Syst. Biol. 0(0):1–19, 2026
© The Author(s) 2026. Published by Oxford University Press on behalf of the Society of Systematic Biologists. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
https://doi.org/10.1093/sysbio/syag027
Advance Access publication March 10, 2026

Integrating Secondary Structure Information Enhances Phylogenetic Signal in Mitochondrial Protein Coding Genes

Claudio Cucini [1,2], Francesco Nardi [1,2], and Joan Pons [3,*]

[1]*Department of Life Sciences, University of Siena, Via Aldo Moro 2, 53100 Siena, Italy*
[2]*National Biodiversity Future Center (NBFC), Piazza Marina 61, 90133 Palermo, Italy*
[3]*Departament de Biodiversitat Animal i Microbiana, Institut Mediterrani d'Estudis Avançats, IMEDEA (UIB-CSIC), Miquel Marqués 21, 07190 Esporles, Spain*
**Correspondence to be sent to: Departament de Biodiversitat Animal i Microbiana, Institut Mediterrani d'Estudis Avançats, IMEDEA (UIB-CSIC), Miquel Marqués 21, 07190 Esporles, Spain; E-mail: jpons@imedea.uib-csic.es.*

Received 16 September 2024; reviews returned 7 October 2025; accepted 15 February 2026
Associate Editor: Lars Jermiin

Abstract.— Accurate phylogenetic inference requires models that account for heterogeneity in molecular evolution. Mitochondrial protein-coding genes, which encode membrane-bound proteins composed of multiple transmembrane α-helices, exhibit considerable compositional and functional variation across structural regions, variation that is often overlooked in standard partitioning strategies. Here, we introduce TRAMPO (TRAnsMembrane Protein Order), a novel pipeline that incorporates predicted secondary structural features (i.e., matrix-facing, transmembrane, and intermembrane-facing domains) into phylogenetic partitioning schemes. We applied TRAMPO to seven mitochondrial datasets, spanning crustaceans, hexapods, and vertebrates, and evaluated eight partitioning strategies based on combinations of codon position, strand, and secondary structure. Transmembrane helices showed pronounced thymine enrichment at second codon positions and hydrophobic amino-acid composition, reflecting domain-specific evolutionary constraints. To assess whether these structural patterns influence phylogenetic reconstruction, we performed maximum likelihood analyses under Markov models with various degrees of complexity (ranging from standard Markow models, via Lie Markov and General Heterogeneous evolution On a Single Topology Markovs models, to profile mixture Markov models). We also evaluated different models of rate-heterogeneity across sites (including the invariable sites model, gamma-distribution model, and FreeRate model) to examine their interaction with partitioning strategies and overall model performance. Incorporating structural information into partitioning schemes consistently improved model fit and reduced apparent heterogeneity, as reflected in lower AIC values and more compositionally homogeneous partitions. These improvements translated into more consistent and topologically congruent phylogenetic trees across most datasets, while also reducing computational time. Notably, second codon positions in DNA that encodes transmembrane helices were consistently retained as distinct partitions during model optimization, even in Mammals and Vertebrates, where secondary structure contributed little to overall model performance, underscoring their strong and conserved evolutionary signal. Surveys of tree space using quartet distances further supported these findings, with structurally informed models yielding more tightly clustered and internally consistent tree topologies. The benefits of structural partitioning were most pronounced in lineages of intermediate evolutionary depth and declined in ancient vertebrate and mammalian clades, where substitutional saturation accumulates with evolutionary time and strand asymmetry tends to emerge more frequently. In some cases, models with the lowest AIC did not yield the most congruent topologies, underscoring the limitations of information criteria when comparing models of different complexity. Overall, our findings demonstrate that secondary structural features, particularly the repetitive architecture of transmembrane helices, harbour meaningful phylogenetic signal. Incorporating this information into partitioning schemes improves tree reconstruction and mitigates underlying heterogeneity. TRAMPO provides a scalable, open-source tool to implement this approach in mitochondrial phylogenetics. [transmembrane regions; secondary structure; partitioning; mitochondrial genome; phylogenetics.]

In standard phylogenetic analyses, each site in a multiple sequence alignment (MSA) is typically assumed to evolve under independent Markovian conditions, yet under a uniform model of character frequencies, substitution process, and variation in rates among sites (Felsenstein 1978). However, using an inappropriate substitution model can compromise phylogenetic inference, highlighting the need for models that adequately reflect evolutionary processes (Posada and Crandall 1998; Minin et al. 2003; Kalyaanamoorthy et al. 2017; Abadi et al. 2019). Early in the development of molecular phylogenetics, researchers recognized that sequences from different genomes—such as nuclear vs. mitochondrial DNA—often exhibit distinct evolutionary dynamics, making the assumption of a uniform model across all sites overly simplistic (Shapiro et al. 2006). For example, pancrustacean mitochondrial genes display a strong A + T mutation bias, particularly at third codon positions (Pons et al. 2014). In contrast, certain vertebrates—such as anseriform birds—show G + C-rich mitochondrial genomes (Jermiin et al. 1994). Consequently, third codon positions often require distinct models of sequence evolution, including both substitution models and models of rate-heterogeneity among sites (e.g., Pons

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

et al. 2019), to account for compositional differences (e.g., Foster 2004) and to ensure appropriate model selection (e.g., Shapiro et al. 2006; Jayaswal et al. 2014).

Following the notion that ignoring potential heterogeneity in sequence evolution can lead to analytical errors (Ho and Jermiin 2004; Jermiin et al. 2004), two main strategies have been developed to address this issue: (1) a priori partitioning (e.g., Lanfear et al. 2012; Kainer and Lanfear 2015), where MSAs are divided based on molecular, structural, or functional features (e.g., codon positions or RNA secondary structure), and (2) model-based approaches, which dynamically assign sites to different categories according to shared evolutionary characteristics, such as nucleotide composition and rate variation. Notable examples of the latter include the General Heterogeneous evolution On a Single Topology (GHOST; Crotty et al. 2020), profile mixture models (CAT; Le et al. 2008), and other mixture models (Ren et al. 2025). The latter also accounts for heterotachy—evolutionary rate variation over time.

