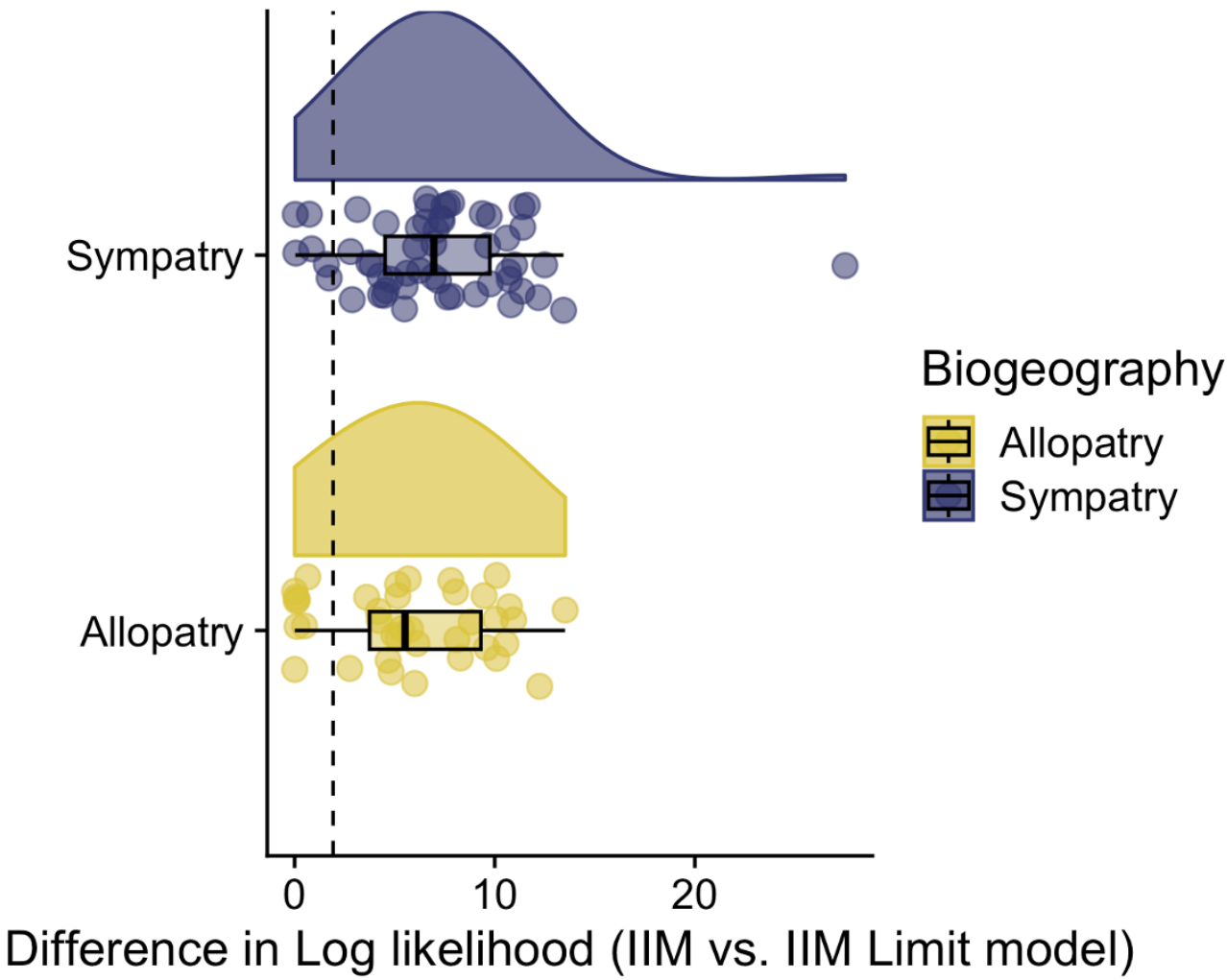
Supplementary Figure 7: Speciation models considered in the analysis. Models include a Strict Isolation (SI) model representing allopatric speciation without any gene flow (2 parameters), an isolation-with-migration (IM) model with continuous gene flow (3 parameters), an isolation-with-initial-migration model (IIM) which includes an initial phase of migration followed by no gene flow (4 parameters), and a secondary contact (SC) model which includes a recent pulse of migration (4 parameters).



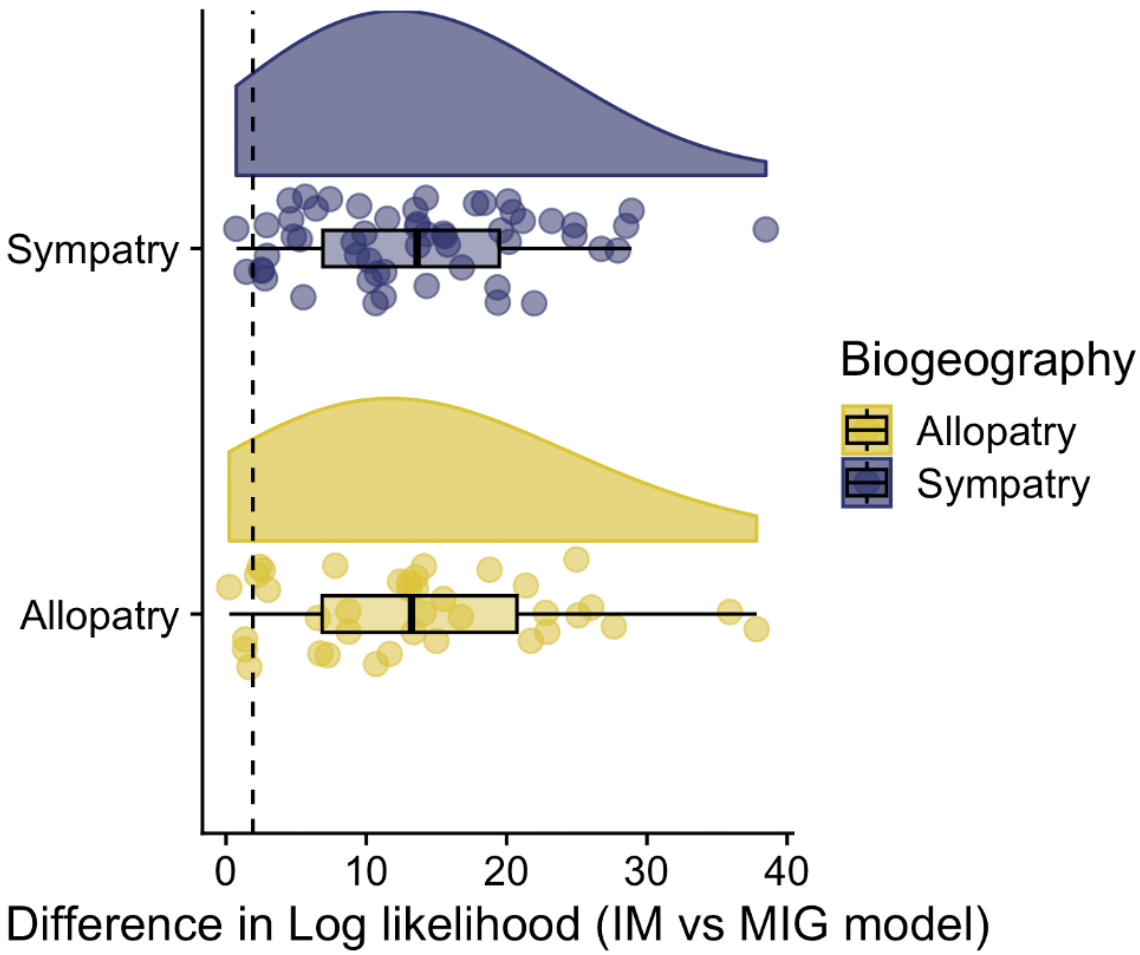
Supplementary Figure 8: Diagnostic speciation limit models. Models include a MIG model, which contains the same parameters as an IM model, but with no definable initial divergence time. In this model, gene flow is continuous and has no beginning or definable stopping point. The IIM Limit model contains the same parameters as an IIM model, but with no definable initial divergence time. For the IIM limit model, there is no identifiable start time to continuous gene flow, but a defined cessation time ($\tau 1$ parameter).



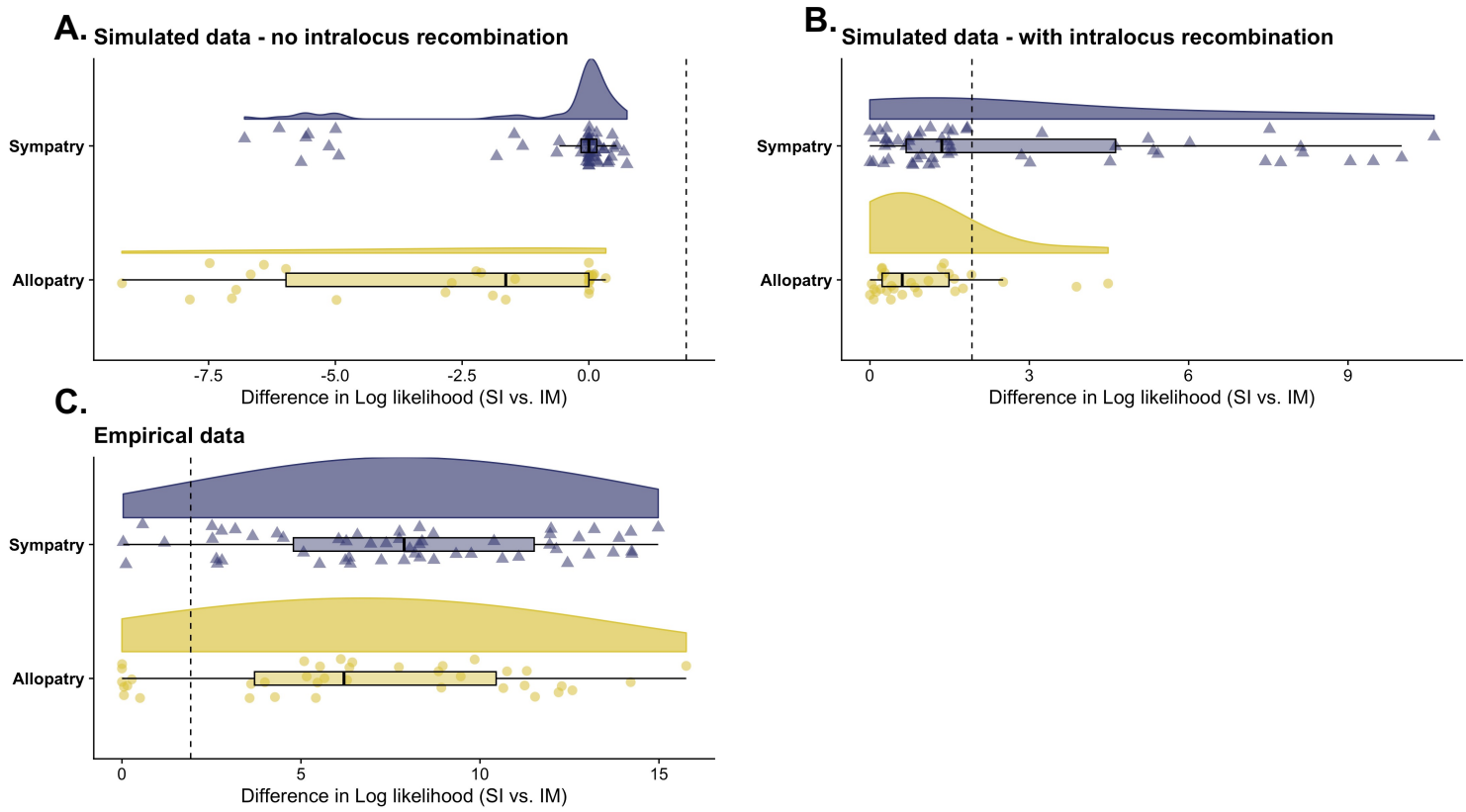
Supplementary Figure 9: Duration of migration for allopatric and sympatric pairs. Duration of migration ($\tau 0 - \tau 1$) was calculated for all pairs under the IIM model and is shown without being converted into scaled geological time estimates.



