Using Variable Window Sizes for Phylogenomic Analyses of Whole Genome Alignments

Jeremias Ivan[1], Robert Lanfear[1]

[1]*Research School of Biology, Australian National University, ACT, Australia*

Corresponding author: Jeremias Ivan

Research School of Biology, Australian National University, ACT, Australia

Ivan.Jeremias@anu.edu.au

ABSTRACT

Many phylogenomic studies used non-overlapping windows to address gene tree discordance across a set of aligned genomes. Recently, Ivan et al. (2025) proposed an information theoretic approach to choose an optimal window size given the alignment. However, this approach selects only a single fixed window size per chromosome, which is a useful first step but fails to account for variation in the size of non-recombining regions along each chromosome. Such variation is expected to occur due to the stochastic nature of recombination as well as the variation in recombination rates along chromosomes. In this study, we extend the approach of Ivan et al. (2025) to allow window sizes to vary across the chromosome, using a splitting-and-merging strategy that allows for each window to be of an arbitrary length. We showed that the new method outperformed the fixed-window approach in recovering gene tree topologies on a wide range of simulated datasets. Applying the new method on the genomes of seven *Heliconius* butterflies, we found that the average window sizes for the group ranged between 538-808bp, but with a very similar distribution of gene tree topologies compared to previous studies that used fixed window sizes. For the genomes

of great apes, the average window sizes ranged from 4.2kb to 6.2kb, with the proportion of the major topology (i.e., grouping human and chimpanzee together) reaching approximately 80%. In conclusion, our study highlights the limitations of using a fixed window size when recombination rates vary across the chromosomes, and proposes a splitting-and-merging approach that allows for variable window sizes across whole genome alignments.

Keywords: gene tree discordance, recombination rate, simulation, splitting, merging, AIC

INTRODUCTION

Many phylogenomic studies of closely-related taxa begin by aligning the genomes of species or individuals, and then use those alignments to infer evolutionary events such as speciation, introgression, and demographic changes (e.g., Pease et al. 2016; Morales-Briones et al. 2021; Lescroart et al. 2023; Herrig et al. 2024). As different sites in the alignments may reflect different evolutionary histories (Maddison 1997; Degnan and Rosenberg 2009; Mallet et al. 2016; Hibbins et al. 2023; Steenwyk et al. 2023), methods that account for this variation become important. For example, Thawornwattana et al. (2018, 2022) extracted 100-locus blocks (where a locus was defined as a coding or non-coding region with a minimum length of 100bp and spaced at least 2kb apart) and inferred the species tree for each block using BPP (Flouri et al. 2018). This approach is useful for identifying introgressed regions, but it excludes most parts of the genomes. To infer the complete histories of whole genome alignments – which is important for analyses such as PhyloGWAS (Pease et al. 2016) – methods that use Hidden Markov Models (HMMs) and Ancestral Recombination Graphs (ARGs) are powerful for inferring site-specific histories by considering the similarities between neighbouring sites, potentially under a population genetic model of evolutionary change. However, these approaches have their own limitations, including accuracy and

scalability (Siepel and Haussler 2005; Brandt et al. 2024). A relatively simpler alternative is to partition the genome alignment into non-overlapping blocks (or windows) and infer a single gene tree for each window. This approach is commonly-used in phylogenomic studies to assess various evolutionary processes, such as incomplete lineage sorting (ILS) and introgression (e.g., Fontaine et al. 2015; Pease et al. 2016; Edelman et al. 2019; Meleshko et al. 2021; Yang et al. 2021; Feng et al. 2023; Lescroart et al. 2023; Herrig et al. 2024). However, one of the main challenges for this approach is to select a window size that best reflects the recombination breakpoints of given alignment, with a trade-off between concatenation preferring the primary tree topology (when the window size is too long, thus recombination is ignored) and gene tree estimation error (when the window size is too short, thus limiting the information available to estimate a tree for each window). Most previous studies have selected a single window size somewhat arbitrarily, in an attempt to balance these tradeoffs. However, the selection of window size can have significant effects on downstream inferences. For example, Ivan et al. (2025) showed that in the genomes of great apes (i.e., human, chimpanzee, gorilla, orangutan), only 50% of gene trees grouped human and chimpanzee together when using 250bp non-overlapping windows across all chromosomes. However, this proportion increased to almost 100% when the window size was 64kb. This highlights the importance of carefully selecting a window size in such analyses.

Recently, Ivan et al. (2025) proposed that the AIC (Akaike Information Criteraion; Akaike 1974) – which reflects how the substitution models and the trees fit the observed alignment data – could be used to select a window size in phylogenomic non-overlapping window analyses. They showed that the AIC was a good predictor of the accuracy of different window sizes in recovering the true histories of individual sites along simulated

chromosomes. Applying this approach on empirical datasets, Ivan et al. (2025) further showed that the best window sizes for chromosome alignments of *Heliconius* butterflies ranged between ≤125-250bp, while the window sizes for great apes ranged between 500bp-1kb (and 4kb for the mitochondrial genome). These windows are significantly shorter than the ones used in previous studies (e.g., Edelman et al. (2019) used 10kb and 50kb windows to analyse the genomes of *Heliconius* butterflies), but still longer than the estimated lengths of non-recombining blocks – sometimes called 'coalescent genes' or 'c-genes' – of each group (e.g., Springer and Gatesy (2016) estimated the average length of c-genes between human, chimpanzee, and gorilla to be ~109bp, but see also Edwards et al. (2016)). While the AIC provides a less arbitrary way to choose a window size for non-overlapping window analysis, the method proposed by Ivan et al. (2025) has two key limitations: it was evaluated using simulated chromosomes with constant recombination rates, and it selects a single fixed window size for each chromosome. A single fixed window size is problematic as some non-recombining blocks will be much longer than others even under constant recombination rate, due to the stochastic nature of recombination. And since recombination rates vary along the chromosomes for a variety of reasons (Jensen-Seaman et al. 2004; McVean et al. 2004; Chan et al. 2012; Edelman et al. 2019), the best window size may also vary systematically along chromosomes.

In this study, we extend the method proposed in Ivan et al. (2025) to allowing the window size to vary along chromosomes via a splitting-and-merging approach (Fig. 1). We do this by first partitioning the alignment into non-overlapping windows of a fixed size. Then, we split each window into two sub-windows and build gene trees for both the initial window and the two sub-windows. If the total AIC of the two sub-windows is better than the AIC of the initial window, we keep the split; otherwise, we retain the initial window. Once no further AIC

improvement could be achieved through splitting, we then iteratively merge neighbouring windows. Similar to the splitting step, for every pair of neighbouring windows, we compare the total AIC of the two initial windows and the AIC of the newly-merged window, keeping the merge only if the AIC improves. Through this splitting-and-merging procedure, we generate a set of non-overlapping windows that can vary in length along the chromosome and should better represent the underlying recombination breakpoints of the alignment compared to using a single window size. This approach is similar to GARD (Genetic Algorithm for Recombination Detection) (Kosakovsky Pond et al. 2006) that aims to detect recombination breakpoints from multiple sequence alignments. However, our method differs in using: (i) Maximum Likelihood (ML) instead of Neighbour-Joining (NJ) to build the gene trees; (ii) the AIC instead of the corrected AIC (AIC_c) to score the split because both criteria yielded very similar results in prior work (Ivan et al. 2025, results not shown); and (iii) a greedy algorithm instead of a genetic algorithm to locate the putative genomic breakpoints, which allows our method to scale to whole genome alignments in reasonable amount of time.