Initially implemented in phylogenetic programs such as p4 (Foster 2004), many of these advanced models have since been incorporated into user-friendly software like IQ-TREE 2 (Minh et al. 2020). This platform now supports a range of sophisticated models, including FreeRates (Yang 1995; Soubrier et al. 2012), Lie Markov models (Woodhams et al. 2015), GHOST (Crotty et al. 2020), codon (Muse and Gaut 1994), and mixture models (Ren et al. 2025), thereby enhancing the accuracy and flexibility of estimates accommodating evolutionary heterogeneity. The FreeRate model avoids the constraints of the model that assumes a gamma distribution by sampling site-specific rates from flexible discrete distributions (Soubrier et al. 2012). Lie Markov models allow non-reversible substitutions assuming stationary conditions and preserve purine/pyrimidine symmetry. The GHOST models combine site- and lineage-specific variation in stationary substitutions under a unified topology by probabilistically assigning sites to different discrete rate categories. Codon models explicitly incorporate the genetic code, estimating synonymous and nonsynonymous substitution rates, though they are computationally intensive. Mixture models, such as empirical profile mixture models (Le et al. 2008), dynamically group sites into evolutionary classes using separate models and parameters. These models typically obviate the need for user-defined partitions.

Choosing between these analytical approaches is particularly important when analysing sequences that include distinct functional or structural domains, which often exhibit substantial heterogeneity. For instance, in mitochondrial genomes, protein-coding genes (PCGs) display variation across codon positions (Leavitt et al. 2013; Jermiin et al. 2017), and ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) genes show variation across stem-loop structures (Gowri-Shankar and Rattray 2006; Erpenbeck et al. 2007). The secondary structure of proteins may also influence evolutionary constraints, with α-helices, β-sheets, and loops evolving differently (Le and Gascuel 2010; Barbera and Cucini 2022). Recently, incorporating tertiary structure has been shown to improve model fit and reduce systematic error (Puente-Lelievre et al. 2023), supporting the relevance of structure-informed partitioning (Pandey and Braun 2020).

When sites are automatically assigned to groups (point 2 above), the model and groupings are inferred directly from the data. By contrast, a priori partitioning (point 1 above) requires predefined criteria based on prior biological knowledge rather than patterns emerging from the data itself. In phylogenetic analyses involving large multigene datasets, sequences are often partitioned according to functional, structural, or biochemical characteristics (e.g., Stokkan et al 2018; Pons et al. 2019). Tools like PartitionFinder 2 (Lanfear et al. 2017) are then used to optimize partitioning schemes by balancing model fit and complexity, typically by using information criteria such as Akaike Information Criterion (AIC; Akaike 1973) or Bayesian Information Criterion (BIC; Schwarz 1978). As an example, partitioning by codon position and strand orientation has proven especially effective in mitochondrial phylogenetics (Pons et al. 2010; Bauzà-Ribot et al. 2012; Bover et al. 2019), capturing well-established heterogeneity in nucleotide composition, substitution process, and among-site rate variation (Pons et al. 2014), improving phylogenetic signal (Manuel et al. 2025). However, these approaches focus on evaluating whether predefined partitions should be merged, rather than identifying novel ones (Crotty and Holland 2022), highlighting the importance of incorporating biologically meaningful features—such as structural domains—during initial partition and model selection.

Mitochondrial PCGs, central to oxidative phosphorylation (OXPHOS), encode membrane proteins with transmembrane α-helices (TM) and extramembrane loops projecting into the mitochondrial matrix (MA) or intermembrane space (IM) (Lloyd and McGeehan 2013; Dennerlein and Rehling 2015). Based on primary sequences, transmembrane helices, and extramembrane loops, boundaries can be predicted using tools such as TMHMM (Möller et al. 2001), DeepTMHMM (Hallgren et al. 2022), HMMTOP (Tusnády and Simon 2001), and CCTOP (Dobson et al. 2015), or retrieved from curated resources.

Mitochondrial genes are widely used in phylogenetics because they occur in many copies per cell, are uniparentally inherited without recombination, and evolve faster than nuclear genes. Their conserved regions facilitate primer design, and complete mitogenomes can often be assembled from low-coverage data, making them especially valuable when nuclear sequences are difficult to obtain (e.g., fossils, museum specimens, environmental DNA, or large-scale taxon screenings; Blair 2023). Here, we present TRAMPO (TRAnsMembrane Protein Order), a pipeline for generating struc-

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

TABLE 1. Number of taxa and total alignment length of the concatenated sequences from the 13 mitochondrial protein-coding genes across the seven datasets, presented at both the nucleotide (nt) and amino-acid (aa) levels.