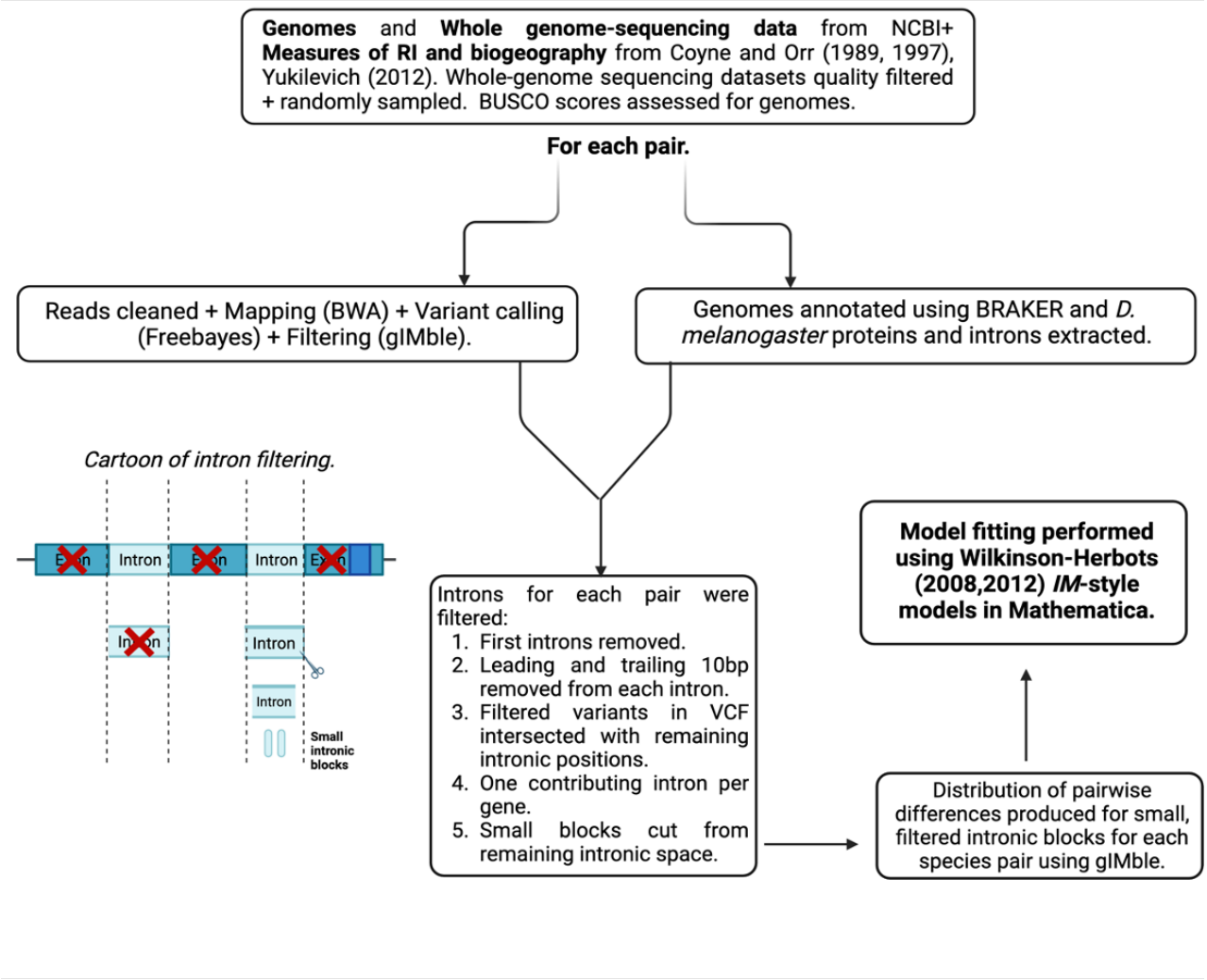
Supplementary Figure 10: Difference in log-likelihood between the isolation-with-initial-migration model (IIM) and the IIM Limit model, which contains the same parameters as the IIM model except an initial divergence time parameter (T0). In the IIM Limit model, there is no identifiable start to divergence, but there is a point at which gene flow ceases (T1), as with the IIM model. The dashed line is the critical threshold ($p < 0.05$). Points past the dashed line show support for the IIM model over the limit model. Most pairs show evidence for an initial start to speciation-with-gene-flow, and pairs that do not fit an IIM model better in this pairwise comparison show overall best support for strict isolation model (SI).



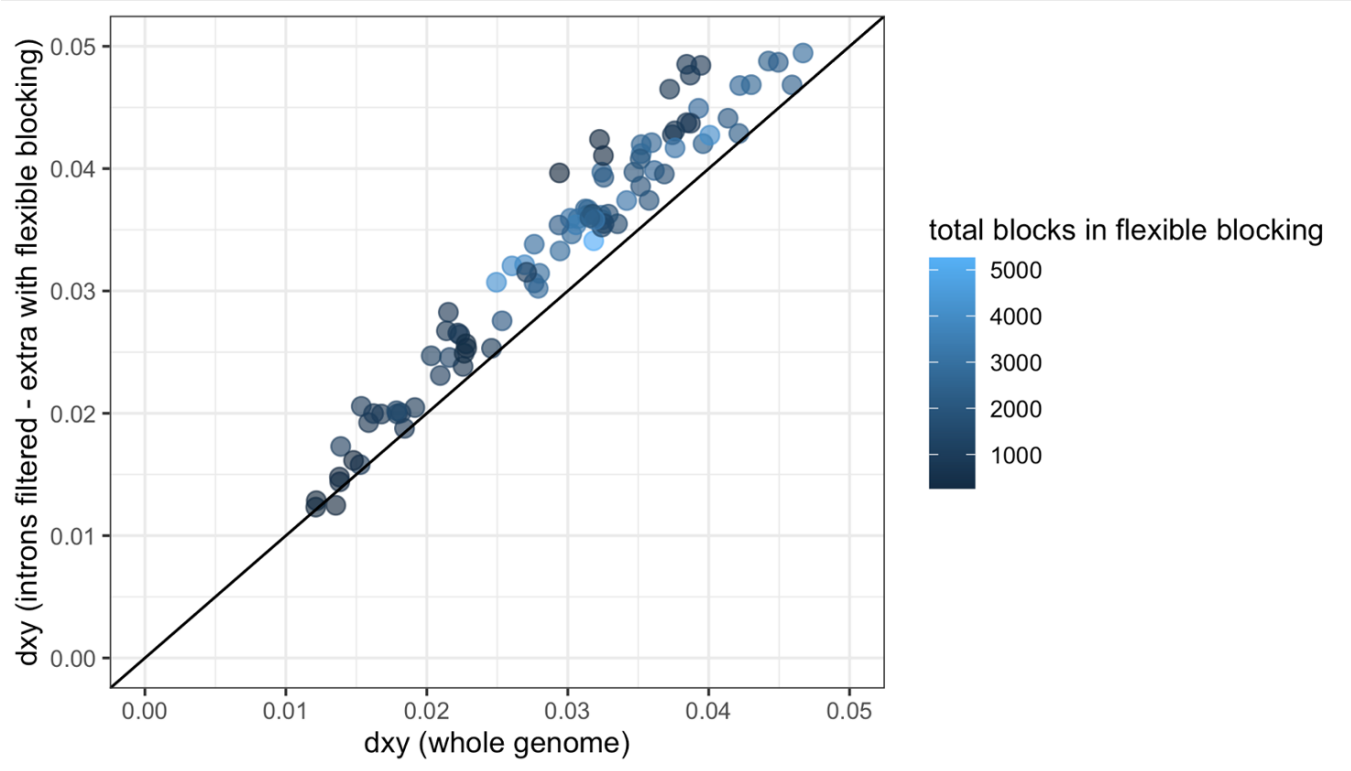
Supplementary Figure 11: Difference in log-likelihood between the isolation-with-migration model (IM) and the migration-only (MIG) model, which contains the same parameters as the IM model except an initial divergence time parameter (T0). In the MIG model, there is no clear identifiable start to divergence, such that there is only identifiable gene flow that is with no identifiable onset or cessation, and an effective population size parameter. The dashed line is the critical threshold ($p < 0.05$). Points past the dashed line show support for the IM model over the MIG model. Most pairs show evidence for an initial start to speciation-with-gene-flow, and pairs that do not fit an IM model better in this pairwise comparison show overall best support for strict isolation model (SI). Only one pair shows overall best support for a MIG model.



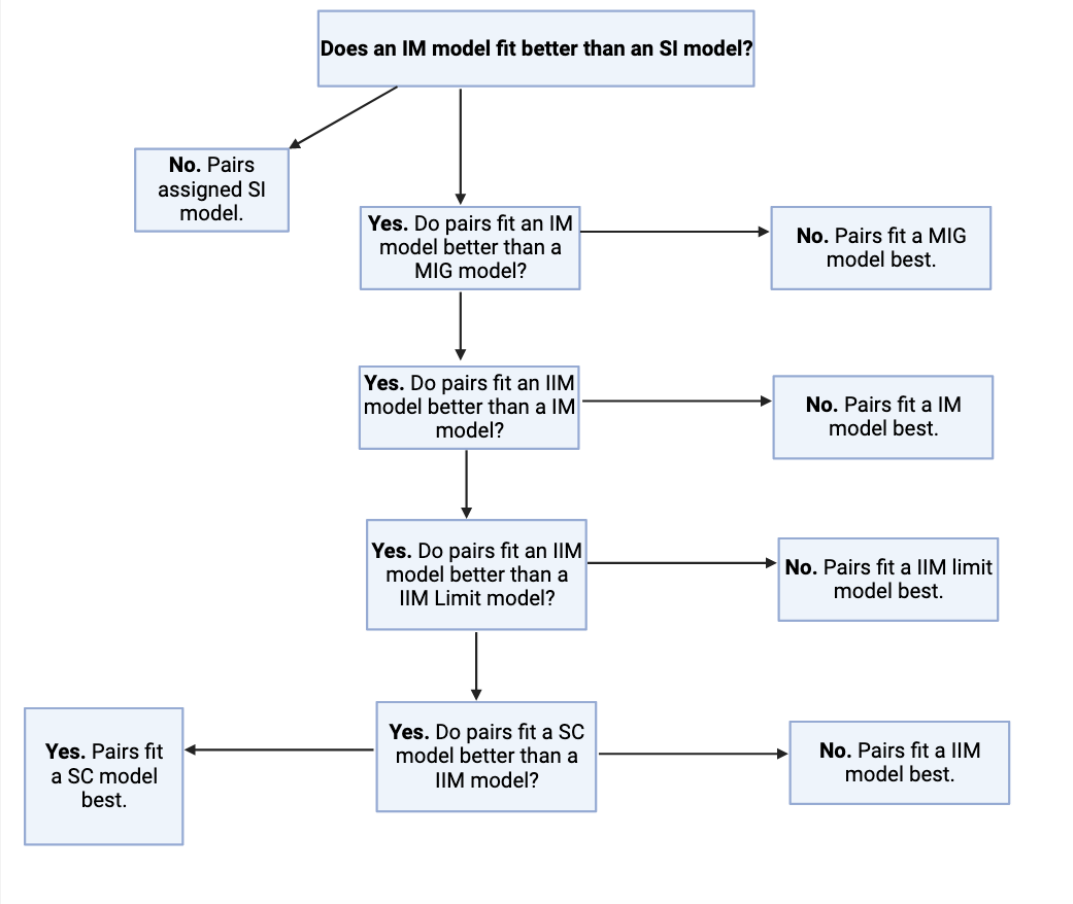
Supplementary Figure 12: Comparison of the difference in log-likelihood between a strict isolation (SI) model and an isolation-with-migration model (IM) between **A.** Data simulated according to optimized strict divergence history parameter estimates but with no recombination, **B.** Data simulated under an optimized strict divergence history with *D. melanogaster* levels of recombination, and **C.** empirical data. The dashed line indicates the critical threshold ($p < 0.05$). **A.** Simulations with no recombination show a false positive rate of detecting a history of speciation-with-gene-flow where there is none is 0%. B. Simulated data with realistic levels of recombination show a higher false positive rate of 25%. Comparison of **B.** and **C.** demonstrate that simulations with recombination do not recapitulate as strong support for a speciation history with gene flow across allopatric and sympatric pairs as the empirical data.



Supplementary Figure 13: Flowchart of analysis pipeline. Describing workflow for the analysis beginning with gathering data on reproductive isolation and biogeography for *Drosophila* pairs from previous comparative analyses, and whole-genome sequencing datasets and genomes from NCBI, and ending with model fitting for remaining pairs. Intron filtering cartoon alongside the description highlights main filtering decisions taken to produce the datasets used for the model fitting.



Supplementary Figure 14: The relationship between d_{XY} calculated in fixed 200bp blocks across the whole genome and d_{XY} calculated in filtered, intronic blocks across 93 species pairs. Points are coloured based on how many intronic blocks remained and were used in the model fitting analysis for each pair.



Supplementary Figure 15: Decision tree for assigning a particular model to a given pair of species in our analyses through a series of pairwise model comparisons.

Supplementary Tables

Table legends.

Supplementary Table 1: Species pairs and SRA accessions for whole-genome sequencing samples used in this study.