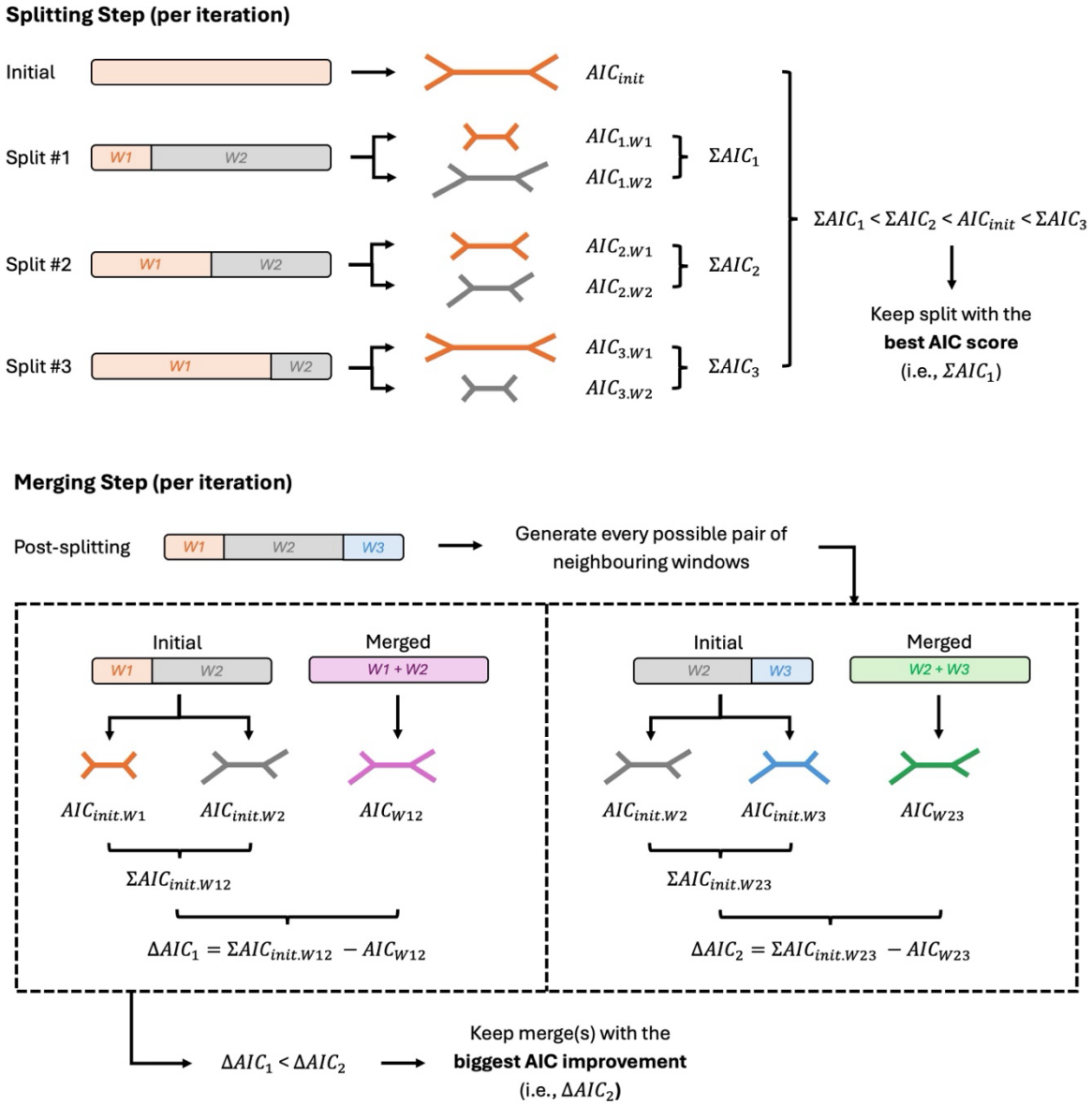

Figure 1. *Overview of the splitting-and-merging approach proposed in this study. The splitting step is iteratively applied to every window in the alignment – selecting split configuration with the best AIC score for each window – until no further improvement can be achieved from any new split. The merging step is then iteratively applied to every possible pair of neighbouring windows in the alignment, prioritising merges with the biggest AIC improvement for each iteration until no further improvement can be achieved from any new merge.*

In order to assess if using variable windows sizes would improve the accuracy of non-overlapping window method to recover the true topology of every site in the alignment (as compared to when using only one fixed window size), we benchmarked our splitting-and-merging approach with two simulation schemes: one with homogeneous recombination rates from Ivan et al. (2025), and one with heterogeneous recombination rates generated in this study. For each approach, we measured the accuracy of the final set of window sizes using the two metrics used in Ivan et al. (2025): *site accuracy* (i.e., the proportion of sites from non-overlapping windows that recover the true simulated topology) and *RMSE* (i.e., the root mean squared error of the observed tree topology distribution from non-overlapping windows compared to the true simulated distribution). Finally, we applied our method to the empirical datasets from *erato-sara* clade of *Heliconius* butterflies and great apes to infer the variation of window sizes along each chromosome alignment, and assess how this variation affects the distribution of gene tree topologies of each group.

MATERIALS & METHODS

Overview

Our method works by first partitioning the genome alignment into non-overlapping windows of a fixed size. We then iteratively refine this partitioning by either splitting each window into two smaller sub-windows or merging two neighbouring windows together, keeping the split and/or merge only if it improves the AIC of the gene trees. This approach generates a partitioning scheme in which the window size can be longer in some regions due to merging, and shorter in other regions due to splitting. In order to evaluate if using variable window sizes improves the accuracy of non-overlapping windows to recover the true histories from the alignment (as compared to using only one fixed window size), we ran our splitting-and-

merging approach on simulated alignments with varying recombination rates and measured the accuracy of the final sets of window sizes using the two metrics used in Ivan et al. (2025): site accuracy and RMSE. We then applied our method on the genomes of *erato-sara* clade of *Heliconius* butterflies (Edelman et al., 2019) and great apes (Waterson et al. 2005; Locke et al. 2011; Scally et al. 2012; Schneider et al. 2017). All the code necessary to reproduce the methods in this paper is available at https://github.com/jeremiasivan/SimNOW.

Simulating Chromosomes with Varying Recombination Rates

We designed seven scenarios to simulate 10Mb alignments: four with homogeneous (i.e., constant) recombination rates and three with heterogeneous recombination rates (Fig. 2). We used ms (Hudson 2002) to simulate the gene trees and AliSim (Ly-Trong et al. 2022) to simulate the alignments in every scenario. The four scenarios with homogeneous recombination rates (Fig. 2) were taken directly from Ivan et al. (2025). The three scenarios with heterogenous recombination rates comprise: (i) *Scenario 1* (Fig. 2), a simple artificial scenario in which the first half of the 10Mb chromosome was simulated with no recombination, and the second half was simulated with an intermediate recombination rate of ρ=100; (ii) *Scenario 2* (Fig. 2) crudely mimics a pattern of recombination often seen in empirical datasets where the recombination rates are higher towards the ends of the chromosomes and lower towards the center, in which 2.5Mb at each end of the chromosome was simulated with ρ=50, and the central 5Mb region was simulated with ρ=0; (iii) *Scenario 3* (Fig. 2) mimics the same scenario, but with a 1Mb region at each end of the chromosome was simulated with ρ=200, the next 2.5Mb region at each end simulated with ρ=50, and the central 3Mb of the chromosome simulated with ρ=0.

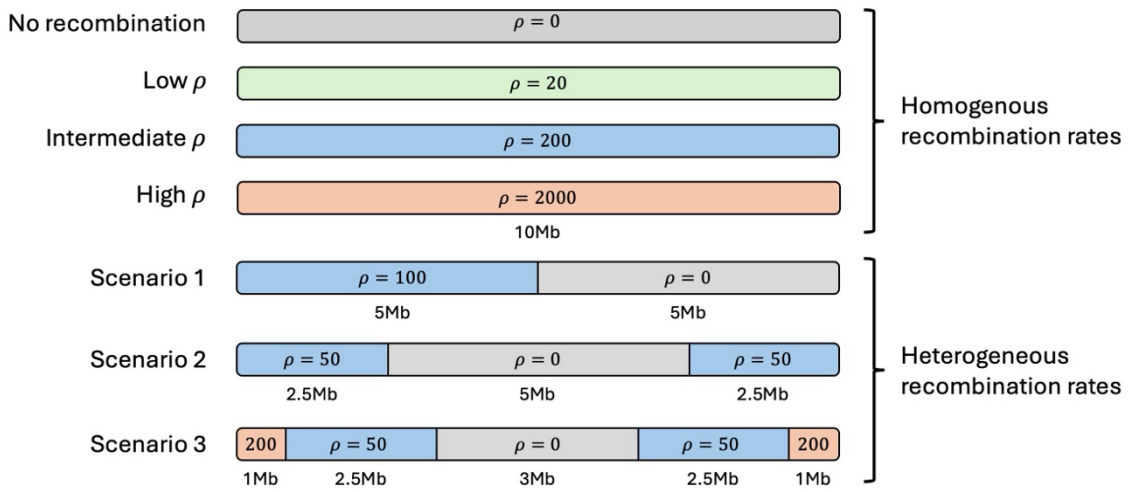

Figure 2. *Scenarios for simulating chromosome alignments. Simulated chromosomes with homogenous recombination rates were derived from Ivan et al. (2025), while ones with heterogeneous recombination rates were generated in this study. Gray: zero recombination rate; green: low recombination rate; blue: intermediate recombination rate; red: high recombination rate.*

For each simulation, we simulated gene trees following the ms (Hudson 2002) parameters used in Ivan et al. (2025) to mimic the evolutionary histories between seven *Heliconius* butterflies with medium level of ILS. Then, we used these sets of gene trees as input for AliSim (Ly-Trong et al. 2022) to simulate alignments under the Jukes-Cantor (JC) model (Jukes and Cantor 1969) for computational efficiency. Note that we multiplied the coalescent branch lengths used in ms by a factor of 0.005 to produce branch lengths in substitutions per site used in AliSim. This scaling factor was shown previously to produce alignments that closely mimic empirical alignments, with roughly 5% informative sites (Ivan et al. 2025). We did ten independent replicates for each scenario, resulting in 70 simulated chromosomes (i.e., 10 replicates for each of the 7 simulated chromosomes depicted in Figure 2).

Running Non-overlapping Window Analyses with Fixed Window Sizes

We ran non-overlapping window analysis on each simulated chromosome with 16 different fixed window sizes, ranging from 100bp to 10Mb. Then, we built the gene tree for each window using IQ-TREE v2.4.0 (Minh et al. 2020) with the following command: `iqtree2 -S alndir -m JC -blmin 1/window_size`, where `-S alndir` refers to the directory storing all window alignments, `-m JC` sets the substitution model to be JC (Jukes and Cantor 1969), and `-blmin 1/window_size` sets the minimum branch length for all trees to represent at least one substitution per branch, penalising gene trees that were inferred from windows with limited phylogenetic signal (Ivan et al. 2025).

For each simulated chromosome, we then calculated the accuracy of each window size using site accuracy (i.e., the proportion of sites from non-overlapping window analysis that correctly recover the true simulated topology) and the RMSE of the gene tree distribution from the non-overlapping window analysis compared to the simulated truth (Ivan et al. 2025). As Ivan et al. (2025) proposed that the AIC is a good predictor of both accuracy metrics, we also retrieved the AIC of each window size from each simulation and calculated the loss of accuracy if we always chose the window size with the lowest AIC score (compared to the window sizes with the highest site accuracy and/or lowest RMSE). To avoid unnecessary computation, we retrieved the results for the 40 simulated chromosomes with homogenous recombination rates from Ivan et al. (2025) and did not redo these analyses.