Dataset	# taxa	nt (total)	aa (total)	aa (MA)	aa (TM)	aa (IM)
Collembola	39	11103	3701	934	1945	822
Hyalella	33	10944	3648	912	1923	813
Metacrangonyx	23	10902	3634	913	1918	803
Pseudoniphargus	32	10956	3652	901	1936	815
Mammals	134	11367	3789	990	1937	862
Primates	69	11367	3789	990	1937	862
Vertebrates	150	11367	3789	990	1937	862

Note: The table also reports the number of amino-acid sites assigned to each secondary structural domain type, as defined by TRAMPO: matrix-facing (MA) loops, transmembrane (TM) helices, and intermembrane space-facing (IM) loops.

turally informed partitions of mitochondrial PCGs using predicted secondary structure. TRAMPO classifies aligned sequences into three functional classes—MA, TM, and IM—and assigns each site to its corresponding codon position and strand orientation. We use TRAMPO to analyze seven datasets encompassing 13 mitochondrial PCGs from three crustacean (amphipods) lineages (*Metacrangonyx*, *Pseudoniphargus*, and *Hyalella*), one hexapod taxon (Collembola, springtails), and three vertebrate clades (Vertebrates, Mammals, and Primates), spanning a broad taxonomic and temporal scale. For each of these data sets, we answer two questions:

1) Do the MA, TM, and IM domains differ in nucleotide composition and patterns of among-site rate variation, thereby reflecting heterogeneity associated with secondary structure?
2) Using TRAMPO, does incorporating structural information improve phylogenetic inference relative to models such as GHOST and profile mixture models, which are specifically designed to account for heterogeneity?

To address these questions, we analysed compositional skews (A- and G-skew; e.g., Sueoka 1995), amino-acid composition, and codon usage (Sharp and Li 1987) across distinct functional regions. Based on these patterns, we defined alternative partitioning schemes informed by functional regions and conducted comparative phylogenetic analyses. Our results support the idea that partitioning by functional region, alongside codon position and strand orientation, can significantly improve model fit and, hence, the phylogenetic reconstruction.

Materials and Methods

Datasets

The mitochondrial PCGs in our datasets—*Metacrangonyx* (Pons et al. 2019), *Pseudoniphargus* (Stokkan et al. 2018), *Hyalella* (Zapelloni et al. 2021), and Collembola (Cucini et al. 2021a)—were retrieved from previous studies (Table 1). Three additional datasets (Vertebrates, Mammals, and Primates) were assembled from complete mitochondrial genomes available in GenBank as of November 2021. For the Vertebrates dataset (150 species), one species per order was selected; for Mammals (134 species), one per family; and for Primates (69 species), one per genus. Only one species per order, family, or genus (depending on the dataset) was randomly selected to minimize user-driven bias. Although the Mammal and Vertebrate datasets share 27 species and the Mammals and Primates datasets share 14, these overlaps represent only a small fraction of the total number of species in each dataset. Species names and GenBank accession numbers are provided in Supplementary Table S1. Both this table and the multiple sequence alignments (MSAs, in FASTA and Nexus formats; Supplementary Data S1) are also available at https://github.com/dbajpp0/TRAMPO.

In *Metacrangonyx*, *Pseudoniphargus*, *Hyalella*, and Collembola, most PCGs are encoded on the positive strand, except for *nad1*, *nad4*, *nad4L*, and *nad5*, consistent with the ancestral pancrustacean gene order. In *Metacrangonyx*, the *cob* gene is also encoded on the negative strand. In all vertebrate lineages, PCGs are located on the positive strand, except for *nad6*. Within each dataset, the coding strand from which PCGs are transcribed is consistent across genomes. Altogether, these seven datasets span a wide taxonomic breadth and evolutionary timescale—from ∼25 Ma (*Hyalella*) to ∼440 Ma (Vertebrata) (Brazeau and Friedman 2015; Zapelloni et al. 2021).

TRAMPO Pipeline

TRAMPO (TRAnsMembrane Protein Order) is a bioinformatic pipeline implemented in Bash and Python 3, incorporating previously published components (Cucini et al. 2021b; Narakusumo et al. 2020). Its primary goal is to partition aligned mitochondrial PCGs into TM, MA and IM regions. These outputs are intended as data partitions to be incorporated in downstream phylogenetic analyses. The pipeline outputs partition files in Nexus format, including codon and encoding strand information, and generates comprehensive tables and plots summarizing nucleotide and amino-acid composition, A- and G-skews, Relative Synonymous Codon Usage (RSCU; Sharp and Li 1987), and domain-specific features. Furthermore, TRAMPO classifies amino-acids into six physicochemical groups based on Dayhoff et

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

al. (1978): G1: Small (A, G, P, S, T), G2: Acid/Amide (D, E, N, Q), G3: Basic (H, K, R), G4: Hydrophobic (I, L, M, V), G5: Aromatic (F, W, Y), and G6: Sulfur (C). The pipeline operates within a dedicated Conda environment, which is automatically created during installation. TRAMPO is freely available at https://github.com/dbajpp0/TRAMPO. An overview of the workflow is summarized in Fig. 1.

Briefly, TRAMPO removes terminal stop codons, translates nucleotide PCG sequences into amino acids using the appropriate NCBI mitochondrial code, and aligns them with MAFFT v7.475 using default parameters (Katoh and Standley 2013). For each gene, the alignment is anchored by the orthologous sequence of a representative model organism such as *Homo sapiens* (vertebrates), *Drosophila melanogaster* (pancrustaceans), *Patiria pectinifera* (echinoderms), *Albinaria caerulea* (mollusks), *Lumbricus terrestris* (annelids), *Caenorhabditis elegans* (nematodes), and *Metridium senile* (cnidarians), sourced from UniProt-SwissProt (Boutet et al., 2007) via MAFFT v7.475 using the –add option. Accession numbers of the representative protein sequences are provided in Supplementary Table S2. Nucleotide sequences are then codon-realigned to preserve the reading frame. Structural domains—TM, MA, and IM—are not predicted anew, but assigned from a curated reference database integrated into the pipeline. This database was built from UniProt-SwissProt mitochondrial PCGs, with domain boundaries first inferred by TMHMM v2.0 (Möller et al. 2001) and HMMTOP v2.0 (Tusnády and Simon 2001), refined with CCTOP v1.1.0 (Dobson et al. 2015), and manually adjusted to correct systematic biases (e.g., merging adjacent MA–TM–IM segments in *nad4* and *nad2*, and enforcing matrix-facing N-termini). During analysis, TRAMPO applies these curated boundaries to assign each aligned site to its domain, and produces partitioning schemes by gene, codon position, domain, and—when gene order is specified—strand. For each dataset analysed (Table 1), sequences were translated using the appropriate NCBI mitochondrial code (-c option in TRAMPO; invertebrates = 5, vertebrates = 2), and partitioned according to the domain structure of *D. melanogaster* or *H. sapiens* for invertebrate and vertebrate datasets, respectively.