Supplementary Table 2: Species pairs and summary statistics calculated from filtered intronic data. Includes estimates of intronic absolute genetic divergence (D_{XY}) for each sample, and the number and size of filtered intronic blocks used in the analysis.

Supplementary Table 3: Species pairs and the genomes used to map whole-genome sequences to. Includes species' genome which was used for analysis, NCBI BioSample identification of the genome used for each pair and the publications where the data sequenced and assembled (where possible).

Supplementary Table 4: Species pairs and sample specific coverage information. Mean, standard deviation, minimum and maximum coverage are provided for each species in a pair-specific manner.

Supplementary Table 5: Information on reproductive isolation (pre- and postzygotic isolation), Nei's D based on allozymes and range overlap classifications for every pair considered in the analysis.

Supplementary Table 6: Difference in model support for five different model comparisons calculated for each pair. SI refers to the strict isolation model, IM to the isolation-with-migration model, IIM to the isolation-with-initial-migration model, SC to the secondary contact model, MIG to the migration-only model and IIM limit to the isolation-with-migration limit model.

Supplementary Table 7: For each pair, the difference in model support between SI model and IM model for simulated data with and without recombination.

Supplementary Table 8: Contingency tables showing the number of allopatric and sympatric pairs that best fit each model in pairwise model comparisons.

3.2 Tables: 1 - 8

Pair	Accession Species A	Accession Species B
D. affinis - D. athabasca	ERR127384	SRR9967640
D. ambigua - D. obscura	SRR13070667	SRR13070671
D. ambigua - D. persimilis	SRR12899691	SRR13070667
D. ambigua - D. pseudoobscura	SRR11230572	SRR13070667
D. ambigua - D. tristis	SRR13070667	SRR13070669
D. americana - D. lacicola	ERR2610695	SRR20226091
D. americana - D. littoralis	ERR2610695	SRR13070670
D. americana - D. montana	ERR2610695	SRR3218635
D. americana - D. novamexicana	ERR2610692	ERR2610695
D. americana - D. virilis	ERR2610695	SRR9426114
D. arizonae - D. mojavensis	SRR2070760	SRR6425997
D. auraria - D. rufa	SRR13070713	SRR13188937
D. auraria - D. triauraria	SRR13188922	SRR13188937
D. bipectinata - D. malerkotliana (pallens)	SRR13070716	SRR15048573
D. bipectinata - D. parabipectinata	SRR13070721	SRR15048573
D. bipectinata - D. pseudoananassae (nigrens)	SRR13070715	SRR15048573
D. borealis - D. lacicola	SRR20226101	SRR20226091
D. borealis - D. littoralis	SRR13070670	SRR20226101
D. borealis - D. montana	SRR3218635	SRR20226101
D. erecta - D. orena	SRR1977592	SRR6425990
D. flavomontana - D. montana	SRR3218635	SRR20226092
D. heteroneura - D. planitibia	SRR2233094	SRR2233332
D. heteroneura - D. silvestris	SRR2233094	SRR2233314
D. insularis - D. willistoni	SRR13070692	SRR6426003
D. jambulina - D. punjabiensis	SRR15048354	SRR9997923
D. kanekoi - D. ezoana	SRR20226095	SRR20226093

D. kanekoi - D. littoralis	SRR13070670	SRR20226093
D. kanekoi - D. virilis	SRR9426114	SRR20226093
D. kepulauna - D. albomicans	DRR076000	SRR10100373
D. kikkawai - D. leontia	SRR15048447	SRR1810494
D. kohkoa - D. albomicans	DRR076000	DRR160734
D. lacicola - D. montana	SRR3218635	SRR20226091
D. littoralis - D. ezoana	SRR13070670	SRR20226095
D. littoralis - D. montana	SRR13070670	SRR3218635
D. loweii - D. pseudoobscura	SRR11230572	SRR9967663
D. lummei - D. ezoana	SRR5278982	SRR20226095
D. lummei - D. littoralis	SRR13070670	SRR5278982
D. lummei - D. virilis	SRR5278982	SRR9426114
D. malerkotliana (pallens) - D. parabipectinata	SRR13070716	SRR13070721
D. malerkotliana (pallens) - D. pseudoananassae (nigrens)	SRR13070715	SRR13070716
D. mauritiana - D. erecta	SRR1560109	SRR6425990
D. mauritiana - D. orena	SRR1560109	SRR1977592
D. mauritiana - D. sechellia	SRR14138507	SRR1560109
D. mauritiana - D. teissieri	SRR1560109	SRR5860615
D. mauritiana - D. yacuba	SRR1560109	SRR9700025
D. melanogaster - D. erecta	SRR3931586	SRR6425990
D. melanogaster - D. orena	SRR1977592	SRR3931586
D. melanogaster - D. sechellia	SRR14138507	SRR3931586
D. melanogaster - D. simulans	SRR11456787	SRR3931586
D. melanogaster - D. teissieri	SRR3931586	SRR5860615
D. melanogaster - D. yacuba	SRR3931586	SRR9700025
D. miranda - D. bifasciata	DRR061004	SRR789669
D. miranda - D. persimilis	SRR12899691	SRR789669
D. miranda - D. pseudoobscura	SRR11230572	SRR789669
D. miranda - D. subobscura	SRR789669	SRR9967664

D. nasuta - D. albomicans	DRR076000	SRR10100355
D. novamexicana - D. lacicola	ERR2610692	SRR20226091
D. novamexicana - D. littoralis	ERR2610692	SRR13070670
D. novamexicana - D. montana	ERR2610692	SRR3218635
D. novamexicana - D. virilis	ERR2610692	SRR9426114
D. obscura - D. tristis	SRR13070669	SRR13070671
D. parabipectinata - D. pseudoananassae	SRR13070715	SRR13070721
D. persimilis - D. bifasciata	DRR061004	SRR12899691
D. persimilis - D. pseudoobscura	SRR11230572	SRR12899691
D. persimilis - D. subobscura	SRR12899691	SRR9967664
D. planitibia - D. silvestris	SRR2233314	SRR2233332
D. pseudoobscura - D. bifasciata	DRR061004	SRR11230572
D. pseudoobscura - D. subobscura	SRR11230572	SRR9967664
D. sechellia - D. erecta	SRR14138507	SRR6425990
D. sechellia - D. orena	SRR14138507	SRR1977592
D. sechellia - D. teissieri	SRR14138507	SRR5860615
D. sechellia - D. yacuba	SRR14138507	SRR9700025
D. simulans - D. erecta	SRR11456787	SRR6425990
D. simulans - D. orena	SRR11456787	SRR1977592
D. simulans - D. sechellia	SRR11456787	SRR14138507
D. simulans -D. teissieri	SRR11456787	SRR5860615
D. simulans - D. yacuba	SRR11456787	SRR9700025
D. subobscura - D. ambigua	SRR13070667	SRR9967664
D. subobscura - D. bifasciata	DRR061004	SRR9967664
D. subobscura - D. obscura	SRR13070671	SRR9967664
D. subobscura - D. tristis	SRR13070669	SRR9967664
D. sulfurigaster albostrigata - D. kepulauna	SRR10100373	SRR10100407
D. sulfurigaster albostrigata - D. pulaua	SRR10100335	SRR10100407

D. sulfurigaster albostrigata - D. sulfurigaster bilimbata	SRR10100308	SRR10100323
D. sulfurigaster albostrigata - D. sulfurigaster neonasuta	SRR10100407	SRR5059342
D. sulfurigaster albostrigata - D. sulfurigaster sulfurigaster	SRR10100323	SRR10100401
D. sulfurigaster bilimbata - pulaua	SRR10100308	SRR10100335
D. sulfurigaster bilimbata - sulfurigaster neonasuta	SRR10100308	SRR5059342
D. sulfurigaster bilimbata - D. sulfurigaster sulfurigaster	SRR10100308	SRR10100401
D. sulfurigaster neonasuta - pulaua	SRR10100335	SRR5059342
D. sulfurigaster neonasuta - D. sulfurigaster sulfurigaster	SRR10100401	SRR5059342
D. sulfurigaster sulfurigaster - D. pulaua	SRR10100335	SRR10100401
D. teissieri - D. erecta	SRR5860615	SRR6425990
D. teissieri - D. orena	SRR1977592	SRR5860615
D. triauraria - D. rufa	SRR13070713	SRR13188922
D. virilis - D. ezoana	SRR9426114	SRR20226095
D. virilis - D. lacicola	SRR9426114	SRR20226091
D. virilis - D. littoralis	SRR13070670	SRR9426114
D. yacuba - D. erecta	SRR6425990	SRR9700025
D. yacuba - D. orena	SRR1977592	SRR9700025
D. yacuba - D. santomea	SRR5860642	SRR9700025
D. yacuba - D. teissieri	SRR5860615	SRR9700025

Pair	**Intronic D_{XY}**	**No. of blocks**	**Block size**
D. affinis - D. athabasca	0.028	1300	137
D. ambigua - D. obscura	0.0119	2460	99
D. ambigua - D. persimilis	0.0128	738	70
D. ambigua - D. pseudoobscura	0.0126	725	70
D. americana - D. lacicola	0.0128	2130	82
D. americana - D. littoralis	0.0132	2230	84
D. americana - D. montana	0.0115	2140	82
D. americana - D. virilis	0.0129	2530	84
D. arizonae - D. mojavensis	0.0145	3120	121
D. auraria- D. rufa	0.0132	2080	83
D. bipectinata - D. malerkotliana	0.0108	2670	157
D. bipectinata - D. parabipectinata	0.0116	2660	165
D. bipectinata - D. pseudoananassae	0.0117	1920	116
D. borealis - D. lacicola	0.0133	3270	94
D. borealis - D. littoralis	0.0139	2070	72
D. borealis - D. montana	0.01	3280	93
D. erecta - D. orena	0.0143	2440	95
D. flavomontana- D. montana	0.0103	3460	98
D. insularis - D. willistoni	0.0147	1720	80
D. jambulina - D. punjabiensis	0.0122	2590	109
D. kanekoi - D. ezoana	0.0138	2450	84
D. kanekoi - D. littoralis	0.0144	2160	75
D. kanekoi - D. lummei	0.012	2230	81
D. kanekoi - D. virilis	0.0141	1890	67

D. kepulauna - D. albomicans	0.0125	3510	185
D. kikkawai- D. leontia	0.00865	3010	205
D. kohkoa - D. albomicans	0.0128	3790	111
D. littoralis - D. ezoana	0.0131	2480	89
D. littoralis - D. montana	0.0118	2160	76
D. loweii - D. pseudoobscura	0.0135	2150	98
D. lummei - D. ezoana	0.0125	2450	83
D. lummei - D. littoralis	0.0125	2350	87
D. lummei - D. virilis	0.0107	2680	117
D. malerkotliana - D. parabipectinata	0.0139	2950	159
D. malerkotliana - D. pseudoananassae	0.0123	1950	116
D. mauritiana - D. erecta	0.0138	1580	82
D. mauritiana - D. orena	0.0143	1590	78
D. mauritiana - D. sechellia	0.0148	3140	155
D. mauritiana - D. teissieri	0.0139	1930	72
D. mauritiana - D. yacuba	0.0132	1400	83
D. melanogaster - D. erecta	0.0141	1430	85
D. melanogaster - D. orena	0.0135	3990	73
D. melanogaster - D. sechellia	0.0148	6390	88
D. melanogaster - D. simulans	0.011	2010	118
D. melanogaster - D. teissieri	0.0123	1300	86
D. melanogaster - D. yacuba	0.0126	1180	85
D. miranda - D. persimilis	0.0149	2670	155
D. miranda - D. pseudoobscura	0.0146	2560	146
D. nasuta - D. albomicans	0.0143	3350	218
D. novamexicana - D. lacicola	0.0142	2430	72
D. novamexicana - D. littoralis	0.0144	2560	81
D. novamexicana - D. montana	0.0128	2620	84
D. novamexicana - D. virilis	0.0146	2810	85