Running the Splitting-and-Merging Approach on Simulated Datasets

The splitting-and-merging approach requires a starting window size to be chosen by the user. We evaluated two starting window sizes for this study: (i) the best fixed window size selected by the AIC (Ivan et al. 2025) for each simulated chromosome, and (ii) a single window

comprising the entire chromosome (10Mb in this case). These relatively extreme values represent two ends of a spectrum – the former tends to be very small as it should account for the short non-recombining loci in the chromosome, while the latter is the largest possible starting value. Applying both allows us to evaluate the sensitivity of the splitting-and-merging approach to the choice of starting window size.

Once a starting window size has been selected, the splitting-and-merging approach proceeds as follows for each simulated chromosome. Every initial window is divided into two sub-windows using three size ratios: 0.25:0.75, 0.50:0.50, and 0.75:0.25 (e.g., if the initial window size is 1kb and the ratio is 0.25:0.75, then the first sub-window would be 250bp while the second sub-window would be 750bp). If the initial window was not divisible by the given ratio, the length of the first sub-window was rounded up (e.g., if the initial window size is 250bp and the ratio is 0.25:0.75, then the first sub-window would be 63bp while the second sub-window would be 187bp). Next, we inferred the gene trees for the original window and each of the three split ratios using IQ-TREE2 with the following command: `iqtree2 -s file_fasta -blmin 1/window_size`, where `-s` specifies the alignment file and `-blmin` sets the minimum branch length based on the window size. We also set the minimum number of parsimony informative sites per window to be equal to the total number of branches on the (unrooted) tree, ensuring, in principle, that each branch is supported by at least one informative site. If either of the sub-windows from a split could not be analysed (e.g., due to missing data or lack of parsimony informative sites), we skipped that splitting configuration. We then compared the AIC scores of the resulting four partitioning schemes (i.e., the original unsplit window and up to three split configurations) for each window and kept the scheme with the lowest AIC for the next iteration. After applying this approach to

every window in the initial alignment, we iteratively repeated the splitting procedure until no further splits led to an improvement in the AIC.

Following this, we used the AIC to guide a procedure in which neighbouring windows are iteratively merged. For each pair of neighbouring windows in the chromosome, we first calculated the change in AIC that would result from merging them into a single window. To do this, we concatenated their alignments using SeqKit v2.4.0 (Shen et al. 2016) and inferred the gene tree of the newly-merged window using IQ-TREE2 with the same parameters as above. We then extracted the AIC of the merged window and compared it with the total AIC of the two original (unmerged) windows. This procedure provided a list of neighbouring windows whose merging would improve the AIC score. However, not all these merging events were necessarily mutually compatible, because the AIC could be improved by merging the same window with the neighbouring window to the left *and* to the right of it. To resolve these conflicts, we used a greedy algorithm that prioritises merges with the biggest AIC improvements. The algorithm proceeded by merging the pair of windows with the largest AIC improvement, then removing the pair and all conflicting merges from the list of merges that could improve the AIC score. This procedure was iterated until all possible merges that improved the AIC score were exhausted. We then iteratively repeated the entire merging step until we could find no pair of neighbouring windows whose merging would improve the AIC score.

We measured the accuracy of the final partitioning scheme for each simulated chromosome using site accuracy and RMSE (Ivan et al. 2025), and assessed if using variable window sizes improved the two accuracy measures of non-overlapping window analysis compared to using

a single fixed window size. This method is available in the GitHub repository under https://github.com/jeremiasivan/SimNOW/tree/main/codes/3_variable_window_size/.

Empirical Analyses

Erato-sara clade of Heliconius butterflies

We downloaded the genome alignments from six species of *erato-sara* clade of *Heliconius* butterflies (with *H. melpomene* outgroup) from Edelman et al. (2019) following the steps described in Ivan et al. (2025). While the best window sizes for this group ranged between ≤125-250bp according to the stepwise non-overlapping window approach (Ivan et al. 2025), we found that using very short window sizes as the starting window sizes could lead to low accuracy (see Results). Thus, we ran the splitting-and-merging approach on each of the 21 chromosome alignments with the whole chromosome (i.e., full concatenation) as the starting window size.

After no further AIC improvement could be achieved from either splitting or merging, we extracted the final set of window sizes for each chromosome and built the gene tree of each window using IQ-TREE2 with the following command: `iqtree2 -s file_fasta -blmin 1/window_size -bb 1000`, where `-bb` reflects the number of UFBoot replicates (Hoang et al. 2018). We rebuilt the gene trees to extract the bootstrap support for each gene tree. We then retrieved the topology of every gene tree, and compared the proportion of each topology across the chromosomes to those from Edelman et al. (2019) and Ivan et al. (2025).

Great apes

We downloaded genome alignments of four great apes (human, chimpanzee, gorilla, and orangutan) from UCSC Genome Browser (Kent et al. 2002) following the steps described in Ivan et al. (2025). Similar to the analyses on the *Heliconius* butterflies, we ran the splitting-and-merging approach on each of the 25 chromosomes (22 autosomes, chromosome X, chromosome Y, and mitochondrial DNA) but with a minimum of 5 parsimony informative sites per window, using the entire chromosome as the starting point. We then built the gene trees based on the final partitioning scheme with 1,000 UFBoot replicates using IQ-TREE2 and calculated the topology distribution of the group across the chromosomes.

RESULTS

Variable window sizes outperform fixed window sizes on simulated datasets

For simulated chromosomes with zero recombination rate, non-overlapping window analyses using either fixed window sizes (from Ivan et al. 2025) or variable window sizes (from the splitting-and-merging approach proposed in this study) always prefer full concatenation, as expected (Fig. 3-4). However, as variable window sizes consistently outperform fixed window sizes in any other scenario (see below), we recommend to always use variable window sizes when running non-overlapping window approach. For simulations with non-zero homogeneous recombination rates, the variable window size approach increases the site accuracy of non-overlapping window method by on average 3.0-4.7% across scenarios compared to using fixed window sizes from the AIC (Fig. 3), and decreases the RMSE by on average 0.05-0.35% across scenarios (except for $\rho = 2000$ with on average 0.14% RMSE increase; Fig. 4). For simulated chromosomes with heterogeneous recombination rates, the variable window size approach also outperforms the fixed window size method. Using variable window sizes increases the site accuracy of non-overlapping window method by on

average 4.7-8.3% across scenarios compared to using fixed window sizes from the AIC (Fig. 3), and decreases the RMSE by on average 0.57-0.81% (Fig. 4). In one replicate from Scenario 3, running the variable window size approach results in an 22.3% site accuracy increase (Fig. 3) and 4.1% RMSE decrease (Fig. 4) compared to using a single fixed window size. This trend is also reflected by the AIC scores, where partitioning schemes from the variable window size approach have better AIC than the fixed window sizes (Fig. S1-S2).

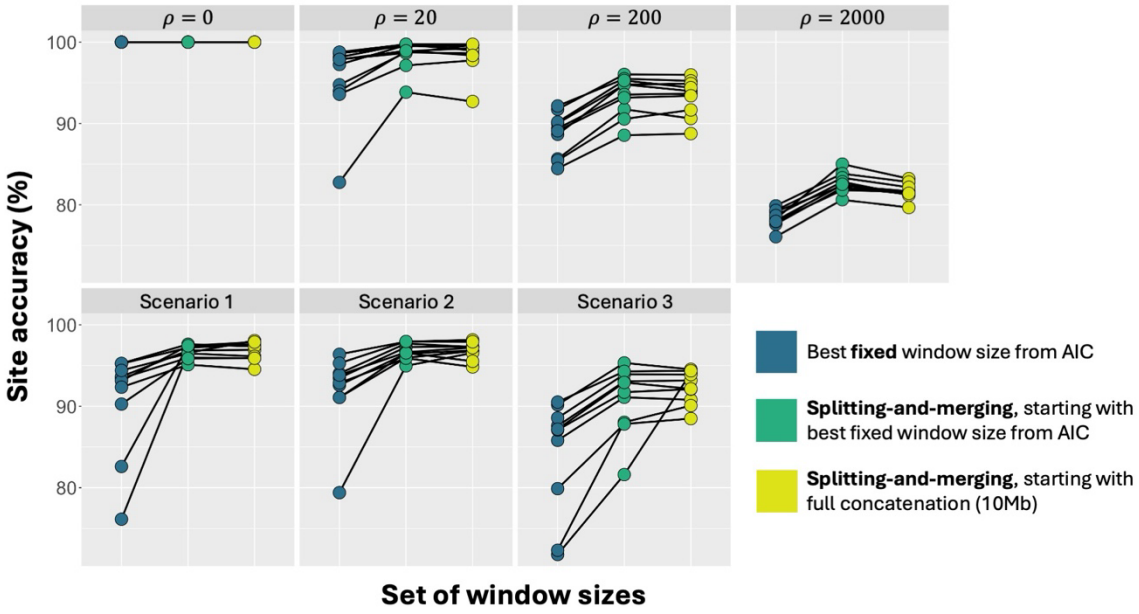

Figure 3. *Site accuracy from running the non-overlapping window analyses with fixed and variable window sizes across 7 simulation scenarios (Fig. 2). Colour denotes the set of window sizes used when running the non-overlapping window method, either using a fixed window size (Ivan et al. 2025) or the splitting-and-merging approach with different starting window sizes. Black lines connect the same replicates.*