Partitioning Schemes

Eight partitioning schemes were generated using TRAMPO by first splitting each of the 13 mitochondrial PCGs according to codon position (first, second, and third), encoding strand (positive or negative), and transmembrane domain type (MA, TM, and IM). The resulting gene alignments were then concatenated into a single alignment of the 13 PCGs, in which the partitioning schemes were incorporated. Designed to reflect commonly used strategies in mitochondrial phylogenomics, these schemes were employed to assess heterogeneity and phylogenetic informativeness, with particular emphasis on domain-based partitioning (Table 2). A visual summary of these schemes is provided in Supplementary Fig. S1.

Phylogenetic Inference

All phylogenetic analyses were done under the maximum likelihood (ML) criterion, as implemented in IQ-TREE v2.4.0 (Minh et al. 2020), and initiated from the same parsimony tree with a fixed seed (-t PARS -seed 13579 -n 0). We rely on information criteria such as AIC because they provide an objective and reproducible statistical framework to evaluate the fit of alternative nested and non-nested models of the data. This approach can be applied consistently across datasets and taxonomic groups, balancing likelihood improvement against model complexity (over-parameterization) while accounting for heterogeneity. Phylogenies, in turn, were not compared against a reference tree, as the true tree is inherently unknown and represents the ultimate objective of phylogenetic inference, making such comparisons potentially biased. We acknowledge, however, that relying solely on AIC has limitations, as it evaluates model fit rather than the accuracy of the inferred topology. Model selection was based on AIC, which has been shown to perform better than BIC in small datasets, since most of our subsets (partitions) are under a thousand sites (Posada and Buckley 2004; Sullivan and Joyce 2005). To enable comparisons across lineages, AIC values were normalized as follows. For each lineage, the AIC of the ML tree under the most subdivided scheme in PartitionFinder 2 (the 18-partition scheme) was used as the reference, with all other values expressed relative to it. This normalization indicates the difference of each partition scheme from reference one within its lineage, providing a consistent scale for cross-lineage comparisons (AIC_norm(i) = (AIC_i - AIC_ref) / AIC_ref).

Analysis of concatenated dataset without partitioning. First, analyses were conducted on the concatenated dataset of the 13 PCGs without partitioning, using four distinct modelling approaches:

Approach 1 *A standard phylogenetic analysis was conducted using the most parameter-rich standard Markov substitution model: GTR + G + FO.*

Approach 2 *We conducted a selection among several standard Markov models. ModelFinder (Kalyaanamoorthy et al. 2017; option -m MPF) was used to select the best-fitting substitution model among a set of standard time-reversible Markov models (F81, HKY, TN, K3Puf, TIM, TVM, and GTR) with either empirical (+F) or optimized (+FO) base frequencies. Only models allowing for unequal nucleotide composition were assessed, as this is a characteristic feature of mitochondrial genomes in metazoans (Smith 2012). In addition, various models of among-site rate variation were evaluated, including a proportion of invariable sites (+I), a gamma distribution (+G), the FreeRates model (+R), and GHOST models*