D. obscura - D. tristis	0.0123	2530	102
D. parabipectinata - D. pseudoananassae	0.0139	2070	120
D. persimilis - D. bifasciata	0.0133	682	69
D. persimilis - D. pseudoobscura	0.0143	2700	228
D. persimilis - D. subobscura	0.0124	377	72
D. pseudoobscura - D. bifasciata	0.0131	703	68
D. pseudoobscura - D. subobscura	0.0123	367	76
D. sechellia - D. erecta	0.0144	1800	60
D. sechellia - D. orena	0.0147	4900	70
D. sechellia - D. teissieri	0.0133	1660	61
D. sechellia - D. yacuba	0.0139	1600	61
D. simulans - D. erecta	0.0134	1680	64
D. simulans - D. orena	0.0123	1470	75
D. simulans - D. teissieri	0.0119	1710	68
D. simulans - D. yacuba	0.013	1730	70
D. subobscura - D. ambigua	0.0122	673	63
D. subobscura - D. bifasciata	0.0125	669	62
D. subobscura - D. obscura	0.0124	658	64
D. subobscura - D. tristis	0.0118	643	62
D. sulfurigaster albostrigata - D. kepulauna	0.0106	2670	188
D. sulfurigaster albostrigata - D. pulaua	0.0102	1460	291
D. sulfurigaster albostrigata - D. sulfurigaster bilimbata	0.0125	2980	212
D. sulfurigaster albostrigata - D. sulfurigaster neonasuta	0.0111	2920	186
D. sulfurigaster albostrigata - D. sulfurigaster sulfurigaster	0.0126	2950	217
D. sulfurigaster bilimbata - D. pulaua	0.0131	1750	314

D. sulfurigaster bilimbata - D. sulfurigaster neonasuta	0.0137	3300	120
D. sulfurigaster bilimbata - D. sulfurigaster sulfurigaster	0.0117	1710	341
D. sulfurigaster neonasuta - D. pulaua	0.0137	3340	121
D. sulfurigaster neonasuta - D. sulfurigaster sulfurigaster	0.0138	3400	122
D. sulfurigaster sulfurigaster - D. pulaua	0.0125	1620	325
D. teissieri- D. erecta	0.0134	1850	64
D. teissieri - D. orena	0.0125	1640	75
D. triauraria - D. rufa	0.0126	1830	83
D. virilis - D. ezoana	0.0133	1930	85
D. virilis - D. lacicola	0.0142	1800	77
D. virilis - D. littoralis	0.014	1920	72
D. yacuba - D. erecta	0.014	1800	64
D. yacuba - D. orena	0.0136	1590	74
D. yacuba - D. santomea	0.0118	3050	147
D. yacuba - D. teissieri	0.0125	2440	90

Pair	***Species (Genome)***	**Source**	**Reference**
D. affinis - D. athabasca	*D. athabasca*	SAMN12214156	Bracewell, R., Chatla, K., Nalley, M.J. and Bachtrog, D., 2019. Dynamic turnover of centromeres drives karyotype evolution in Drosophila. *Elife*, *8*, p.e49002.
D. ambigua - D. obscura	*D. obscura*	SAMN16729671	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. ambigua - D. persimilis	*D. persimilis*	SAMN16729701	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. ambigua - D. pseudoobscura	*D.pseudoobscura*	SAMN16729702	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. americana - D. lacicola	*D. lacicola*	SAMN29503225	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. americana - D. littoralis	*D. littoralis*	SAMN29503227	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. americana - D. montana	*D. montana*	SAMN29503231	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. americana - D. virilis	*D. virilis*	SAMN16729705	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. arizonae - D. mojavensis	*D. mojavaensis*	SAMN16729700	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. auraria- D. rufa	*D. auraria*	SAMN12263403	Bronski, M.J., Martinez, C.C., Weld, H.A. and Eisen, M.B., 2020. Whole genome sequences of 23 species from the Drosophila montium species group (Diptera: Drosophilidae): a resource for testing evolutionary hypotheses. *G3: Genes, Genomes, Genetics*, *10*(5), pp.1443-1455.
D. bipectinata - D. malerkotliana	*D. bipectinata*	SAMN16729609	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. bipectinata - D. parabipectinata	*D. bipectinata*	SAMN16729609	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. bipectinata - D. pseudoananassae	*D. bipectinata*	SAMN16729609	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. borealis - D. lacicola	*D. lacicola*	SAMN29503225	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. borealis - D. littoralis	*D. littoralis*	SAMN29503227	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. borealis - D. montana	*D. montana*	SAMN29503231	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. erecta - D. orena	*D. erecta*	SAMN16729697	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. flavomontana-D. montana	*D. montana*	SAMN29503231	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.

D. insularis - D. willistoni	*D. willistoni*	SAMN16729706	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. jambulina - D. punjabiensis	*D. punjabiensis*	SAMN12263408	Bronski, M.J., Martinez, C.C., Weld, H.A. and Eisen, M.B., 2020. Whole genome sequences of 23 species from the Drosophila montium species group (Diptera: Drosophilidae): a resource for testing evolutionary hypotheses. *G3: Genes, Genomes, Genetics*, *10*(5), pp.1443-1455.
D. kanekoi - D. ezoana	*D. ezoana*	SAMN29503221	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. kanekoi - D. littoralis	*D. littoralis*	SAMN29503227	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. kanekoi - D. lummei	*D. lummei*	SAMN29503229	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.

D. kanekoi - D. virilis	*D. virilis*	SAMN16729705	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. kepulauna - D. albomicans	*D. albomicans*	SAMN12703415 (legacy)	Mai, D., Nalley, M.J. and Bachtrog, D., 2020. Patterns of genomic differentiation in the Drosophila nasuta species complex. *Molecular biology and evolution*, *37*(1), pp.208-220; Wei, K.H.C., Mai, D., Chatla, K. and Bachtrog, D., 2022. Dynamics and impacts of transposable element proliferation in the Drosophila nasuta species group radiation. Molecular Biology and Evolution, 39(5), p.msac080.
D. kikkawai- D. leontia	*D. kikkawai*	SAMN16729629	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405; Suvorov, A., Kim, B.Y., Wang, J., Armstrong, E.E., Peede, D., D'agostino, E.R., Price, D.K., Waddell, P.J., Lang, M., Courtier-Orgogozo, V. and David, J.R., 2022. Widespread introgression across a phylogeny of 155 Drosophila genomes. Current Biology, 32(1), pp.111-123.

D. kohkoa - D. albomicans	*D. albomicans*	SAMN12703415 (legacy)	Mai, D., Nalley, M.J. and Bachtrog, D., 2020. Patterns of genomic differentiation in the Drosophila nasuta species complex. *Molecular biology and evolution*, *37*(1), pp.208-220; Wei, K.H.C., Mai, D., Chatla, K. and Bachtrog, D., 2022. Dynamics and impacts of transposable element proliferation in the Drosophila nasuta species group radiation. Molecular Biology and Evolution, 39(5), p.msac080.
D. littoralis - D. ezoana	*D. ezoana*	SAMN29503221	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. littoralis - D. montana	*D. montana*	SAMN29503231	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. loweii - D. pseudoobscura	*D. pseudoobscura*	SAMN16729702	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. lummei - D. ezoana	*D. lummei*	SAMN29503229	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. lummei - D. littoralis	*D. lummei*	SAMN29503229	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. lummei - D. virilis	*D. lummei*	SAMN29503229	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. malerkotliana - D. parabipectinata	*D. parabipectinata*	SAMN16729611	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. malerkotliana - D. pseudoananassae	*D. malerkotliana*	SAMN16729617	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. mauritiana - D. erecta	*D. erecta*	SAMN16729697	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. mauritiana - D. orena	*D. mauritiana*	SAMN06827991	No specific publication; see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_004382145.1/
D. mauritiana - D. sechellia	*D. mauritiana*	SAMN06827991	No specific publication; see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_004382145.1/
D. mauritiana - D. teissieri	*D. mauritiana*	SAMN06827991	No specific publication; see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_004382145.1/
D. mauritiana - D. yacuba	*D. yacuba*	SAMN16729707	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. melanogaster - D. erecta	*D. erecta*	SAMN16729697	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. melanogaster - D. orena	*D. melanogaster*	dmel_r6.39_FB2021_02/	Flybase; v6.39.
D. melanogaster - D. sechellia	*D. melanogaster*	dmel_r6.39_FB2021_02/	Flybase; v6.39.
D. melanogaster - D. simulans	*D. simulans*	SAMN16729704	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. melanogaster - D. teissieri	*D. teissieri*	SAMN16729662	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405; Suvorov, A., Kim, B.Y., Wang, J., Armstrong, E.E., Peede, D., D'agostino, E.R., Price, D.K., Waddell, P.J., Lang, M., Courtier-Orgogozo, V. and David, J.R., 2022. Widespread introgression across a phylogeny of 155 Drosophila genomes. Current Biology, 32(1), pp.111-123.

D. melanogaster - D. yacuba	*D. yacuba*	SAMN16729707	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. miranda - D. persimilis	*D. persimilis*	SAMN16729701	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. miranda - D. pseudoobscura	*D.pseudoobscura*	SAMN16729702	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. nasuta - D. albomicans	*D. albomicans*	SAMN12703415 (legacy)	Mai, D., Nalley, M.J. and Bachtrog, D., 2020. Patterns of genomic differentiation in the Drosophila nasuta species complex. *Molecular biology and evolution*, *37*(1), pp.208-220; Wei, K.H.C., Mai, D., Chatla, K. and Bachtrog, D., 2022. Dynamics and impacts of transposable element proliferation in the Drosophila nasuta species group radiation. Molecular Biology and Evolution, 39(5), p.msac080.

D. novamexicana - D. lacicola	*D. novamexicana*	SAMN09383009	No specific publication; see https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_003285875.2/
D. novamexicana - D. littoralis	*D. novamexicana*	SAMN09383009	No specific publication; see https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_003285875.2/
D. novamexicana - D. montana	*D. novamexicana*	SAMN09383009	No specific publication; see https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_003285875.2/
D. novamexicana - D. virilis	*D. novamexicana*	SAMN09383009	No specific publication; see https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_003285875.2/
D. obscura - D. tristis	*D. obscura*	SAMN16729671	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. parabipectinata - D. pseudoananassae	*D. parabipectinata*	SAMN16729611	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. persimilis - D. bifasciata	*D. persimilis*	SAMN16729701	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. persimilis - D. pseudoobscura	*D. persimilis*	SAMN16729701	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. persimilis - D. subobscura	*D. persimilis*	SAMN16729701	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. pseudoobscura - D. bifasciata	*D.pseudoobscura*	SAMN16729702	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. pseudoobscura - D. subobscura	*D.pseudoobscura*	SAMN16729702	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. sechellia - D. erecta	*D. erecta*	SAMN16729697	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. sechellia - D. orena	*D. melanogaster*	dmel_r6.39_FB2021_02/	Flybase; v6.39.
D. sechellia - D. teissieri	*D. teissieri*	SAMN16729662	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405; Suvorov, A., Kim, B.Y., Wang, J., Armstrong, E.E., Peede, D., D'agostino, E.R., Price, D.K., Waddell, P.J., Lang, M., Courtier-Orgogozo, V. and David, J.R., 2022. Widespread introgression across a phylogeny of 155 Drosophila genomes. Current Biology, 32(1), pp.111-123.

D. sechellia - D. yacuba	*D. yacuba*	SAMN16729707	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. simulans - D. erecta	*D. simulans*	SAMN16729704	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. simulans - D. orena	*D. simulans*	SAMN16729704	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. simulans - D. teissieri	*D. simulans*	SAMN16729704	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. simulans - D. yacuba	*D. simulans*	SAMN16729704	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. subobscura - D. ambigua	*D. subobscura*	SAMN16729670	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. subobscura - D. bifasciata	*D. subobscura*	SAMN16729670	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
SAMN16729617	*D. subobscura*	SAMN16729670	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. subobscura - D. tristis	*D. subobscura*	SAMN16729670	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. sulfurigaster albostrigata - D. kepulauna	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. sulfurigaster albostrigata - D. pulaua	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. sulfurigaster albostrigata - D. sulfurigaster bilimbata	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. sulfurigaster albostrigata - D. sulfurigaster neonasuta	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/

D. sulfurigaster albostrigata - D. sulfurigaster sulfurigaster	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. sulfurigaster bilimbata - D. pulaua	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. sulfurigaster bilimbata - D. sulfurigaster neonasuta	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. sulfurigaster bilimbata - D. sulfurigaster sulfurigaster	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. sulfurigaster neonasuta - D. pulaua	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/

D. sulfurigaster neonasuta - D. sulfurigaster sulfurigaster	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. sulfurigaster sulfurigaster - D. pulaua	*D. neonasuta*	SAMN06052360	see https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_005889595.1/
D. teissieri- D. erecta	*D. teissieri*	SAMN16729662	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405; Suvorov, A., Kim, B.Y., Wang, J., Armstrong, E.E., Peede, D., D'agostino, E.R., Price, D.K., Waddell, P.J., Lang, M., Courtier-Orgogozo, V. and David, J.R., 2022. Widespread introgression across a phylogeny of 155 Drosophila genomes. Current Biology, 32(1), pp.111-123.