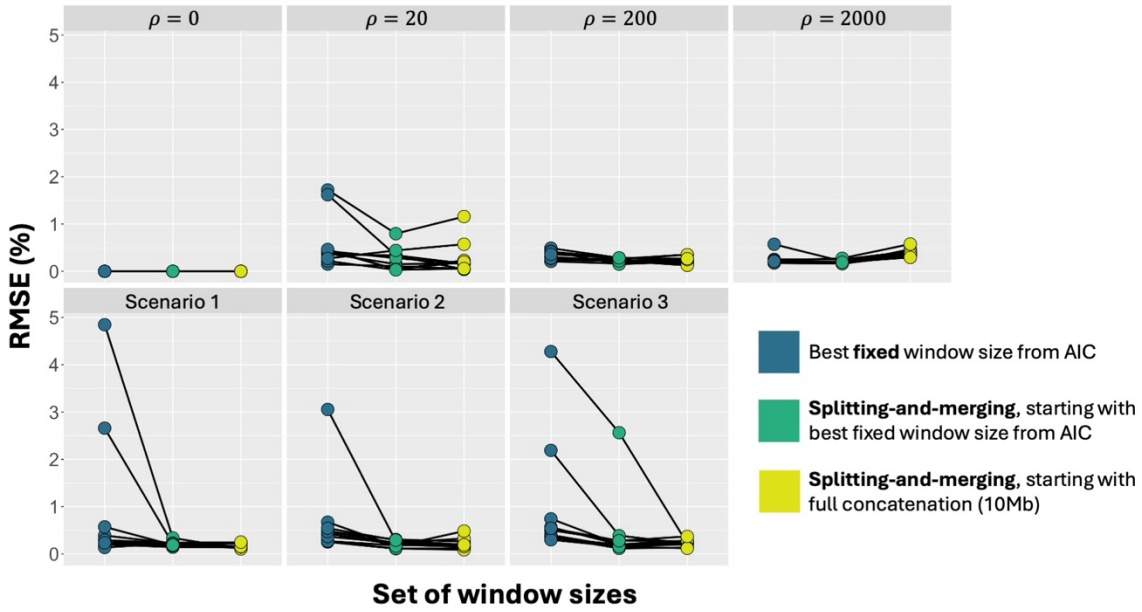

Figure 4. *RMSE from running the non-overlapping window analyses with fixed and variable window sizes across 7 simulation scenarios (Fig. 2). Colour denotes the set of window sizes used when running the non-overlapping window method, either using a fixed window size (Ivan et al. 2025) or the splitting-and-merging approach with different starting window sizes. Black lines connect the same replicates.*

Full concatenation is preferable as starting window size for splitting-and-merging approach

In general, our simulations show that there is very little difference in accuracy when using different starting window sizes for the splitting-and-merging algorithm. On average, the two methods differ in site accuracy by 0.1-1.4% (Fig. 3) and in RMSE by 0.01-0.23% (Fig. 4). However, there is one simulated chromosome from Scenario 3 where using full concatenation as the starting window size has 12.48% higher site accuracy and 2.34% lower RMSE compared to using the fixed window sizes from AIC. This suggests that in certain cases, starting the splitting-and-merging algorithm with a small window size can lead it to become stuck in a local optimum. We also found that starting the splitting-and-merging approach with short windows often fails to recover the long chromosomal regions with zero recombination

346 rate as a single window (Fig. S3-S5 (left), Table S1). This is improved when starting the
347 algorithm with the whole chromosome as a single window (Fig. S3-S5 (right)). Additionally,
348 starting with a single concatenated window has the advantage of being far more
349 computationally efficient, as no prior computation needs to be done to choose the starting
350 window size. Because of this, we used this approach on the empirical datasets whose analyses
351 we describe in the following sections.
352

353 ***Fixed window sizes do not represent the size variation of non-recombining loci across***
354 ***empirical chromosomes***

355 As the stepwise non-overlapping window method from Ivan et al. (2025) excluded windows
356 that cannot be analysed due to missing data, we could not compare the AIC of the
357 partitioning schemes between the fixed window sizes and the variable window sizes. For
358 *erato-sara* clade of *Heliconius* butterflies, the variable window size approach produces
359 substantially different window size results compared to the fixed window size method of Ivan
360 et al. (2025). Using the variable window size method, the shortest window on each
361 chromosome ranges from 18bp to 40bp, while the longest window ranges from 31kb to
362 106kb (Table S1; Fig. S6). Based on the distributions of window sizes across chromosomes,
363 the best fixed window sizes selected by the stepwise non-overlapping window method (Ivan
364 et al. 2025) are 2-5x shorter than the average window sizes from the variable window size
365 approach, and much closer the first quartile of the distribution than the median (Fig. S7;
366 Table S1). Despite these differences, the ten most common topologies recovered from the
367 variable window size approach are consistent with those from previous studies (Edelman et
368 al. 2019; Ivan et al. 2025), accounting for 53.0% of the sites (or 40.8% of the windows) when
369 including all gene trees (Fig. 5, S8). The most common topology is T2 (which is regarded as
370 the species tree topology (Edelman et al. 2019)), followed by T1 and T3. When we filtered

out gene trees with <95 average UFBoot support, the top ten topologies account for 97.9% of the sites (or 92.9% of the windows) from the highly-supported windows (Fig. 5, S8), with T1 as the most common topology. Our results highlight that the distribution of tree topologies can differ substantially depending on whether it is quantified by individual windows or individual sites. For example, the species tree topology (T2) that dominates the sex chromosome (chr21) is only recovered by 15.37% of windows on that chromosome (followed by 12.68% of T3 and 12.19% of T7; Table S2). However, these windows account for 26.85% of the sites (followed by 18.44% of T3 and 13.16% of T7; Table S3), reflecting that windows that are associated with these topologies tend to be longer than average.

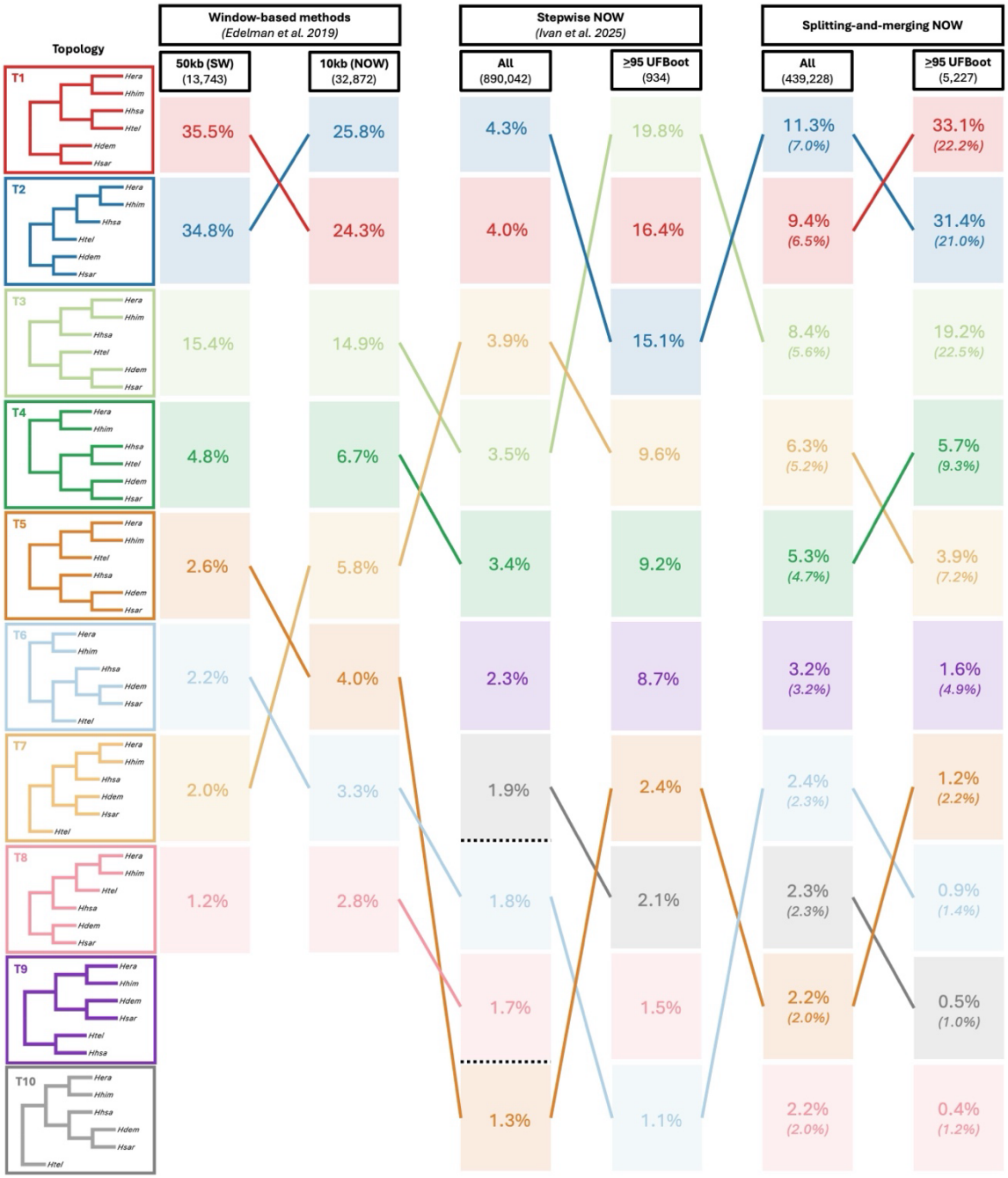

Figure 5. *Comparison of topology frequency of the most common tree topologies from the genomes of erato-sara clade of Heliconius between 50kb sliding windows (SW) and 10kb non-overlapping windows (NOW) from Edelman et al. (2019), stepwise NOW from Ivan et al. (2025), and the splitting-and-merging approach from this study. For splitting-and-merging approach, the percentages reflect the proportions of sites that support each topology, while the numbers in bracket reflect the proportions of windows that support each topology. Topologies*

are rooted with H. melpomene (not shown) as the outgroup. The number under each method reflects the total number of windows. Black dotted lines refer to other topologies not listed.