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

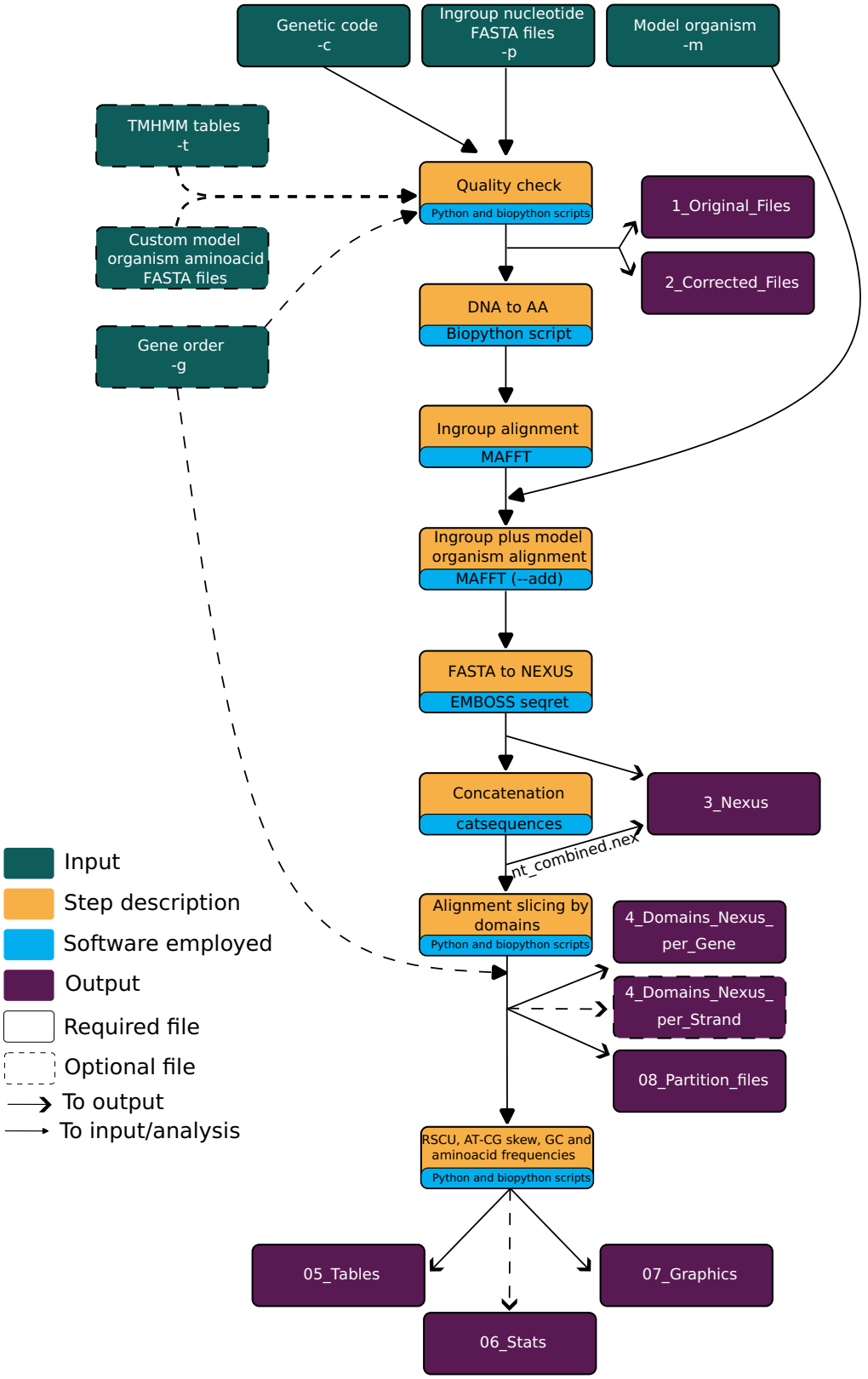


FIGURE 1. Overview of the TRAMPO pipeline workflow. Input, output and analytical steps are colour-coded. Optional input/output files are represented by dashed boxes while required input files as well as default outputs are indicated by solid lines.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

TABLE 2. Partitioning schemes of the 13 mitochondrial protein-coding genes generated by the TRAMPO pipeline and evaluated using maximum likelihood-based phylogenetic analyses

Partition scheme	Description	Subsets
1	codon position (first, second, third)	3
2	transmembrane domain (MA, TM, IM)	3
3	codon position × encoding strand (positive, negative)	6
4	transmembrane domain × encoding strand	6
5	codon position × domain	9
6	gene	13
7	codon position × encoding strand × domain	18
8	codon position × encoding strand × simplified domain *	12

*Simplified domain was obtained by combining MA + IM into MAIM; i.e., MAIM vs TM.

*using either linked (+H) or unlinked (*H) parameter configurations.*

Approach 3 *ModelFinder was restricted to the set of Lie Markov models (-m, from 1.1 to 12.12), combined with optimized base frequencies (+FO) and FreeRates (+R) to account for among-site rate variation.*

Approach 4 *We evaluated mixtures of standard Markov matrices (-m MIX + MFP) with unequal base compositions (+F or + FO), in combination with models of rate heterogeneity (+I, +G, or + R).*

Phylogenetic analyses of partition schemes defined in TRAMPO using ModelFinder. Partitioned analyses were performed for each of the eight partitioning schemes, evaluating both standard and Lie Markov models in separate analyses. The best-fitting substitution model for each partition was identified using ModelFinder (Kalyaanamoorthy et al. 2017). The -Q option enabled independent branch length estimation per partition. In the Lie Markov analysis, only FreeRates (+R) and + FO were used, excluding + I, +G, or + F based on the results obtained with the standard model selection. Combined I + G models were excluded because invariant sites are inherently represented within the lowest rate category of the gamma distribution (Sullivan et al. 1999).

Phylogenetic analyses of partition schemes defined in TRAMPO using PartitionFinder. We used PartitionFinder 2 (Lanfear et al. 2017) along ModelFinder, both integrated within IQ-TREE 2 (-Q option), to merge compatible subsets and optimize substitution models. We repeated independent analyses using both standard and Lie Markov models with specific settings in PartitionFinder 2. Models and rate categories were allowed to merge by enabling the -m*erge-model* and -m*erge-rate* options. The number of FreeRates categories was constrained between three and five (*-cmin 3 -cmax 5*), based on results previously obtained with ModelFinder. To promote a more exhaustive search of possible partition merging combinations, we applied a greedy *rcluster* algorithm with high thresholds (*-rcluster 100 -rcluster-max 10000*).

Analysis of GHOST Categories

Using GHOST's posterior site probabilities (*-wspm*), we extracted the category with the highest probability per site. We developed custom Python 3 scripts (Supplementary Data S1) to analyze the distribution of site probability categories (p1–p10) across different genomic regions defined by strand, structural domain, and codon position. For each site, the category with the highest assignment probability was first selected and recorded. The resulting site-level classifications were then summarized to obtain observed counts of categories within each genomic region. Expected counts were calculated under the null hypothesis of independence, based on the marginal totals of the contingency tables. Standardized residuals were computed to quantify deviations between observed and expected counts, and statistical significance was assessed using a two-sided Bonferroni correction to adjust for multiple comparisons across categories. Categories showing significant over- or under-representation were expressed as percentage deviations from expectation. Results were visualized as bar plots, and summary tables containing observed, expected, residual, and significance values were also generated. All analyses were performed in a single Python 3 script, and figures were produced with Matplotlib (Hunter 2007).