D. teissieri - D. orena	*D. teissieri*	SAMN16729662	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405; Suvorov, A., Kim, B.Y., Wang, J., Armstrong, E.E., Peede, D., D'agostino, E.R., Price, D.K., Waddell, P.J., Lang, M., Courtier-Orgogozo, V. and David, J.R., 2022. Widespread introgression across a phylogeny of 155 Drosophila genomes. Current Biology, 32(1), pp.111-123.
D. triauraria - D. rufa	*D. triauraria*	SAMN16744625	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. virilis - D. ezoana	*D. ezoana*	SAMN29503221	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.
D. virilis - D. lacicola	*D. lacicola*	SAMN29503225	Yusuf, L.H., Tyukmaeva, V., Hoikkala, A. and Ritchie, M.G., 2022. Divergence and introgression among the virilis group of Drosophila. *Evolution Letters*, *6*(6), pp.537-551.

D. virilis - D. littoralis	*D. virilis*	SAMN16729705	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. yacuba - D. erecta	*D. yacuba*	SAMN16729707	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. yacuba - D. orena	*D. yacuba*	SAMN16729707	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.
D. yacuba - D. santomea	*D. yacuba*	SAMN16729707	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

D. yacuba - D. teissieri	*D. yacuba*	SAMN16729707	Kim, B.Y., Wang, J.R., Miller, D.E., Barmina, O., Delaney, E., Thompson, A., Comeault, A.A., Peede, D., D'Agostino, E.R., Pelaez, J. and Aguilar, J.M., 2021. Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, *10*, p.e66405.

Pairs	Sample	Coverage (mean)	Coverage (SD)	Coverage (min)	Coverage (max)
D. affinis - D. athabasca	SRR9967640	13.35	7.86	2	29
D. affinis - D. athabasca	ERR127384	16.24	16.86	2	49
D. ambigua - D. obscura	SRR13070667	18.55	15.92	2	50
D. ambigua - D. obscura	SRR13070671	16.3	10.81	2	37
D. ambigua - D. persimilis	SRR12899691	24.45	19.76	2	63
D. ambigua - D. persimilis	SRR13070667	16.89	20.92	2	58
D. ambigua - D. pseudoobscura	SRR11230572	25.82	14.29	2	54
D. ambigua - D. pseudoobscura	SRR13070667	16.93	22.35	2	61
D. ambigua - D. pseudoobscura	SRR13070669	15.01	30.76	2	76
D. ambigua - D. pseudoobscura	SRR13070667	14.68	25.11	2	64
D. americana - D. lacicola	ERR2610695	16.82	44.13	2	105
D. americana - D. lacicola	SRR20226091	22	8.95	2	39
D. americana - D. littoralis	ERR2610695	16.71	22.26	2	61

D. americana - D. littoralis	SRR13070670	16.28	33.6	2	83
D. americana - D. montana	SRR3218635	28.99	56.14	2	141
D. americana - D. montana	ERR2610695	16.66	209.99	2	436
D. americana - D. novamexicana	ERR2610692	19.69	61.5	2	142
D. americana - D. novamexicana	ERR2610695	19.83	48.49	2	116
D. americana - D. virilis	SRR9426114	9.94	8.52	2	26
D. americana - D. virilis	ERR2610695	18.23	47.92	2	114
D. arizonae - D. mojavensis	SRR2070760	10.93	10.28	2	31
D. arizonae - D. mojavensis	SRR6425997	30.01	15.74	2	61
D. auraria - D. rufa	SRR13188937	11.34	10.05	2	31
D. auraria - D. rufa	SRR13070713	14.91	16.77	2	48
D. auraria - D. triauraria	SRR13188937	11.65	9.47	2	30
D. auraria - D. triauraria	SRR13188922	16.05	12.22	2	40
D. bipectinata - D. malerkotliana (pallens)	SRR15048573	9.74	8.91	2	27

D. bipectinata - D. malerkotliana (pallens)	SRR13070716	19.17	14.54	2	48
D. bipectinata - D. parabipectinata	SRR13070721	12.33	12.51	2	37
D. bipectinata - D. parabipectinata	SRR15048573	9.74	8.91	2	27
D. bipectinata - D. pseudoananassae (nigrens)	SRR15048573	9.74	8.91	2	27
D. bipectinata - D. pseudoananassae (nigrens)	SRR13070715	11.19	17.17	2	45
D. borealis - D. lacicola	SRR13070670	21.49	18.57	2	58
D. borealis - D. lacicola	SRR20226091	22	8.95	2	39
D. borealis - D. littoralis	SRR13070670	16.28	33.6	2	83
D. borealis - D. littoralis	SRR13070670	20.35	17.55	2	55
D. borealis - D. montana	SRR3218635	28.99	56.13	2	141
D. borealis - D. montana	SRR13070670	20.76	14.82	2	50
D. erecta - D. orena	SRR1977592	9.8	8.49	2	26
D. erecta - D. orena	SRR6425990	40.14	17.62	2	75

D. flavomontana - D. montana	SRR3218635	28.99	56.15	2	141
D. flavomontana - D. montana	SRR20226092	21.67	87.7	2	197
D. heteroneura - D. planitibia	SRR2233094	17.54	12.91	2	43
D. heteroneura - D. planitibia	SRR2233332	29.52	202.99	2	435
D. heteroneura - D. silvestris	SRR2233094	17.54	12.91	2	43
D. heteroneura - D. silvestris	SRR2233314	18.35	12.57	2	43
D. insularis - D. willistoni	SRR6426003	23.82	19.33	2	62
D. insularis - D. willistoni	SRR13070692	31.88	37.87	2	107
D. jambulina - D. punjabiensis	SRR9997923	18.72	9.44	2	37
D. jambulina - D. punjabiensis	SRR15048354	7	23.94	2	54
D. kanekoi - D. ezoana	SRR20226095	22.75	10.2	2	43
D. kanekoi - D. ezoana	SRR20226093	19.5	461.55	2	942
D. kanekoi - D. littoralis	SRR13070670	16.28	33.6	2	83
D. kanekoi - D. littoralis	SRR20226093	20.22	24.8	2	69

D. kanekoi - D. virilis	SRR20226093	19.6	282.26	2	584
D. kanekoi - D. virilis	SRR5278982	16.68	12.33	2	41
kanekoi_virilis	SRR9426114	9.94	8.52	2	26
kanekoi_virilis	SRR20226093	20.62	12.68	2	45
D. kepulauna - D. albomicans	DRR076000	82.24	454.35	2	990
D. kepulauna - D. albomicans	SRR10100373	9.19	5.66	2	20
D. kikkawai - D. leontia	SRR1810494	9.04	8.68	2	26
D. kikkawai - D. leontia	SRR15048447	9.23	10.13	2	29
D. kohkoa - D. albomicans	DRR160734	49.97	68.1	2	186
D. kohkoa - D. albomicans	DRR076000	82.24	454.36	2	990
D. lacicola - D. montana	SRR3218635	28.99	56.14	2	141
D. lacicola - D. montana	SRR20226091	20.65	10.58	2	41
D. littoralis - D. ezoana	SRR20226095	22.75	10.2	2	43
D. littoralis - D. ezoana	SRR13070670	15.09	58.14	2	131
D. littoralis - D. montana	SRR13070670	15.21	36.84	2	88
D. littoralis - D. montana	SRR3218635	28.99	56.13	2	141

D. loweii - D. pseudoobscura	SRR11230572	25.82	14.29	2	54
D. loweii - D. pseudoobscura	SRR9967663	12.31	18.12	2	48
D. lummei - D. ezoana	SRR20226095	21.72	19.62	2	60
D. lummei - D. ezoana	SRR5278982	16.68	12.33	2	41
D. lummei - D. littoralis	SRR13070670	15.58	34.12	2	83
D. lummei - D. littoralis	SRR5278982	16.68	12.33	2	41
D. lummei - D. virilis	SRR9426114	8.86	8.21	2	25
D. lummei - D. virilis	SRR5278982	16.68	12.33	2	41
D. malerkotliana (pallens) - D. parabipectinata	SRR13070721	11.7	8.39	2	28
D. malerkotliana (pallens) - D. parabipectinata	SRR13070716	18.9	13.71	2	46
D. malerkotliana (pallens) - D. pseudoananassae (nigrens)	SRR13070716	19.54	11.85	2	43
D. malerkotliana (pallens) - D. pseudoananassae (nigrens)	SRR13070715	11.3	15.78	2	42

D. mauritiana - D. erecta	SRR1560109	12.82	8.14	2	29
D. mauritiana - D. erecta	SRR6425990	40.14	17.62	2	75
D. mauritiana - D. orena	SRR1560109	15.23	7.23	2	29
D. mauritiana - D. orena	SRR1977592	8.85	7.3	2	23
D. mauritiana - D. sechellia	SRR1560109	15.23	7.23	2	29
D. mauritiana - D. sechellia	SRR14138507	22.38	13.35	2	49
D. mauritiana - D. teissieri	SRR5860615	11.89	12.18	2	36
D. mauritiana - D. teissieri	SRR1560109	15.23	7.23	2	29
D. mauritiana - D. yacuba	SRR1560109	12.79	23.88	2	60
D. mauritiana - D. yacuba	SRR9700025	16.48	7.73	2	31
D. melanogaster - D. erecta	SRR3931586	13.19	132.76	2	278
D. melanogaster - D. erecta	SRR6425990	40.14	17.62	2	75
D. melanogaster - D. orena	SRR1977592	8.98	7.93	2	24
D. melanogaster - D. orena	SRR3931586	12.21	12.3	2	36

D. melanogaster - D. sechellia	SRR14138507	21.87	34.25	2	90
D. melanogaster - D. sechellia	SRR3931586	12.21	12.3	2	36
D. melanogaster - D. simulans	SRR11456787	26.86	12.34	2	51
D. melanogaster - D. simulans	SRR3931586	13.15	13.95	2	41
D. melanogaster - D. teissieri	SRR5860615	12.82	7.06	2	26
D. melanogaster - D. teissieri	SRR3931586	12.83	89.37	2	191
D. melanogaster - D. yacuba	SRR9700025	16.48	7.73	2	31
D. melanogaster - D. yacuba	SRR3931586	12.84	78.35	2	169
D. miranda - D. bifasciata	SRR789669	9.27	10.01	2	29
D. miranda - D. bifasciata	DRR061004	46.87	24.65	2	96
D. miranda - D. persimilis	SRR12899691	24.45	19.76	2	63
D. miranda - D. persimilis	SRR789669	13.26	11	2	35
D. miranda - D. pseudoobscura	SRR11230572	25.82	14.29	2	54
D. miranda - D. pseudoobscura	SRR789669	13.23	11.16	2	35

D. miranda - D. subobscura	SRR789669	9.3	8.52	2	26
D. miranda - D. subobscura	SRR9967664	14.94	8.94	2	32
D. nasuta - D. albomicans	SRR10100355	11.68	6.74	2	25
D. nasuta - D. albomicans	DRR076000	82.24	454.36	2	990
D. novamexicana - D. lacicola	SRR20226091	20.59	19.38	2	59
D. novamexicana - D. lacicola	ERR2610692	19.7	61.49	2	142
D. novamexicana - D. littoralis	ERR2610692	19.69	61.48	2	142
D. novamexicana - D. littoralis	SRR13070670	16.35	20.22	2	56
D. novamexicana - D. montana	SRR3218635	25.21	95.57	2	216
D. novamexicana - D. montana	ERR2610692	19.7	61.51	2	142
D. novamexicana - D. virilis	SRR9426114	9.32	8.61	2	26
D. novamexicana - D. virilis	ERR2610692	19.7	61.48	2	142
D. obscura - D. tristis	SRR13070671	16.3	10.81	2	37
D. obscura - D. tristis	SRR13070669	19.21	18.36	2	55

D. parabipectinata - D. pseudoananassae	SRR13070715	10.96	17.22	2	45
D. parabipectinata - D. pseudoananassae	SRR13070721	11.7	8.39	2	28
D. persimilis - D. bifasciata	SRR12899691	24.45	19.76	2	63
D. persimilis - D. bifasciata	DRR061004	37.59	29.98	2	97
D. persimilis - D. pseudoobscura	SRR11230572	25.43	15.99	2	57
D. persimilis - D. pseudoobscura	SRR12899691	24.45	19.76	2	63
D. persimilis - D. subobscura	SRR12899691	24.45	19.76	2	63
D. persimilis - D. subobscura	SRR9967664	11.32	8.56	2	28
D. planitibia - D. silvestris	SRR2233314	18.35	12.57	2	43
D. planitibia - D. silvestris	SRR2233332	29.52	203.01	2	435
D. pseudoobscura - D. bifasciata	SRR11230572	25.82	14.29	2	54
D. pseudoobscura - D. bifasciata	DRR061004	37.7	28.83	2	95
D. pseudoobscura - D. subobscura	SRR11230572	25.82	14.29	2	54
D. pseudoobscura - D. subobscura	SRR9967664	11.35	9.13	2	29