For the genomes of great apes, the shortest window ranges from 17bp to 141bp across chromosomes, while the longest window ranges from 138kb to 4.5Mb, except for chrY (chromosome Y) and mtDNA (mitochondrial DNA; Table S6; Fig. S9). For chrY, a major proportion of the alignment was not split due to missing data (Ivan et al. 2025), driving the average window size to reach almost 200kb. On the other hand, the 16kb-long mtDNA was partitioned into 8 windows, with an average window size of 2.1kb (Table S6). Similar to the results from the *Heliconius* butterflies, the best fixed window sizes from the stepwise non-overlapping window method (Ivan et al. 2025) are 4-11x shorter than the average window sizes from the splitting-and-merging approach (Fig. S10). When we consider all gene trees across all chromosomes, the major topology (grouping human and chimpanzee together) is supported by 78.9% of the sites (or 75.3% of the windows), while the minor topologies account for 10.8% and 10.3% of the sites (or 12.5% and 12.2% of the windows), respectively (Fig. 6, S11; Table S7-S8). This contrasts with the topology distribution of 60.7%, 19.8%, and 19.2% when using the fixed window size approach from Ivan et al. (2025) (Fig. 6). Similar to the results from Ivan et al. (2025), filtering out gene trees with <95 average UFBoot support drives the proportion of the major topology to approximately 92%, while the minor topologies account for 4.7% and 3.7% of the sites, respectively (Fig. 6; Table S7-S8). Interestingly, for mtDNA, six of the eight windows recover the major topology (comprising 80.08% of the sites), while the other two windows result in each of the minor topologies (Table S7-S8; Fig. S11). This variation of gene trees in mtDNA is likely caused systematic error rather than biological processes (Richards et al. 2018), and serves as a useful reminder of some of the limitations of current phylogenetic inference methods.

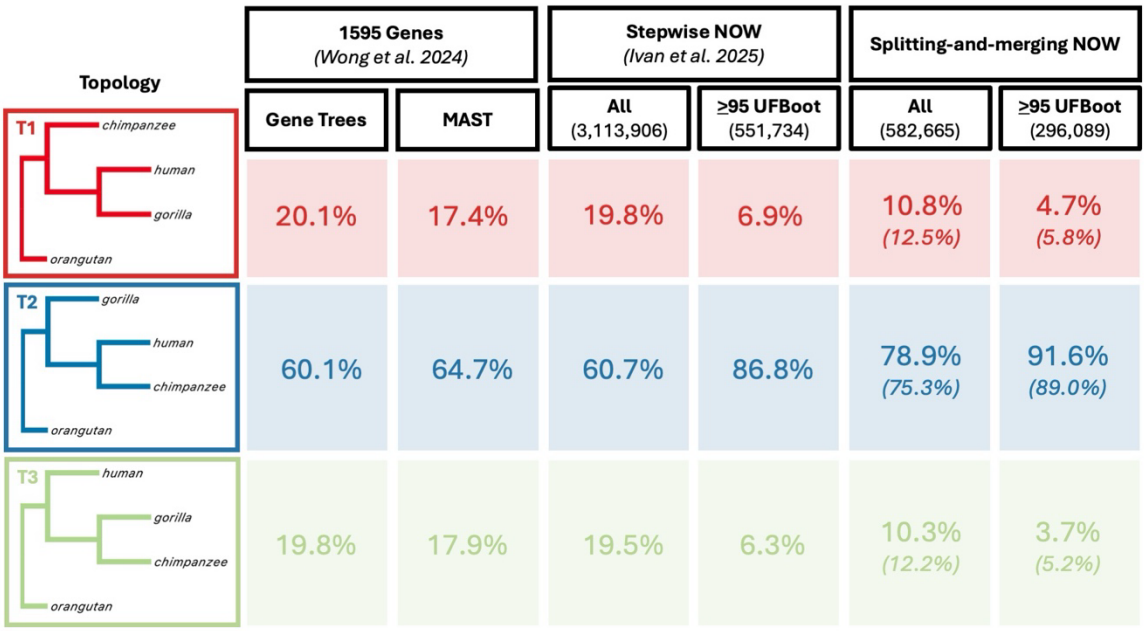

Figure 6. *Comparison of unrooted tree topology frequency from the genomes of great apes using different approaches: single gene tree inference and MAST (Wong et al. 2024), the stepwise non-overlapping windows (NOW) that includes all gene trees and only highly-supported trees (Ivan et al. 2025), and the splitting-and-merging approach from this study. For splitting-and-merging approach, the percentages reflect the proportions of sites that support each topology, while the numbers in bracket reflect the proportions of windows that support each topology. The number under each method reflects the total number of windows. Topologies are visualised using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree) without branch lengths.*

Variable window sizes can identify regions of biological importance

Similar to Edelman et al. (2019), we found that the proportion of T1 on the chromosomes of *Heliconius* butterflies generally goes down as the chromosome becomes longer (R^2=0.743, ρ<0.01 for all gene trees; and R^2=0.520, ρ<0.01 for highly-supported trees), while T2 is positively correlated with the chromosome length (R^2=0.712, ρ<0.01 for all gene trees; and

R^2=0.392, ρ<0.01 for highly-supported trees) (Fig. S12). When each chromosome was divided into 10 equally-sized bins along its length, we also observed that T1 has lower proportion on the ends of the chromosomes, in contrast to T2 that shows the opposite pattern (Fig. S12). Moreover, we identified two genomic regions that are dominated by the minor topologies (i.e., neither T1 nor T2; Fig. 7). The first region spans nearly 2Mb on chromosome 2 and is dominated by T3 (Fig. 7a), while the second region comprises 376 windows and is dominated by T4, covering approximately 400kb on chromosome 15 (Fig. 7b). The latter is of particular interest as it spans the loci that control the wing colour pattern in *Heliconius* butterflies (Edelman et al. 2019). To assess whether the region is best fit by a single tree topology (as our splitting-and-merging approach recovered multiple different topologies along this region despite being dominated by the T4 topology), we analysed the whole ~400kb region as one partition, and compared the AIC to the partitioning scheme recovered from our algorithm. This analysis shows that analysing the region as a single partition has a much worse AIC score than the partitioning scheme recovered by the variable window size method (AIC score of 1928355 compared to 1913217, a difference of 15,138 units).

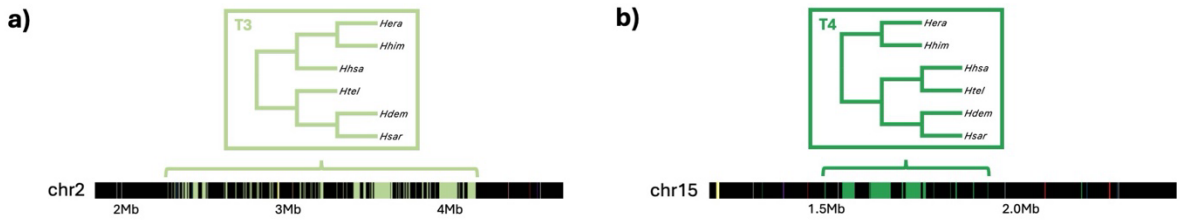

Figure 7. *Inversion topologies on (a) chromosome 2 and (b) chromosome 15 of erato-sara clade of Heliconius butterflies (Edelman et al. 2019). Colouring is based on Fig. S8.*

450 **DISCUSSION**

451 The non-overlapping window method is a common approach to infer gene trees across a set
452 of aligned chromosomes. While the selection of window size is often arbitrary, a recent study
453 by Ivan et al. (2025) showed that the total AIC score of the gene trees is a good indicator of
454 the ability of different window sizes to recover the 'true' topology of each site in the aligned
455 chromosomes. In this study, we extended this approach by allowing the window size to vary
456 across the chromosomes. We achieved this by iteratively splitting each window and merging
457 neighbouring windows until the total AIC of the whole alignment stops improving from any
458 new split and/or merge (Fig. 1). Our results showed that using variable window sizes from
459 the splitting-and-merging algorithm had better accuracy in recovering the 'true' gene tree
460 topology from each site of simulated datasets compared to using fixed window sizes. The
461 fixed window sizes selected by the AIC had at least 70% site accuracy across simulated
462 chromosomes with homogeneous and heterogenous recombination rates (Fig. 3-4). Using
463 variable window sizes consistently increased this accuracy further by on average 3-8% across
464 scenarios (Fig. 3), reflecting that variable window sizes better fit the size variation of non-
465 recombining loci across the chromosomes (Fig. S3-S5). While the different starting window
466 sizes tested in this study did not significantly affect the accuracy of the splitting-and-merging
467 approach (Fig. 3-4), we proposed future phylogenomic analyses to start with full
468 concatenation compared to the best window size from the AIC (Ivan et al. 2025) because: (i)
469 we found one simulation with low accuracy when starting with short window size (Fig. 3),
470 which might be caused by initial over-splitting of the chromosome alignment that makes it
471 harder to improve the AIC via merging (i.e., if a non-recombining locus is split into multiple
472 windows, merging two neighbouring windows might not improve the AIC due to lack of
473 phylogenetic signal, as opposed to merging *all* windows together); (ii) it better captured the
474 long non-recombining loci from the simulated chromosomes as splitting these regions would

make the AIC worse (Fig. S3-S5); and (iii) it is computationally more efficient because we do not have to test and compare the AIC of various fixed window sizes for the given alignment.