Topological Congruence and Tree Space Exploration

To assess whether different partitioning schemes resulted in significantly different tree topologies, we applied the Approximately Unbiased (AU) test (Shimodaira 2002) in IQ-TREE 2 using 10,000 replicates. To visualize topological differences, we used the TreeDist package (Smith 2022), which calculates quartet distances among all possible sets of four taxa. The software then projects trees into tree space by reducing dimensionality with Principal Coordinates Analysis (PCoA), and finally groups trees using the k-means clustering algorithm. PCoA was selected over alternatives such as Sammon (1969) or Kruskal (1964) mappings due to its superior ability to preserve pairwise topological distances (Smith 2022). We also evaluated previously published tree topologies inferred from the same datasets for *Hyalella* (Zapelloni et al. 2021), *Metacrangonyx* (Pons et al. 2019), and *Pseudoniphargus* (Stokkan et al. 2018), which were constructed using a relaxed uncorrelated molecular clock model as an alternative approach to account for evolutionary rate heterogeneity. A similar strategy was applied to Collembola, using a non-ultrametric topology from which taxa not included in the present study were pruned (Cucini et al. 2021a).

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

Results

Nucleotide and Amino-Acid Composition, Intrastrand Skew, and Codon Usage of Secondary Structures

TRAMPO successfully assigned aligned positions to the same functional domains—MA, TM and IM—as predicted in *Homo sapiens* (for Vertebrates, Mammals, and Primates) and *Drosophila melanogaster* (for springtails and amphipods). The final MSAs and domain assignments are provided in Supplementary Data S1. As expected, TMs contained more positions than the MA and IM loops across all datasets and genes, except in short genes such as *atp8* and *cox2*, which contained only one or two TM helices, respectively (Table 1).

Nucleotide composition and A- and G-skew values varied across codon positions and encoding strands (Fig. 2; Supplementary Fig. S2). As anticipated, third codon positions—regardless of domain or strand—had the highest A + T content (Supplementary Fig. S3), especially in the four pancrustacean datasets (Collembola and Amphipoda), which were notably A + T rich compared to vertebrates. The most striking and consistent result was that second codon positions in TM helices exhibited uniquely low A-skew values (∼0.2), reflecting a strong intra-strand bias toward thymine over adenine (Fig. 2; Supplementary Fig. S2). In pancrustaceans, this trend was more pronounced between positive and negative strands. Notably, in the Collembola dataset, second codon sites on the negative strand of MA domains exhibited even more extreme A-skew values (<0.2). First codon positions showed more balanced nucleotide composition, while third codon positions in TM helices often formed distinct compositional clusters (Fig. 2; Supplementary Fig. S2-S3). In summary, second codon positions in TM helices consistently showed elevated thymine content and distinct nucleotide composition across lineages.

Amino-acid composition also varied across domains (Fig. 3; Supplementary Fig. S4). TM helices were enriched in hydrophobic amino-acids (G4 group) and depleted in acidic, amide (G2), and basic (G3) residues relative to MA and IM loop domains. Differences in small (G1) and aromatic (G5) groups were minor. One exception was the Collembolan dataset, where MA loop domains on the negative strand showed a deviation in G4 group composition, corresponding with low A-skew at second codon positions. Across all datasets, strand-based differences were less pronounced than those between domains.

Codon usage analysis revealed a general preference for A + T-rich codons over GC-rich ones across all regions. Some lineage-specific trends were also observed, as in case of vertebrates, which differ from pancrustaceans in the relative usage of Leucine and Serine codons (data not shown) due to differences between the genetic codes.

Inference of Phylogenetic Trees Under Maximum Likelihood

Phylogenetic analyses of PCGs included four unpartitioned analyses, 16 partitioned analyses using ModelFinder (eight partition schemes, see Table 2, each evaluated with standard and Lie Markov models), and 16 additional partitioned analyses using PartitionFinder 2 with the same schemes and model types—resulting in a total of 36 ML analyses per lineage (see Table 3).

Phylogenetic analyses of unpartitioned dataset of the 13 mitochondrial PCGs.— We first inferred ML trees for each dataset without partitioning, using the most complex standard Markov model, GTR + G + FO (Approach 1 in Material and Methods). In Approach 2, which involved ModelFinder-based selection among standard Markov models, the best-fitting models were typically TVM + FO or GTR + FO combined with GHOST-based among-site rate variation. Most datasets comprised 7–10 categories with unlinked parameters, except for Mammals, in which parameters were linked (+H10). These results suggest that GHOST more effectively captured site heterogeneity compared to the Invariant sites model (+I), Gamma distribution (+G), or FreeRates (+R). Approach 3, based on ModelFinder selection among Lie Markov models, consistently identified the fully non-reversible model 12.12 with 4–10 FreeRates categories as the best fit. In this approach, FreeRates (+R) was the only model tested for among-site rate variation, as it consistently outperformed the invariable sites model (+I) and the gamma distribution (+G) in all previous analyses across lineages using standard Markov models. Finally, the mixture model approach (Approach 4) applied a combination of 4–10 substitution matrices—primarily GTR + FO and TVM + FO—alongside 4–10 FreeRates categories.

Phylogenetic analyses based on eight partitioning schemes of the 13 mitochondrial PCGs.— We also evaluated the impact of data partitioning based on codon position, encoding strand, and secondary structure using eight predefined partitioning schemes (Table 2 and Table 3). For each partition, we first identified the best-fitting evolutionary model using ModelFinder and AIC scores, followed by ML tree estimation. Next, we investigated whether merging initial partitions using the PartitionFinder 2 algorithm improved model fit, as indicated by lower AIC values. The best-fit models before and after merging were largely consistent. Most included standard models such as GTR + FO, TIM + FO, and TVM + FO, or Lie Markov models such as 12.12 and RY8.8, typically combined with 3–5 FreeRates categories.