D. sechellia - D. erecta	SRR14138507	21.37	17.45	2	56
D. sechellia - D. erecta	SRR6425990	40.14	17.62	2	75
D. sechellia - D. orena	SRR1977592	8.98	7.93	2	24
D. sechellia - D. orena	SRR14138507	21.87	34.25	2	90
D. sechellia - D. teissieri	SRR5860615	12.82	7.06	2	26
D. sechellia - D. teissieri	SRR14138507	21.06	17.34	2	55
D. sechellia - D. yacuba	SRR14138507	21.14	23.78	2	68
D. sechellia - D. yacuba	SRR9700025	16.48	7.73	2	31
D. simulans - D. erecta	SRR6425990	34.9	19.93	2	74
D. simulans - D. erecta	SRR11456787	26.86	12.34	2	51
D. simulans - D. orena	SRR1977592	9.22	8.4	2	26
D. simulans - D. orena	SRR11456787	26.86	12.34	2	51
D. simulans - D. sechellia	SRR14138507	24.04	15.67	2	55
D. simulans - D. sechellia	SRR11456787	26.86	12.34	2	51

D. simulans - D. teissieri	SRR5860615	12.52	11.5	2	35
D. simulans - D. teissieri	SRR11456787	26.86	12.34	2	51
D. simulans - D. yacuba	SRR11456787	26.86	12.34	2	51
D. simulans - D. yacuba	SRR9700025	12.97	8.79	2	30
D. subobscura - D. ambigua	SRR9967664	14.94	8.94	2	32
D. subobscura - D. ambigua	SRR13070667	16.97	25.77	2	68
D. subobscura - D. bifasciata	SRR9967664	14.94	8.94	2	32
D. subobscura - D. bifasciata	DRR061004	36.99	33.81	2	104
D. subobscura - D. obscura	SRR13070671	14.36	31.75	2	77
D. subobscura - D. obscura	SRR9967664	14.94	8.94	2	32
D. subobscura - D. tristis	SRR9967664	14.94	8.94	2	32
D. subobscura - D. tristis	SRR13070669	17.07	42.16	2	101
D. sulfurigaster albostrigata - D. kepulauna	SRR10100407	5.58	6.02	2	17

D. sulfurigaster albostrigata - D. kepulauna	SRR10100373	9.41	13.34	2	36
D. sulfurigaster albostrigata - D. pulaua	SRR10100407	5.58	6.01	2	17
D. sulfurigaster albostrigata - D. pulaua	SRR10100335	17.39	91.72	2	200
D. sulfurigaster albostrigata - D. sulfurigaster bilimbata	SRR10100323	26.51	21.29	2	69
D. sulfurigaster albostrigata - D. sulfurigaster bilimbata	SRR10100308	15.76	19.18	2	54
D. sulfurigaster albostrigata - D. sulfurigaster neonasuta	SRR5059342	13.07	180.59	2	374
D. sulfurigaster albostrigata - D. sulfurigaster neonasuta	SRR10100407	5.58	6.01	2	17
D. sulfurigaster albostrigata - D.	SRR10100323	26.51	21.29	2	69

sulfurigaster sulfurigaster					
D. sulfurigaster albostrigata - D. sulfurigaster sulfurigaster	SRR10100401	14.69	46.96	2	108
D. sulfurigaster bilimbata - pulaua	SRR10100335	17.39	91.72	2	200
D. sulfurigaster bilimbata - pulaua	SRR10100308	15.76	19.18	2	54
D. sulfurigaster bilimbata - sulfurigaster neonasuta	SRR5059342	13.07	180.59	2	374
D. sulfurigaster bilimbata - sulfurigaster neonasuta	SRR10100308	15.76	19.18	2	54
D. sulfurigaster bilimbata - D. sulfurigaster sulfurigaster	SRR10100308	15.76	19.18	2	54
D. sulfurigaster bilimbata - D. sulfurigaster sulfurigaster	SRR10100401	14.69	46.96	2	108
D. sulfurigaster neonasuta - pulaua	SRR5059342	13.07	180.59	2	374

D. sulfurigaster neonasuta - pulaua	SRR10100335	17.39	91.72	2	200
D. sulfurigaster neonasuta - D. sulfurigaster sulfurigaster	SRR10100401	14.69	46.96	2	108
D. sulfurigaster neonasuta - D. sulfurigaster sulfurigaster	SRR5059342	13.07	180.59	2	374
D. sulfurigaster sulfurigaster - D. pulaua	SRR10100401	14.69	46.96	2	108
D. sulfurigaster sulfurigaster - D. pulaua	SRR10100335	17.39	91.72	2	200
D. teissieri - D. erecta	SRR5860615	12.82	7.06	2	26
D. teissieri - D. erecta	SRR6425990	34.55	20.9	2	76
D. teissieri - D. orena	SRR5860615	12.82	7.06	2	26
D. teissieri - D. orena	SRR1977592	9.24	8.39	2	26
D. triauraria - D. rufa	SRR13188922	16.05	12.22	2	40
D. triauraria - D. rufa	SRR13070713	14.98	16.41	2	47
D. virilis - D. ezoana	SRR20226095	22.75	10.2	2	43
D. virilis - D. ezoana	SRR9426114	8.38	9.33	2	27

D. virilis - D. lacicola	SRR9426114	8.68	8.47	2	25
D. virilis - D. lacicola	SRR20226091	22	8.95	2	39
D. virilis - D. littoralis	SRR9426114	9.94	8.52	2	26
D. virilis - D. littoralis	SRR13070670	16.41	21.28	2	58
D. yacuba - D. erecta	SRR9700025	16.48	7.73	2	31
D. yacuba - D. erecta	SRR6425990	34.64	64.51	2	163
D. yacuba - D. orena	SRR1977592	9.24	8.4	2	26
D. yacuba - D. orena	SRR9700025	16.48	7.73	2	31
D. yacuba - D. santomea	SRR9700025	16.48	7.73	2	31
D. yacuba - D. santomea	SRR5860642	24.27	14.01	2	52
D. yacuba - D. teissieri	SRR5860615	13	9.41	2	31
D. yacuba - D. teissieri	SRR9700025	16.48	7.73	2	31

Pair	Premating Isolation (Yukilevich 2014)	Postzygotic Isolation (Yukilevich 2014)	Nei's *D* *(Yukilevich 2014)*	Biogeographic classification (Yukilevich 2014)
D. affinis - D. athabasca	0.775519466	NA	0.74	Sympatry
D. ambigua - D. obscura	0.993894994	NA	0.535	Sympatry
D. ambigua - D. persimilis	NA	1	1.66	Allopatry
D. ambigua - D. pseudoobscura	NA	1	1.66	Allopatry
D. americana - D. lacicola	1	NA	1.42	Sympatry
D. americana - D. littoralis	0.756484913	NA	1.21	Allopatry
D. americana - D. montana	0.992359932	1	1.48	Sympatry
D. americana - D. virilis	0.64428754	0	0.54	Allopatry
D. arizonae - D. mojavensis	0.766970618	0.375	0.279	Allopatry
D. auraria- D. rufa	0.987677141	NA	0.9245	Sympatry
D. bipectinata - D. malerkotliana	0.838798283	0.5	0.104	Sympatry
D. bipectinata - D. parabipectinata	0.55012987	0.5	0.145	Sympatry
D. bipectinata - D. pseudoananassae	0.840869565	NA	0.271	Sympatry
D. borealis - D. lacicola	1	NA	0.34	Sympatry
D. borealis - D. littoralis	1	NA	0.68	Allopatry
D. borealis - D. montana	0.994594595	0.5	0.21	Sympatry
D. erecta - D. orena	1	NA	1.03	Sympatry
D. flavomontana- D. montana	0.992857143	0.5	0.29	Sympatry

D. insularis - D. willistoni	0.923036649	NA	1.07	Sympatry
D. jambulina - D. punjabiensis	0.991326531	NA	0.944	Sympatry
D. kanekoi - D. ezoana	0.9945	NA	1.22	Sympatry
D. kanekoi - D. littoralis	0.5515	NA	1.22	Allopatry
D. kanekoi - D. lummei	0.722052705	NA	1.1	Sympatry
D. kanekoi - D. virilis	0.691660291	NA	1.1	Sympatry
D. kepulauna - D. albomicans	0.939064857	0	0.792	Allopatry
D. kikkawai- D. leontia	0.971875	0	0.244	Sympatry
D. kohkoa - D. albomicans	1	0.75	0.833	Sympatry
D. littoralis - D. ezoana	0.84125	NA	0.65	Sympatry
D. littoralis - D. montana	0.784	0.75	0.66	Sympatry
D. loweii - D. pseudoobscura	0.98	NA	1.05	Sympatry
D. lummei - D. ezoana	1	NA	1.22	Sympatry
D. lummei - D. littoralis	0.629	NA	1.22	Sympatry
D. lummei - D. virilis	0.263	0	0.35	Sympatry
D. malerkotliana - D. parabipectinata	0.853214286	0.625	0.226	Sympatry
D. malerkotliana - D. pseudoananassae	0.956950673	0.95	0.282	Sympatry
D. mauritiana - D. erecta	0.984042553	NA	1.59	Allopatry
D. mauritiana - D. orena	0.9953125	NA	1.07	Allopatry
D. mauritiana - D. sechellia	0.920463693	0.5	0.32	Allopatry
D. mauritiana - D. teissieri	0.875842697	NA	1.24	Allopatry

D. mauritiana - D. yacuba	0.989247312	NA	0.88	Allopatry
D. melanogaster - D. erecta	1	NA	1.63	Sympatry
D. melanogaster - D. orena	1	NA	1.14	Sympatry
D. melanogaster - D. sechellia	0.619439164	1	0.62	Allopatry
D. melanogaster - D. simulans	0.826903313	1	0.55	Sympatry
D. melanogaster - D. teissieri	1	NA	1.01	Sympatry
D. melanogaster - D. yacuba	0.734042553	NA	0.94	Sympatry
D. miranda - D. persimilis	1	1	0.56	Sympatry
D. miranda - D. pseudoobscura	0.99537037	1	0.56	Sympatry
D. nasuta - D. albomicans	0.068400226	0	0.45	Allopatry
D. novamexicana - D. lacicola	1	NA	1.2	Allopatry
D. novamexicana - D. littoralis	0.788918206	NA	0.97	Allopatry
D. novamexicana - D. montana	1	0.75	1.22	Sympatry
D. novamexicana - D. virilis	0.507460036	0	0.5	Allopatry
D. obscura - D. tristis	0.989746683	NA	0.47	Sympatry
D. parabipectinata - D. pseudoananassae	0.762171946	NA	0.366	Sympatry

D. persimilis - D. bifasciata	0.975	NA	1.95	Allopatry
D. persimilis - D. pseudoobscura	0.785042646	0.5	0.41	Sympatry
D. persimilis - D. subobscura	0.99137931	NA	1.66	Allopatry
D. pseudoobscura - D. bifasciata	0.941935484	NA	1.95	Allopatry
D. pseudoobscura - D. subobscura	0.962121212	NA	1.66	Allopatry
D. sechellia - D. erecta	1	NA	1.51	Allopatry
D. sechellia - D. orena	1	NA	1.27	Allopatry
D. sechellia - D. teissieri	0.995	NA	1.36	Allopatry
D. sechellia - D. yacuba	1	NA	1.27	Allopatry
D. simulans - D. erecta	0.997382199	NA	1.5	Sympatry
D. simulans - D. orena	0.953846154	NA	1.01	Sympatry
D. simulans - D. teissieri	0.961325967	NA	1.24	Sympatry
D. simulans - D. yacuba	0.958730159	NA	1	Sympatry
D. subobscura - D. ambigua	0.929569267	NA	1.44	Sympatry
D. subobscura - D. bifasciata	0.934782609	NA	1.33	Allopatry
D. subobscura - D. obscura	1	NA	1.25	Sympatry
D. subobscura - D. tristis	0.98273878	NA	1.1	Sympatry
D. sulfurigaster albostrigata - D. kepulauna	0.979381443	1	0.902	Sympatry