When we ran the splitting-and-merging algorithm on the six species from *erato-sara* clade of *Heliconius* butterflies (with *H. melpomene* outgroup), we found that the window sizes varied substantially across the chromosomes (Fig. S6-S7). For instance, the window sizes on chromosome 1 ranged from 29bp to 70kb with an average of 581±11bp (95% CI), while the best fixed window size from the stepwise non-overlapping window method (Ivan et al. 2025) was 250bp (Table S1). This best fixed window size was ~10x longer than the smallest window – which prompted the concatenation of multiple windows during gene tree inference – but 280x smaller than the biggest window – which might increase gene tree estimation error during gene tree inference. As both concatenation and gene tree estimation error might bias tree inferences (Ivan et al. 2025), this result further highlights the importance of using variable window sizes when running non-overlapping window method on chromosome alignments. Unfortunately, we cannot compare the AIC between the two methods, because the stepwise non-overlapping window method (Ivan et al. 2025) excluded windows that cannot be analysed while the splitting-and-merging approach keep them in the analysis. Interestingly, the topology distribution of the group is generally consistent across studies when considering the top ten most common topologies (Fig. 5), with either T1 or T2 as the most common topology depending on the methods being used. This might infer that the gene tree inference for this group is quite robust to the selection of window sizes. However, we note that the *Heliconius* chromosomes from Edelman et al. (2019) were not phased. Huang et al. (2022) demonstrated that random phasing can bias species tree inference, whereas allelic phasing may improve this inference (Andermann et al. 2019; but see also Kates et al. 2018). We hypothesise that running the splitting-and-merging approach on phased chromosomes

would generally result in shorter windows (particularly in highly heterozygous regions) because our approach would also consider the allelic variation derived from phasing. This observation highlights the importance of using variable window sizes on phased genome alignments whenever possible to accurately identify recombination breakpoints across the genomes.

Similar to the results from *Heliconius* butterflies, running the splitting-and-merging approach on the genomes of great apes showed that the window sizes vary substantially across the chromosomes (Fig. S9-S10). Interestingly, the mtDNA was partitioned into eight windows (Table S6), where six of the windows recovered the major topology (grouping human and chimpanzee together) while the other two windows recovered the two minor topologies (Fig. S11). Gene tree discordance in mitochondrial genomes has been observed in different groups, such as tetrapods (Toups et al. 2025), branchiopods (Sun and Cheng 2023), and new world monkeys (Finnegan et al. 2025). In our case, the first window with the minor topology is 3.1kb-long and spans the 12S and 16S ribosomal RNA genes. A recent study by DeSalle and Tessler (2025) observed that the two rRNA genes had low degree of phylogenetic congruence with the other genes in the mtDNA genomes across 89 animal taxonomic groups, which likely resulted in different tree topology. On the other hand, the second window is part of the mtDNA control region. While previous studies have shown the utility of this region as a phylogenetic marker (e.g., Randi et al. 2001; Bronstein et al. 2018), this window only contains six parsimony informative sites in our four-taxon dataset, which is likely not enough to infer a robust tree.

Across the great apes chromosomes, the estimated proportion of the major topology that groups human and chimpanzee together (~79% of the sites) was generally higher than

previous estimates that ranged between 60-77% (Fig. 6; Ebersberger et al. 2007; Vanderpool et al. 2020; Wong et al. 2024; Ivan et al. 2025). There are three potential explanations for the differences in the estimates. First, both window-based and gene-based inferences are likely to be affected by trade-offs between concatenation and gene tree estimation error (Ivan et al. 2025). While our method should more accurately reflect chromosomal recombination breakpoints – and thus provides a more reliable estimate of topology distribution than gene-based or fixed-window approaches – it is not entirely free from these competing factors. For instance, we did not split a window if it cannot be analysed due to missing data or an insufficient number of parsimony informative sites (as clearly shown for chromosome Y; Table S6), which might concatenate multiple non-recombining loci together and favour the topology that is supported by the majority of sites. In contrast, methods like MAST assign the tree for individual sites, allowing weaker incongruent signals to be accounted in the final estimates. Second, our analyses were based on whole genome alignments that include both coding and non-coding regions. In great apes, coding regions were shown to undergo stronger selection than non-coding regions (Cagan et al. 2016). Under the Hill-Robertson effect (Hill and Robertson 1966), regions with stronger selection tend to have reduced effective population size (N_e) (Comeron et al. 2008), which correlates with shortened coalescent time and increased gene tree discordance. However, we cannot fully evaluate this explanation because we did not explicitly differentiate the estimates between the coding and non-coding regions. Third, Reddy et al. (2017) proposed that the current substitution models might be inadequate to explain the evolution of coding regions, and potentially better-fit the non-coding regions. As model violations have been shown to affect phylogenetic analysis (Naser-Khdour et al. 2019), it is possible that analyses of coding regions – on which most previous estimates were based – might underestimate the proportion of the major topology due to higher level of discordance caused by model misspecification. In order to assess the relative

550 contributions of these explanations to our results, further studies should incorporate these
551 factors into the analyses (e.g., running MAST on whole genome alignments and estimating
552 the topological proportions separately for coding and non-coding regions).
553
554 Using variable window sizes, we further confirmed the findings from Edelman et al. (2019),
555 where the proportion of T1 was higher on shorter chromosomes, while T2 was more common
556 on longer chromosomes (Fig. S12 (left)). Martin et al. (2019) showed that chromosome
557 length is negatively correlated with recombination rate in *Heliconius* butterflies (i.e., shorter
558 chromosomes have higher recombination rates). On the other hand, recombination rate is also
559 correlated with the estimated admixture proportion between the three focal *Heliconius*
560 species in that study (Martin et al. 2019). This pattern arises because genomic regions with
561 low recombination rate tend to exhibit increased linkage among loci, which can act as a
562 barrier to gene flow (Barton and Bengtsson 1986; Aeschbacher et al. 2017). In our dataset, T2
563 is the regarded as the species tree topology, while T1 reflects hybridisation between *H.*
564 *hecalesia* and *H. telesiphe* (Edelman et al. 2019). As a result, T1 is enriched on shorter
565 chromosomes with higher recombination rates and weaker barrier to hybridisation, while T2
566 is enriched on longer chromosomes with lower recombination rates and stronger barrier to
567 gene flow. A similar pattern was observed along individual chromosomes (Fig. S12 (right)):
568 T1 was more frequent towards chromosomal ends (where recombination rates are higher),
569 while T2 had higher proportion towards the chromosomal centers (where recombination rates
570 are lower; Martin et al. 2019). We also identified two genomic regions that are dominated by
571 the minor topologies (Fig. 7). The first region spanned approximately 2Mb on chromosome 2
572 and was dominated by T3, likely caused by ILS and ancestral polymorphism (Edelman et al.
573 2019). The second region was 400Mb-long and spanned the *cortex* locus on chromosome 15,
574 which controls the wing colour patterns in *Heliconius* butterflies (Joron et al. 2006; Nadeau et

575 al. 2016; Jay et al. 2018). This region reflects an inversion in *H. demeter*, *H. sara*, *H.*
576 *telesiphe*, and *H. hecalesia*, but not in *H. erato*, *H. himera*, and the outgroup *H. melpomene*
577 (Edelman et al. 2019). As a result, *H. telesiphe* and *H. hecalesia* formed a sister pair, and
578 together become the sister clade of *H. demeter* and *H. sara* (T4; Edelman et al. 2019).
579
580 Overall, our AIC-based splitting-and-merging approach provides a method that uses variable
581 window sizes to infer gene trees from whole genome alignments. The windows (i.e., loci)
582 selected by our method should better reflect recombination breakpoints and thus serve as
583 better approximations of *c*-genes than simply choosing a fixed window size, or selecting loci
584 based on known genomic boundaries such as exons and introns. These loci can then be used
585 in a number of downstream applications. For example, if the aim is to reconstruct different
586 species trees across a genome alignment, loci from our method could be used as the starting
587 point for coalescent-based approaches such as those used by Thawornwattana et al. (2018,
588 2022). If the aim is to infer the evolutionary history of every site in the genome (as may be
589 required for applications such as PhyloGWAS (Pease et al. 2016)), then gene trees could be
590 estimated for each locus. If stochastic error is a concern, gene trees inferred from these loci
591 could be further refined using more complex substitution models that better-explain empirical
592 datasets (e.g., GHOST (Crotty et al. 2020) that models heterotachy, MixtureFinder (Ren et al.
593 2025) that models variation of transition rates among sites), and/or more sophisticated
594 methods that estimate the trees for individual sites (e.g., MAST (Wong et al. 2024)) or refine
595 the gene trees in relation to each other (e.g., Espalier (Rasmussen and Guo 2023), HMMs
596 (Siepel and Haussler 2005)). In all these cases, more accurate approximations of *c*-genes can
597 benefit downstream analyses, and have the potential to improve our understanding of
598 complex genomic histories.
599

SUPPLEMENTARY MATERIALS

Supplementary material will be available online upon the acceptance of the paper.

ACKNOWLEDGEMENTS

We thank Minh Bui, Justin Borevitz, the members of Lanfear and Bui Lab from Australian National University, as well as Matt Hahn from Indiana University, for comments on the study design and the manuscript.