Normalized AIC differences across 36 ML analyses per lineage (as shown in Fig. 4 and Table 3) indicate, as expected, the four non-partitioned analyses had the highest (worst) AIC scores, even when using Lie Markov

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

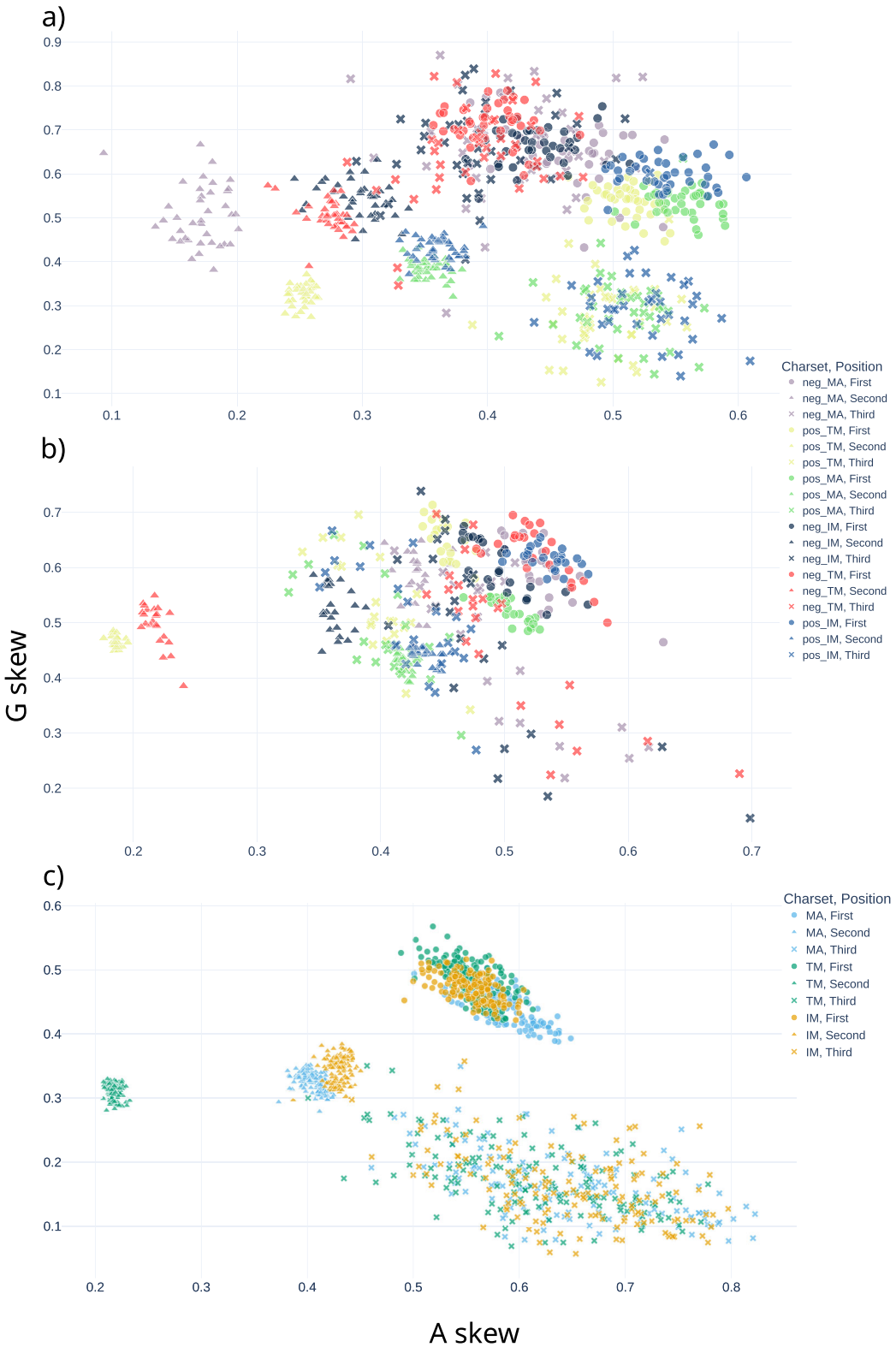


FIGURE 2. Scatter plots of A-skew (x-axis) versus G-skew (y-axis) for two strand orientation (positive "pos" and negative "neg"), three transmembrane domains (IM, MA, TM), and three codon positions across selected datasets: a) Collembola, b) *Metacrangonyx*, and c) Vertebrates. Partitions are labelled as 'strand_domain_codon position'. Results for additional lineages are provided in Supplementary Fig. S2. In panel (c), strand differences are not shown due to the presence of only a single gene encoded on the negative strand in vertebrates.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

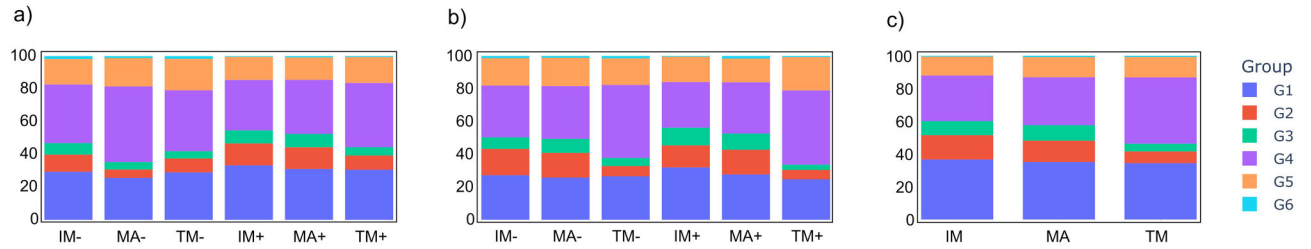


FIGURE 3. Frequency of amino-acids grouped into six chemical classes across two strand orientation (positive "+" and negative "–") and three transmembrane domains (IM, MA, TM) across selected datasets: a) Collembola, b) *Metacrangonyx*, and c) Vertebrates. Results for additional datasets are shown in Supplementary Fig. S4. In panel (c), strand differences are not displayed due to the presence of only a single gene encoded on the negative strand in vertebrates. The six chemical classes are: G1–Small (A, G, P, S, T); G2–Acidic/Amide (D, E, N, Q); G3–Basic (H, K, R); G4–Hydrophobic (I, L, M, V); G5–Aromatic (F, W, Y); and G6–Sulfur-containing (C).