D. sulfurigaster albostrigata - D. pulaua	0.992078001	0	0.389	Sympatry
D. sulfurigaster albostrigata - D. sulfurigaster bilimbata	0.430365297	0.25	0.386	Allopatry
D. sulfurigaster albostrigata - D. sulfurigaster neonasuta	0.183673469	0	0.557	Allopatry
D. sulfurigaster albostrigata - D. sulfurigaster sulfurigaster	0.132816824	0	0.378	Allopatry
D. sulfurigaster bilimbata - D. pulaua	0.921381775	0	0.448	Allopatry
D. sulfurigaster bilimbata - D. sulfurigaster neonasuta	0.595550485	0.5	0.491	Allopatry
D. sulfurigaster bilimbata - D. sulfurigaster sulfurigaster	0.097560976	0	0.555	Allopatry
D. sulfurigaster neonasuta - D. pulaua	0.909866017	0.5	0.705	Allopatry
D. sulfurigaster neonasuta - D. sulfurigaster sulfurigaster	0.415376106	0.5	0.6	Allopatry
D. sulfurigaster sulfurigaster - D. pulaua	0.785467128	0	0.561	Allopatry
D. teissieri- D. erecta	0.977011494	NA	1.54	Sympatry
D. teissieri - D. orena	1	NA	1.47	Sympatry
D. triauraria - D. rufa	0.963319303	NA	0.7335	Sympatry

D. virilis - D. ezoana	0.9645	NA	1.22	Sympatry
D. virilis - D. lacicola	0.799457995	NA	1.22	Allopatry
D. virilis - D. littoralis	0.682291667	0	1.22	Allopatry
D. yacuba - D. erecta	0.994505495	NA	1.4	Sympatry
D. yacuba - D. orena	1	NA	1.12	Sympatry
D. yacuba - D. santomea	0.842032011	0.5	0.3	Sympatry
D. yacuba - D. teissieri	0.988372093	NA	0.39	Sympatry

Pair	IM vs SI (Difference in logL)	IM vs MIG (Difference in logL)	IIM vs IM (Difference in logL)	IIM vs IIM limit (Difference in logL)	IIM vs SC (Difference in logL)
D. affinis - D. athabasca	14.24491308	14.24786115	13.19500615	27.50108462	-11.68692269
D. ambigua - D. obscura	13.04206931	10.77352297	-0.004148171	10.60620122	-2.243587195
D. ambigua - D. persimilis	3.991292005	8.721210027	0.885644309	4.22470393	0.9035271
D. ambigua - D. pseudoobscura	6.28946069	10.69445586	-0.001496552	5.352178621	0.184398621
D. americana - D. lacicola	14.97877606	20.15603897	-0.00098662	13.43291573	-0.543151174
D. americana - D. littoralis	10.75602422	21.37606345	-0.000487668	9.583662108	0.210473543
D. americana - D. montana	13.88673271	16.81566192	-0.000957944	12.1833729	-0.422316355
D. americana - D. virilis	9.8512	22.89938538	0.004887747	8.821927273	0.205844664
D. arizonae - D. mojavensis	7.734467147	6.553580288	-0.12390625	5.998010897	-0.111483654
D. auraria- D. rufa	8.383414663	10.23931827	-0.003951683	6.949618029	-0.188132692
D. bipectinata - D. malerkotliana	1.188225843	0.731077154	-0.001970412	0.732824532	-0.201507303
D. bipectinata - D. parabipectinata	2.663282331	1.459357331	0.123712782	1.579659398	-0.401817857

D. bipectinata - D. pseudoananassae	7.381202604	5.514776042	-0.000413021	5.514017448	-0.185692187
D. borealis - D. lacicola	11.96261789	9.871978593	-0.00380581	9.778157798	-0.749948471
D. borealis - D. littoralis	12.19612536	16.73800749	-0.001010145	10.73420942	0.085071256
D. borealis - D. montana	11.98553796	10.19691418	-0.008699848	9.718868902	11.97683887
D. erecta - D. orena	7.746144262	26.7541541	0.059769672	7.062721107	0.216444262
D. flavomontana- D. montana	8.306011416	7.415114884	-0.005147399	6.583948844	-0.305093786
D. insularis - D. willistoni	10.62688983	14.28305349	-0.00860407	9.039880233	-2.115722384
D. jambulina - D. punjabiensis	2.51483417	2.523468919	0.116212934	2.794003668	-0.140705019
D. kanekoi - D. ezoana	12.76536306	10.66063122	0.129614694	10.79808776	-3.44634102
D. kanekoi - D. littoralis	9.462859722	13.5244338	-0.002183102	8.104767593	0.056571528
D. kanekoi - D. lummei	13.1923296	11.23698587	0.148437892	11.39380538	-4.706782287
D. kanekoi - D. virilis	10.39194365	21.15592037	-5.39683E-05	9.371634392	0.251877778
D. kepulauna - D. albomicans	0.504826781	1.350282336	0.537596296	0.647783476	0.304171225
D. kikkawai- D. leontia	2.527031561	5.263351827	0.000120266	1.723295681	0.003269269

D. kohkoa - D. albomicans	2.783279551	2.575090501	1.709070185	4.286608839	0.096161214
D. littoralis - D. ezoana	11.08127964	9.346042944	0.280150806	9.63438871	-2.342203024
D. littoralis - D. montana	14.21735625	15.80016435	-0.001536574	12.48827014	-0.406300231
D. loweii - D. pseudoobscura	8.322333256	20.13087209	0.068197907	7.460612093	0.481722093
D. lummei - D. ezoana	13.72533959	11.49402143	0.120456531	11.62248714	-3.982517755
D. lummei - D. littoralis	9.755274043	9.09784383	2.248066809	11.36042149	-3.834426596
D. lummei - D. virilis	3.165555037	2.925311567	3.723085448	6.655712127	-0.652589552
D. malerkotliana - D. parabipectinata	2.792611864	2.786774068	0.077203898	2.86976661	0.051583559
D. malerkotliana - D. pseudoananassae	6.266167179	4.52604359	0.039803077	4.565361538	-0.344649744
D. mauritiana - D. erecta	3.565318038	35.91095759	0.95649462	4.175500949	0.16995443
D. mauritiana - D. orena	8.921816038	27.64561667	0.153428616	8.280984906	0.065475472
D. mauritiana - D. sechellia	3.607259076	8.734057643	-0.003653185	2.747926274	-0.58430414
D. mauritiana - D. teissieri	10.6528171	18.78608549	0.000613472	9.463174611	0.130890674

D. mauritiana - D. yacuba	5.532607143	25.122035	-0.000384286	4.958928214	0.015045
D. melanogaster - D. erecta	8.713952797	21.96184126	0.013171678	7.84508986	0.380977622
D. melanogaster - D. orena	7.244511278	28.52324812	0.321815163	6.928543108	0.173494987
D. melanogaster - D. sechellia	6.353749609	26.04337167	0.043107981	5.676643192	0.181411581
D. melanogaster - D. simulans	3.649863433	17.8269005	0.698151741	3.707978856	0.069332338
D. melanogaster - D. teissieri	6.949174615	15.50903538	-0.001310769	6.039254615	0.116134615
D. melanogaster - D. yacuba	6.238870763	18.38612415	0.640359746	6.176389407	0.422518644
D. miranda - D. persimilis	4.504543633	14.30242022	0.173907491	3.862602809	0.133642322
D. miranda - D. pseudoobscura	2.63669043	13.66360508	1.114391211	3.138232813	0.285713086
D. nasuta - D. albomicans	0.05642791	1.363533433	0.080371045	0.114321045	0.030216567
D. novamexicana - D. lacicola	15.76095021	14.10345658	-0.003652058	13.51413272	-0.574221605
D. novamexicana - D. littoralis	14.20994648	14.0704459	-0.001909766	12.24143867	14.2080375
D. novamexicana - D. montana	12.44508798	13.68200935	-0.001257443	10.72931031	-0.936492939
D. novamexicana - D. virilis	12.28400854	13.03889128	-0.00082847	10.55274306	-0.186704448

D. obscura - D. tristis	14.23822549	11.28884506	0.059387945	11.35275929	-4.194614625
D. parabipectinata - D. pseudoananassae	6.379907729	4.648941787	0.130319324	4.785292029	-0.945142029
D. persimilis - D. bifasciata	11.30813563	13.36962097	0.016970674	10.08949194	0.207610704
D. persimilis - D. pseudoobscura	0.110942778	2.905763148	-0.00014537	0.058538889	0.103986111
D. persimilis - D. subobscura	4.272969496	12.40674125	-0.000791114	3.597905172	-0.091715517
D. pseudoobscura - D. bifasciata	11.53854694	11.6669623	-0.001302987	10.04194879	-0.332282361
D. pseudoobscura - D. subobscura	8.960013079	14.9891891	0.012954496	8.036866349	0.249641008
D. sechellia - D. erecta	5.415726944	21.723655	0.069866389	4.821758889	0.182079722
D. sechellia - D. orena	5.090056122	37.82024082	1.156619388	5.783765306	0.192361224
D. sechellia - D. teissieri	6.434443373	24.97958133	0.336902108	6.106837952	0.143043976
D. sechellia - D. yacuba	5.158990938	22.78730844	0.601375937	5.145194688	0.34580125
D. simulans - D. erecta	8.037706845	27.93186369	-0.000111905	7.318430952	0.168446726

D. simulans - D. orena	7.754219048	24.84440306	0.320733673	7.370303741	0.353908503
D. simulans - D. teissieri	6.576912573	20.43733333	0.385203801	6.238796784	0.42937807
D. simulans - D. yacuba	8.185818208	38.47891561	-0.000549133	7.632178035	0.022125145
D. subobscura - D. ambigua	8.700436107	19.36981798	-0.001460624	7.839205052	0.051959138
D. subobscura - D. bifasciata	8.842215994	13.51382063	-0.000243647	7.801730942	0.143239163
D. subobscura - D. obscura	12.1318959	19.35435562	-0.000442249	10.98921581	-0.129221125
D. subobscura - D. tristis	11.94750622	15.59464541	-0.000721617	10.68239502	-0.040728616
D. sulfurigaster albostrigata - D. kepulauna	0.577541573	4.824444195	0.672075843	0.850208614	0.069114045
D. sulfurigaster albostrigata - D. pulaua	0.034882534	5.654154452	-9.86301E-05	0.028909247	0.034785274
D. sulfurigaster albostrigata - D. sulfurigaster bilimbata	0.055452349	2.624910067	0.093507215	0.136031208	-0.066663591
D. sulfurigaster albostrigata - D. sulfurigaster neonasuta	0.274214555	2.208500856	0.512006336	0.486917466	-0.229921575

D. sulfurigaster albostrigata - D. sulfurigaster sulfurigaster	0.002625932	2.996456102	0.000366949	0.002997797	-0.086701695
D. sulfurigaster bilimbata - D. pulaua	-4E-06	2.401808	0.071788286	0.072888286	-0.861522
D. sulfurigaster bilimbata - D. sulfurigaster neonasuta	6.113404242	7.785624848	0.094340455	5.165595758	0.581416818
D. sulfurigaster bilimbata - D. sulfurigaster sulfurigaster	0.150878947	0.228193567	-0.000371053	0.081200877	-0.170971637
D. sulfurigaster neonasuta - D. pulaua	5.657522305	7.244793114	0.33997515	5.151641467	0.58805
D. sulfurigaster neonasuta - D. sulfurigaster sulfurigaster	5.463299118	6.720198529	0.170557794	4.670718529	0.606013088
D. sulfurigaster sulfurigaster - D. pulaua	-3.7037E-06	1.660530247	2.77778E-06	0.001225617	-0.188035802
D. teissieri- D. erecta	6.346162703	19.59764865	0.071657568	5.617326216	0.107873243
D. teissieri - D. orena	4.334272256	24.82675335	0.439310671	4.259476829	0.247660671

D. triauraria - D. rufa	9.332157377	9.488193169	-0.11545847	7.634520765	-0.62499235
D. virilis - D. ezoana	7.882217617	6.412685492	0.121976425	6.541028238	-0.761964767
D. virilis - D. lacicola	12.58077167	13.06604722	-0.000918333	10.93644278	0.089299167
D. virilis - D. littoralis	11.24893828	15.49669583	0.093423958	10.0997362	0.251539844
D. yacuba - D. erecta	5.071429722	23.21115278	0.012961667	4.469561667	0.153796667
D. yacuba - D. orena	6.052963208	28.91291792	0.008044654	5.487500943	0.004097799
D. yacuba - D. santomea	5.51974082	13.61411066	-0.001588525	4.567719672	-0.046796721
D. yacuba - D. teissieri	8.314546516	13.49229365	0.043210451	7.185981352	0.481718648