FUNDING

This research is supported by Australian Government Research Training Program (AGRTP) scholarship via The Australian National University Higher Degree by Research (HDR) program.

REFERENCES

Aeschbacher S., Selby J.P., Willis J.H., Coop G. 2017. Population-genomic inference of the strength and timing of selection against gene flow. Proceedings of the National Academy of Sciences. 114:7061–7066.

Akaike H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 19:716–723.

Andermann T., Fernandes A.M., Olsson U., Töpel M., Pfeil B., Oxelman B., Aleixo A., Faircloth B.C., Antonelli A. 2019. Allele Phasing Greatly Improves the Phylogenetic Utility of Ultraconserved Elements. Systematic Biology. 68:32–46.

Barton N., Bengtsson B.O. 1986. The barrier to genetic exchange between hybridising populations. Heredity. 57:357–376.

Brandt D.Y.C., Huber C.D., Chiang C.W.K., Ortega-Del Vecchyo D. 2024. The Promise of Inferring the Past Using the Ancestral Recombination Graph. Genome Biology and Evolution. 16:evae005.

Bronstein O., Kroh A., Haring E. 2018. Mind the gap! The mitochondrial control region and its power as a phylogenetic marker in echinoids. BMC Evol Biol. 18:80.

Cagan A., Theunert C., Laayouni H., Santpere G., Pybus M., Casals F., Prüfer K., Navarro A., Marques-Bonet T., Bertranpetit J., Andrés A.M. 2016. Natural Selection in the Great Apes. Mol Biol Evol. 33:3268–3283.

Chan A.H., Jenkins P.A., Song Y.S. 2012. Genome-Wide Fine-Scale Recombination Rate Variation in Drosophila melanogaster. PLOS Genetics. 8:e1003090.

Comeron J.M., Williford A., Kliman R.M. 2008. The Hill–Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. Heredity. 100:19–31.

Crotty S.M., Minh B.Q., Bean N.G., Holland B.R., Tuke J., Jermiin L.S., Haeseler A.V. 2020. GHOST: Recovering Historical Signal from Heterotachously Evolved Sequence Alignments. Syst Biol. 69:249–264.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol Evol. 24:332–340.

DeSalle R., Tessler M. 2025. Mitochondrial Gene Phylogenetic Incongruencies Are Linked to Chromosomal Position and Function. Genome Biol Evol. 17:evaf209.

Ebersberger I., Galgoczy P., Taudien S., Taenzer S., Platzer M., von Haeseler A. 2007. Mapping Human Genetic Ancestry. Molecular Biology and Evolution. 24:2266–2276.

Edelman N.B., Frandsen P.B., Miyagi M., Clavijo B., Davey J., Dikow R.B., García-Accinelli G., Van Belleghem S.M., Patterson N., Neafsey D.E., Challis R., Kumar S., Moreira G.R.P., Salazar C., Chouteau M., Counterman B.A., Papa R., Blaxter M.,

Reed R.D., Dasmahapatra K.K., Kronforst M., Joron M., Jiggins C.D., McMillan W.O., Di Palma F., Blumberg A.J., Wakeley J., Jaffe D., Mallet J. 2019. Genomic architecture and introgression shape a butterfly radiation. Science. 366:594–599.

Edwards S.V., Xi Z., Janke A., Faircloth B.C., McCormack J.E., Glenn T.C., Zhong B., Wu S., Lemmon E.M., Lemmon A.R., Leaché A.D., Liu L., Davis C.C. 2016. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. Molecular Phylogenetics and Evolution. 94:447–462.

Feng C., Wang J., Liston A., Kang M. 2023. Recombination Variation Shapes Phylogeny and Introgression in Wild Diploid Strawberries. Molecular Biology and Evolution. 40:msad049.

Finnegan N., Lima M.G.M., Lynch J.W. 2025. Mitochondrial DNA for Phylogeny Building: Assessing Individual and Grouped mtGenes as Proxies for the mtGenome in Platyrrhines. American Journal of Primatology. 87:e70017.

Flouri T., Jiao X., Rannala B., Yang Z. 2018. Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. Molecular Biology and Evolution. 35:2585–2593.

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

Herrig D.K., Ridenbaugh R.D., Vertacnik K.L., Everson K.M., Sim S.B., Geib S.M., Weisrock D.W., Linnen C.R. 2024. Whole Genomes Reveal Evolutionary Relationships and Mechanisms Underlying Gene-Tree Discordance in Neodiprion Sawflies. Systematic Biology.:syae036.

674 Hibbins M.S., Breithaupt L.C., Hahn M.W. 2023. Phylogenomic comparative methods:
675 Accurate evolutionary inferences in the presence of gene tree discordance.
676 Proceedings of the National Academy of Sciences. 120:e2220389120.
677 Hill W.G., Robertson A. 1966. The effect of linkage on limits to artificial selection. Genetics
678 Research. 8:269–294.
679 Hoang D.T., Chernomor O., von Haeseler A., Minh B.Q., Vinh L.S. 2018. UFBoot2:
680 Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution.
681 35:518–522.
682 Huang J., Bennett J., Flouri T., Leaché A.D., Yang Z. 2022. Phase Resolution of
683 Heterozygous Sites in Diploid Genomes is Important to Phylogenomic Analysis under
684 the Multispecies Coalescent Model. Systematic Biology. 71:334–352.
685 Hudson R.R. 2002. Generating samples under a Wright–Fisher neutral model of genetic
686 variation. Bioinformatics. 18:337–338.
687 Ivan J., Frandsen P., Lanfear R. 2025. Selecting a Window Size for Phylogenomic Analyses
688 of Whole Genome Alignments using AIC. Syst Biol.:syaf053.
689 Jay P., Whibley A., Frézal L., Cara M.Á.R. de, Nowell R.W., Mallet J., Dasmahapatra K.K.,
690 Joron M. 2018. Supergene Evolution Triggered by the Introgression of a
691 Chromosomal Inversion. Current Biology. 28:1839-1845.e3.
692 Jensen-Seaman M.I., Furey T.S., Payseur B.A., Lu Y., Roskin K.M., Chen C.-F., Thomas
693 M.A., Haussler D., Jacob H.J. 2004. Comparative Recombination Rates in the Rat,
694 Mouse, and Human Genomes. Genome Res. 14:528–538.
695 Joron M., Papa R., Beltrán M., Chamberlain N., Mavárez J., Baxter S., Abanto M.,
696 Bermingham E., Humphray S.J., Rogers J., Beasley H., Barlow K., ffrench-Constant
697 R.H., Mallet J., McMillan W.O., Jiggins C.D. 2006. A Conserved Supergene Locus
698 Controls Colour Pattern Diversity in Heliconius Butterflies. PLOS Biology. 4:e303.

Jukes T.H., Cantor C.R. 1969. CHAPTER 24 - Evolution of Protein Molecules. In: Munro H.N., editor. Mammalian Protein Metabolism. Academic Press. p. 21–132.

Kates H.R., Johnson M.G., Gardner E.M., Zerega N.J.C., Wickett N.J. 2018. Allele phasing has minimal impact on phylogenetic reconstruction from targeted nuclear gene sequences in a case study of Artocarpus. American Journal of Botany. 105:404–416.

Kent W.J., Sugnet C.W., Furey T.S., Roskin K.M., Pringle T.H., Zahler A.M., Haussler and D. 2002. The Human Genome Browser at UCSC. Genome Res. 12:996–1006.

Kosakovsky Pond S.L., Posada D., Gravenor M.B., Woelk C.H., Frost S.D.W. 2006. Automated Phylogenetic Detection of Recombination Using a Genetic Algorithm. Molecular Biology and Evolution. 23:1891–1901.

Lescroart J., Bonilla-Sánchez A., Napolitano C., Buitrago-Torres D.L., Ramírez-Chaves H.E., Pulido-Santacruz P., Murphy W.J., Svardal H., Eizirik E. 2023. Extensive Phylogenomic Discordance and the Complex Evolutionary History of the Neotropical Cat Genus Leopardus. Molecular Biology and Evolution. 40:msad255.

Locke D.P., Hillier L.W., Warren W.C., Worley K.C., Nazareth L.V., Muzny D.M., Yang S.-P., Wang Z., Chinwalla A.T., Minx P., Mitreva M., Cook L., Delehaunty K.D., Fronick C., Schmidt H., Fulton L.A., Fulton R.S., Nelson J.O., Magrini V., Pohl C., Graves T.A., Markovic C., Cree A., Dinh H.H., Hume J., Kovar C.L., Fowler G.R., Lunter G., Meader S., Heger A., Ponting C.P., Marques-Bonet T., Alkan C., Chen L., Cheng Z., Kidd J.M., Eichler E.E., White S., Searle S., Vilella A.J., Chen Y., Flicek P., Ma J., Raney B., Suh B., Burhans R., Herrero J., Haussler D., Faria R., Fernando O., Darré F., Farré D., Gazave E., Oliva M., Navarro A., Roberto R., Capozzi O., Archidiacono N., Valle G.D., Purgato S., Rocchi M., Konkel M.K., Walker J.A., Ullmer B., Batzer M.A., Smit A.F.A., Hubley R., Casola C., Schrider D.R., Hahn M.W., Quesada V., Puente X.S., Ordoñez G.R., López-Otín C., Vinar T., Brejova B.,