Pair	IM vs SI (Difference in logL -- No Recombination)	IM vs SI (Difference in logL -- With Recombination)
D. affinis - D. athabasca	-0.154304	9.4760575
D. ambigua - D. obscura	-0.63247475	1.8216665
D. ambigua - D. persimilis	-6.669142	0.22948725
D. ambigua - D. pseudoobscura	-7.86996125	0.1155895
D. americana - D. montana	0.1722205	0.9437255
D. americana - D. virilis	-0.00029675	1.1008325
D. americana - D. lacicola	-5.675239	0.293779
D. americana - D. littoralis	0.061838	0.318727
D. arizonae - D. mojavensis	-2.12492575	2.507734
D. auraria- D. rufa	-5.53448175	1.8336325
D. bipectinata - D. malerkotliana	-0.14979975	7.444685
D. bipectinata - D. parabipectinata	0.6888265	2.85252625
D. bipectinata - D. pseudoananassae	0.39919825	10.011253
D. borealis - D. lacicola	-0.582845	3.011703
D. borealis - D. littoralis	-6.40894375	0.89922525
D. borealis - D. montana	-1.304686	1.46773075
D. erecta - D. orena	0.15642725	0.30473525
D. flavomontana- D. montana	-1.47105625	0.28318675
D. insularis - D. willistoni	-6.79094175	1.4804645
D. jambulina - D. punjabiensis	0.18039175	9.0430455
D. kanekoi - D. ezoana	0.491416	4.520206
D. kanekoi - D. littoralis	-6.95620425	1.3461545
D. kanekoi - D. lummei	0.023294	4.623887
D. kanekoi - D. virilis	-0.000216	0.25911775
D. kepulauna - D. albomicans	-2.70380525	1.392413
D. kikkawai- D. leontia	-5.1257895	0.885859
D. kohkoa - D. albomicans	-0.05141025	8.10915425
D. littoralis - D. ezoana	0.4638655	7.7325425
D. littoralis - D. montana	-4.93153525	0.9080145
D. loweii - D. pseudoobscura	0.1259705	1.475313
D. lummei - D. ezoana	0.19643625	5.336632
D. lummei - D. littoralis	0.08678375	8.1485735
D. lummei - D. virilis	-0.0463755	7.52355625
D. malerkotliana - D. parabipectinata	0.13188	10.61952175
D. malerkotliana - D. pseudoananassae	0.29347375	3.239107
D. mauritiana - D. erecta	0.3344945	0.608184
D. mauritiana - D. orena	-0.00181775	0.22230825

D. mauritiana - D. sechellia	-9.2088235	1.4877295
D. mauritiana - D. teissieri	0.0743675	1.60343475
D. mauritiana - D. yacuba	-0.00088875	0.3945455
D. melanogaster - D. erecta	0.04074125	0.0600005
D. melanogaster - D. orena	-0.000964	-0.00286875
D. melanogaster - D. sechellia	-0.0005805	0.4052805
D. melanogaster - D. simulans	0.7517305	0.68122625
D. melanogaster - D. teissieri	0.02389	1.208034
D. melanogaster - D. yacuba	-1.30E-05	0.07559225
D. miranda - D. persimilis	0.4444185	1.53889375
D. miranda - D. pseudoobscura	-0.0001285	1.50660875
D. nasuta - D. albomicans	-1.63932875	1.9127715
D. novamexicana - D. lacicola	-1.455242	4.48309925
D. novamexicana - D. littoralis	-2.22167625	1.75211425
D. novamexicana - D. montana	-5.5826495	0.73650975
D. novamexicana - D. virilis	-4.97667175	0.451082
D. obscura - D. tristis	0.1502035	6.021555
D. parabipectinata - D. pseudoananassae	0.5475705	5.42126225
D. persimilis - D. bifasciata	-5.97366225	3.8885565
D. persimilis - D. pseudoobscura	0.3772995	0.52908975
D. persimilis - D. subobscura	-0.00042525	2.67E-05
D. pseudoobscura - D. bifasciata	-1.89383775	0.03296825
D. pseudoobscura - D. subobscura	0.0205525	0.77406975
D. sechellia - D. erecta	0.106349	0.06987925
D. sechellia - D. orena	-0.00011025	0.229013
D. sechellia - D. teissieri	0.01147	0.075828
D. sechellia - D. yacuba	0.016072	0.1990405
D. simulans - D. erecta	-0.00114175	1.106863
D. simulans - D. orena	-0.00078075	0.80872725
D. simulans - D. teissieri	-0.00128975	1.1766165
D. simulans - D. yacuba	-0.00290275	0.72976775
D. subobscura - D. ambigua	0.29490925	0.79274025
D. subobscura - D. bifasciata	-7.03964625	0.2769065
D. subobscura - D. obscura	0.009463	0.18738925
D. subobscura - D. tristis	-4.9943265	1.55378425
D. sulfurigaster albostrigata - D. kepulauna	-5.10E-05	0.66901075
D. sulfurigaster albostrigata - D. pulaua	-3.273116	1.75750475
D. sulfurigaster albostrigata - D. sulfurigaster bilimbata	-3.31889625	1.5914515
D. sulfurigaster albostrigata - D. sulfurigaster neonasuta	-2.75170775	1.70006225
D. sulfurigaster albostrigata - D. sulfurigaster sulfurigaster	-1.8848695	0.381667
D. sulfurigaster bilimbata - D. pulaua	-1.76603975	1.33646125
D. sulfurigaster bilimbata - D. sulfurigaster neonasuta	-2.7599	2.82219175

D. sulfurigaster bilimbata - D. sulfurigaster sulfurigaster	-0.1085505	3.0562085
D. sulfurigaster neonasuta - D. pulaua	-5.3976675	2.4274
D. sulfurigaster neonasuta - D. sulfurigaster sulfurigaster	-0.73185425	0.7612595
D. sulfurigaster sulfurigaster - D. pulaua	-1.39609025	1.44585225
D. teissieri - D. orena	-0.00012325	0.31169375
D. teissieri- D. erecta	0.0013285	0.34871325
D. triauraria - D. rufa	-1.820373	1.352234
D. virilis - D. ezoana	0.0004845	5.23917
D. virilis - D. lacicola	-2.8273075	0.8447935
D. virilis - D. littoralis	-7.478602	1.587921
D. yacuba - D. erecta	-1.25E-06	0.9683835
D. yacuba - D. orena	-0.00040575	0.36649375
D. yacuba - D. santomea	-0.0001785	0.0026815
D. yacuba - D. teissieri	-6.107981	1.13633825

Current biogeography	IM vs SI (no. of pairs fitting models in each pairwise comparison).	
	IM	**SI**
Sympatry	51	4
Allopatry	30	8
	IM vs SI (no. of pairs fitting models in each pairwise comparison).	
	IM	**IIM**
Sympatry	52	3
Allopatry	38	0
	IIM vs SC (no. of pairs fitting models in each pairwise comparison).	
	IIM	**SC**
Sympatry	27	28
Allopatry	26	12
	IIM vs IIM Limit (no. of pairs fitting models in each pairwise comparison).	
	IIM	**IIM Limit**
Sympatry	49	6
Allopatry	30	8
	IM vs MIG (no. of pairs fitting models in each pairwise comparison).	
	IM	**MIG**
Sympatry	53	2
Allopatry	34	4

References and Notes

1. J. A. Coyne, H. A. Orr, Patterns of Speciation in Drosophila. *Evolution*, doi: 10.2307/2409213 (1989).

2. J. A. Coyne, H. A. Orr, " PATTERNS OF SPECIATION IN DROSOPHILA " REVISITED. **51**, 295–303 (1997).

3. R. Yukilevich, Asymmetrical patterns of speciation uniquely support reinforcement in drosophila. *Evolution* **66** (2012).

4. R. Yukilevich, The rate test of speciation: Estimating the likelihood of non-allopatric speciation from reproductive isolation rates in drosophila. *Evolution* **68**, 1150–1162 (2014).

5. P. Nosil, Degree of sympatry affects reinforcement in drosophila. *Evolution* **67** (2013).

6. D. J. Funk, P. Nosil, W. J. Etges, Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3209–3213 (2006).

7. M. Turelli, J. R. Lipkowitz, Y. Brandvain, On the coyne and orr-igin of species: Effects of intrinsic postzygotic isolation, ecological differentiation, x chromosome size, and sympatry on drosophila speciation. *Evolution* **68** (2014).

8. B. Y. Kim, J. R. Wang, D. E. Miller, O. Barmina, E. Delaney, A. Thompson, A. A. Comeault, D. Peede, E. R. R. D'agostino, J. Pelaez, J. M. Aguilar, D. Haji, T. Matsunaga, E. E. Armstrong, M. Zych, Y. Ogawa, M. Stamenković-Radak, M. Jelić, M. S. Veselinović, M. Tanasković, P. Erić, J. J. Gao, T. K. Katoh, M. J. Toda, H. Watabe, M. Watada, J. S. Davis, L. C. Moyle, G. Manoli, E. Bertolini, V. Košťál, R. S. Hawley, A. Takahashi, C. D. Jones, D. K. Price, N. Whiteman, A. Kopp, D. R. Matute, D. A. Petrov, Highly contiguous assemblies of 101 drosophilid genomes. *eLife* **10** (2021).

9. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).

10. T. Brůna, K. J. Hoff, A. Lomsadze, M. Stanke, M. Borodovsky, BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics Bioinforma.* **3**, 1–11 (2021).

11. S. Neph, M. S. Kuehn, A. P. Reynolds, E. Haugen, R. E. Thurman, A. K. Johnson, E. Rynes, M. T. Maurano, J. Vierstra, S. Thomas, R. Sandstrom, R. Humbert, J. A. Stamatoyannopoulos, BEDOPS: High-performance genomic feature operations. *Bioinformatics* **28** (2012).

12. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, doi: 10.1093/bioinformatics/btq033 (2010).

13. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. **00**, 1–3 (2013).

14. V. Md, S. Misra, H. Li, S. Aluru, "Efficient architecture-aware acceleration of BWA-MEM for multicore systems" in *Proceedings - 2019 IEEE 33rd International Parallel and Distributed Processing Symposium, IPDPS 2019* (2019).

15. A. Tarasov, A. J. Vilella, E. Cuppen, I. J. Nijman, P. Prins, Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).

16. Github, Picard toolkit. *Broad Inst. GitHub Repos.* (2019).

17. J. Köster, S. Rahmann, Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* **28** (2012).

18. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27** (2011).

19. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *GigaScience* **10** (2021).

20. D. R. Laetsch, G. Bisschop, S. H. Martin, S. Aeschbacher, D. Setter, K. Lohse, Demographically explicit scans for barriers to gene flow using gIMble. *PLOS Genet.* **19**, e1010999 (2023).

21. P. R. Haddrill, B. Charlesworth, D. L. Halligan, P. Andolfatto, Patterns of intron sequence evolution in Drosophila are dependent upon length and GC content. *Genome Biol.* **6** (2005).

22. P. Andolfatto, Adaptive evolution of non-coding DNA in Drosophila. *Nature* **437** (2005).

23. D. L. Halligan, P. D. Keightley, Ubiquitous selective constraints in the Drosophila genome revealed by a genome-wide interspecies comparison. *Genome Res.* **16**, 875–884 (2006).

24. H. M. Wilkinson-Herbots, The distribution of the coalescence time and the number of pairwise nucleotide differences in the "isolation with migration" model. *Theor. Popul. Biol.*, doi: 10.1016/j.tpb.2007.11.001 (2008).

25. H. M. Wilkinson-Herbots, The distribution of the coalescence time and the number of pairwise nucleotide differences in a model of population divergence or

speciation with an initial period of gene flow. *Theor. Popul. Biol.* **82**, 92–108 (2012).

26. Y. Wang, P. McNeil, R. Abdulazeez, M. Pascual, S. E. Johnston, P. D. Keightley, D. J. Obbard, Variation in mutation, recombination, and transposition rates in *Drosophila melanogaster* and *Drosophila simulans*. *Genome Res.* **33**, 587–598 (2023).

27. F. Baumdicker, G. Bisschop, D. Goldstein, G. Gower, A. P. Ragsdale, G. Tsambos, S. Zhu, B. Eldon, E. C. Ellerman, J. G. Galloway, A. L. Gladstein, G. Gorjanc, B. Guo, B. Jeffery, W. W. Kretzschumar, K. Lohse, M. Matschiner, D. Nelson, N. S. Pope, C. D. Quinto-Cortés, M. F. Rodrigues, K. Saunack, T. Sellinger, K. Thornton, H. Van Kemenade, A. W. Wohns, Y. Wong, S. Gravel, A. D. Kern, J. Koskela, P. L. Ralph, J. Kelleher, Efficient ancestry and mutation simulation with msprime 1.0. *Genetics* **220**, iyab229 (2022).

28. J. Kelleher, A. M. Etheridge, G. McVean, "Efficient coalescent simulation and genealogical analysis for large sample sizes" (preprint, Evolutionary Biology, 2015); https://doi.org/10.1101/033118.

29. R. R Core Team, R: A language and environment for statistical computing. (2013).

30. A. Kopp, O. Barmina, Evolutionary history of the *Drosophila bipectinata* species complex. *Genet. Res.* **85**, 23–46 (2005).