724 Ratan A., Harris R.S., Miller W., Kosiol C., Lawson H.A., Taliwal V., Martins A.L.,
725 Siepel A., RoyChoudhury A., Ma X., Degenhardt J., Bustamante C.D., Gutenkunst
726 R.N., Mailund T., Dutheil J.Y., Hobolth A., Schierup M.H., Ryder O.A., Yoshinaga
727 Y., de Jong P.J., Weinstock G.M., Rogers J., Mardis E.R., Gibbs R.A., Wilson R.K.
728 2011. Comparative and demographic analysis of orang-utan genomes. Nature.
729 469:529–533.
730 Ly-Trong N., Naser-Khdour S., Lanfear R., Minh B.Q. 2022. AliSim: A Fast and Versatile
731 Phylogenetic Sequence Simulator for the Genomic Era. Molecular Biology and
732 Evolution. 39:msac092.
733 Maddison W.P. 1997. Gene Trees in Species Trees. Systematic Biology. 46:523–536.
734 Mallet J., Besansky N., Hahn M.W. 2016. How reticulated are species? BioEssays. 38:140–
735 149.
736 Martin S.H., Davey J.W., Salazar C., Jiggins C.D. 2019. Recombination rate variation shapes
737 barriers to introgression across butterfly genomes. PLOS Biology. 17:e2006288.
738 McVean G.A.T., Myers S.R., Hunt S., Deloukas P., Bentley D.R., Donnelly P. 2004. The
739 Fine-Scale Structure of Recombination Rate Variation in the Human Genome.
740 Science. 304:581–584.
741 Meleshko O., Martin M.D., Korneliussen T.S., Schröck C., Lamkowski P., Schmutz J.,
742 Healey A., Piatkowski B.T., Shaw A.J., Weston D.J., Flatberg K.I., Szövényi P.,
743 Hassel K., Stenøien H.K. 2021. Extensive Genome-Wide Phylogenetic Discordance Is
744 Due to Incomplete Lineage Sorting and Not Ongoing Introgression in a Rapidly
745 Radiated Bryophyte Genus. Molecular Biology and Evolution. 38:2750–2766.
746 Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A.,
747 Lanfear R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic
748 Inference in the Genomic Era. Molecular Biology and Evolution. 37:1530–1534.

749 Morales-Briones D.F., Kadereit G., Tefarikis D.T., Moore M.J., Smith S.A., Brockington
750 S.F., Timoneda A., Yim W.C., Cushman J.C., Yang Y. 2021. Disentangling Sources
751 of Gene Tree Discordance in Phylogenomic Data Sets: Testing Ancient
752 Hybridizations in Amaranthaceae s.l. Systematic Biology. 70:219–235.
753 Nadeau N.J., Pardo-Diaz C., Whibley A., Supple M.A., Saenko S.V., Wallbank R.W.R., Wu
754 G.C., Maroja L., Ferguson L., Hanly J.J., Hines H., Salazar C., Merrill R.M., Dowling
755 A.J., ffrench-Constant R.H., Llaurens V., Joron M., McMillan W.O., Jiggins C.D.
756 2016. The gene cortex controls mimicry and crypsis in butterflies and moths. Nature.
757 534:106–110.
758 Naser-Khdour S., Minh B.Q., Zhang W., Stone E.A., Lanfear R. 2019. The Prevalence and
759 Impact of Model Violations in Phylogenetic Analysis. Genome Biol Evol. 11:3341–
760 3352.
761 Pease J.B., Haak D.C., Hahn M.W., Moyle L.C. 2016. Phylogenomics Reveals Three Sources
762 of Adaptive Variation during a Rapid Radiation. PLOS Biology. 14:e1002379.
763 Randi E., Lucchini V., Hennache A., Kimball R.T., Braun E.L., Ligon J.D. 2001. Evolution
764 of the Mitochondrial DNA Control Region and Cytochrome *b* Genes and the
765 Inference of Phylogenetic Relationships in the Avian Genus *Lophura* (Galliformes).
766 Molecular Phylogenetics and Evolution. 19:187–201.
767 Rasmussen D.A., Guo F. 2023. Espalier: Efficient Tree Reconciliation and Ancestral
768 Recombination Graphs Reconstruction Using Maximum Agreement Forests.
769 Systematic Biology. 72:1154–1170.
770 Reddy S., Kimball R.T., Pandey A., Hosner P.A., Braun M.J., Hackett S.J., Han K.-L.,
771 Harshman J., Huddleston C.J., Kingston S., Marks B.D., Miglia K.J., Moore W.S.,
772 Sheldon F.H., Witt C.C., Yuri T., Braun E.L. 2017. Why Do Phylogenomic Data Sets

Yield Conflicting Trees? Data Type Influences the Avian Tree of Life more than Taxon Sampling. Systematic Biology. 66:857–879.

Ren H., Wong T.K.F., Minh B.Q., Lanfear R. 2025. MixtureFinder: Estimating DNA Mixture Models for Phylogenetic Analyses. Mol Biol Evol. 42:msae264.

Richards E.J., Brown J.M., Barley A.J., Chong R.A., Thomson R.C. 2018. Variation Across Mitochondrial Gene Trees Provides Evidence for Systematic Error: How Much Gene Tree Variation Is Biological? Systematic Biology. 67:847–860.

Scally A., Dutheil J.Y., Hillier L.W., Jordan G.E., Goodhead I., Herrero J., Hobolth A., Lappalainen T., Mailund T., Marques-Bonet T., McCarthy S., Montgomery S.H., Schwalie P.C., Tang Y.A., Ward M.C., Xue Y., Yngvadottir B., Alkan C., Andersen L.N., Ayub Q., Ball E.V., Beal K., Bradley B.J., Chen Y., Clee C.M., Fitzgerald S., Graves T.A., Gu Y., Heath P., Heger A., Karakoc E., Kolb-Kokocinski A., Laird G.K., Lunter G., Meader S., Mort M., Mullikin J.C., Munch K., O'Connor T.D., Phillips A.D., Prado-Martinez J., Rogers A.S., Sajjadian S., Schmidt D., Shaw K., Simpson J.T., Stenson P.D., Turner D.J., Vigilant L., Vilella A.J., Whitener W., Zhu B., Cooper D.N., de Jong P., Dermitzakis E.T., Eichler E.E., Flicek P., Goldman N., Mundy N.I., Ning Z., Odom D.T., Ponting C.P., Quail M.A., Ryder O.A., Searle S.M., Warren W.C., Wilson R.K., Schierup M.H., Rogers J., Tyler-Smith C., Durbin R. 2012. Insights into hominid evolution from the gorilla genome sequence. Nature. 483:169–175.

Schneider V.A., Graves-Lindsay T., Howe K., Bouk N., Chen H.-C., Kitts P.A., Murphy T.D., Pruitt K.D., Thibaud-Nissen F., Albracht D., Fulton R.S., Kremitzki M., Magrini V., Markovic C., McGrath S., Steinberg K.M., Auger K., Chow W., Collins J., Harden G., Hubbard T., Pelan S., Simpson J.T., Threadgold G., Torrance J., Wood J.M., Clarke L., Koren S., Boitano M., Peluso P., Li H., Chin C.-S., Phillippy A.M.,

Durbin R., Wilson R.K., Flicek P., Eichler E.E., Church D.M. 2017. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome Res. 27:849–864.

Shen W., Le S., Li Y., Hu F. 2016. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLoS ONE. 11:e0163962.

Siepel A., Haussler D. 2005. Phylogenetic Hidden Markov Models. In: Nielsen R., editor. Statistical Methods in Molecular Evolution. New York, NY: Springer. p. 325–351.

Springer M.S., Gatesy J. 2016. The gene tree delusion. Molecular Phylogenetics and Evolution. 94:1–33.

Steenwyk J.L., Li Y., Zhou X., Shen X.-X., Rokas A. 2023. Incongruence in the phylogenomics era. Nat Rev Genet. 24:834–850.

Sun X., Cheng J. 2023. Conflicts in Mitochondrial Phylogenomics of Branchiopoda, with the First Complete Mitogenome of Laevicaudata (Crustacea: Branchiopoda). Current Issues in Molecular Biology. 45:820–837.

Thawornwattana Y., Dalquen D., Yang Z. 2018. Coalescent Analysis of Phylogenomic Data Confidently Resolves the Species Relationships in the Anopheles gambiae Species Complex. Molecular Biology and Evolution. 35:2512–2527.

Thawornwattana Y., Seixas F.A., Yang Z., Mallet J. 2022. Full-Likelihood Genomic Analysis Clarifies a Complex History of Species Divergence and Introgression: The Example of the erato-sara Group of Heliconius Butterflies. Systematic Biology. 71:1159–1177.

Toups B.S., Thomson R.C., Brown J.M. 2025. Complex Models of Sequence Evolution Improve Fit, But Not Gene Tree Discordance, for Tetrapod Mitogenomes. Syst Biol. 74:86–100.

Vanderpool D., Minh B.Q., Lanfear R., Hughes D., Murali S., Harris R.A., Raveendran M., Muzny D.M., Hibbins M.S., Williamson R.J., Gibbs R.A., Worley K.C., Rogers J.,

Hahn M.W. 2020. Primate phylogenomics uncovers multiple rapid radiations and ancient interspecific introgression. PLOS Biology. 18:e3000954.

Waterson R.H., Lander E.S., Wilson R.K., The Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature. 437:69–87.

Wong T.K.F., Cherryh C., Rodrigo A.G., Hahn M.W., Minh B.Q., Lanfear R. 2024. MAST: Phylogenetic Inference with Mixtures Across Sites and Trees. Systematic Biology.:syae008.

Yang W., Feiner N., Pinho C., While G.M., Kaliontzopoulou A., Harris D.J., Salvi D., Uller T. 2021. Extensive introgression and mosaic genomes of Mediterranean endemic lizards. Nat Commun. 12:2762.