models or complex mixture models (Fig. 4). GHOST models improved AIC scores across datasets but were best-performing only in vertebrates (Fig. 4). Analyses based solely on gene or domain partitions, lacking codon-level resolution, consistently performed poorly. Approaches incorporating codon position, and encoding strand, and transmembrane domain information consistently yielded the lowest (i.e., best) AIC scores in *Hyalella*, *Metacrangonyx*, *Pseudoniphargus*, and Primates. For Collembola, Mammals, and Vertebrates, the best-supported models were those incorporating codon position and encoding strand information while excluding transmembrane domain information. However, partitioning schemes that included transmembrane information yielded AIC scores that were only slightly worse, suggesting that this structural feature still carries some phylogenetic signal.

We examined which initial partitions (predefined site sets) were retained as independent or merged after PartitionFinder 2 analysis, as this reflects whether they contain distinct evolutionary features captured in the ML framework. Analyses that included TM information generally retained TM domains as separate partitions, particularly the second codon positions within TM regions, indicating a distinct evolutionary signal that did not lead to partition merging during model optimization (Fig. 5). In summary, second codon positions of TM helices were rarely merged, though MA and IM second positions were often grouped.

After partition merging with PartitionFinder 2, second codon positions within TM helices were consistently grouped into the same partition in *Hyalella* and *Metacrangonyx*, regardless of strand. By contrast, in *Pseudoniphargus*, these sites were split by encoding strand. Additionally, *Pseudoniphargus* showed strand-specific partitioning of second codon positions in the MA + IM domains. In turn, this separation was observed only on the positive strand in vertebrates and springtails. Overall, all these patterns highlight that TM helices—particularly at second codon positions—follow distinctive evolutionary trajectories, justifying the use of separate substitution models (Fig. 5).

Further trends emerged across datasets. In some cases—such as vertebrates—first codon positions on the negative strand tended to form unique partitions or were aggregated with second codon positions. Third codon positions on the negative strand often formed independent partitions, highlighting their distinct evolutionary behaviour. Additionally, first and third codon sites on the positive strand either were joined together or separated depending on their association with transmembrane domains (TM versus MA + IM), suggesting a structural influence on substitution patterns.

Analyses of GHOST categories

Observed counts of positions assigned to the different GHOST categories (p1–p10; see Materials and Methods) were compared with the expected counts under the null hypothesis of independence (random distribution), derived from the marginal totals of the contingency tables. This analysis revealed that several categories were significantly unevenly distributed across codon positions, strand, and transmembrane domains (Fig. 6). The most pronounced over- and under-representations occurred at the three codon positions, while deviations were comparatively smaller for strand and especially for transmembrane domains. Among codons, position 2 showed the most pronounced imbalance, especially in arthropods (Collembola, *Hyalella, Metacrangonyx* and *Pseudonipharsus* datasets), while third codon sites also exhibited significant deviations across all lineages. Notably, third codon deviations often counterbalanced those at second codon, with categories enriched at one codon tending to be depleted at the other. Despite being assessed independently for each lineage, a consistent pattern emerged: Collembola, *Hyalella*, *Metacrangonyx*, and *Pseudoniphargus* displayed similar numbers of over- and under-represented categories, whereas Primates, Mammals, and Vertebrates showed a larger proportion of under-represented categories (Fig. 6). The overall number of categories with significant uneven distributions was lower in lower rank taxonomic groups (*Hyalella*, *Metacrangonyx*, *Pseudoniphargus*) and higher in broader groups (Collembola, Primates, Mammals, Vertebrates). Strand-related effects were strongest in arthropods (particularly Collembola and *Hyalella*, but also

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

TABLE 3. Likelihood (Ln) and normalized AIC values (see Material and Methods) from maximum likelihood analyses conducted across seven lineages, for 36 analyses, four concatenated analyses and 32 partitioned (see Table 2).

number	partition_name	Ln_Coll	AIC_Coll	Ln_Hyal	AIC_Hyal	Ln_Meta	AIC_Meta	Ln_Pseu	AIC_Pseu	Ln_Prim	AIC_Prim	Ln_Mamm	AIC_Mamm	Ln_Vert	AIC_Vert
1	18p_str_dom_cod _merged_Std	-267383.7	0.000000	-101833.1	0.000000	-135612.99	0.000000	-163462.5	0.000000	-297738.82	0.000000	-696389.99	0.000000	-861010.51	0.000000
2	18p_str_dom_cod _merged_Lie	-267189.79	-0.000317	-101800.17	-0.000116	-135737.28	0.000869	-164351.81	0.005226	-299964.96	0.007862	-697562.41	0.001697	-860145.97	-0.000924
3	12p_MAyIM_merged_Std	-267391.63	0.000030	-101833.15	0.000001	-135669.85	-0.000023	-163462.64	0.000001	-297735.78	0.000003	-696302.79	-0.000125	-860978.19	-0.000040
4	12p_MAyIM_merged_Lie	-267198.58	-0.000284	-101800.15	-0.000117	-135629.58	-0.000172	-163690.5	0.000871	-297656.39	-0.000206	-697262.33	0.001266	-861106.98	0.000127
5	9p_dom_cod_merged_Std	-270549.14	0.010460	-102176.55	0.001137	-136302.2	0.003403	-165728.17	0.011902	-300245.07	0.007906	-705967.3	0.012911	-867424.41	0.006707
6	9p_dom_cod_merged_Lie	-270256.46	0.009752	-102072.2	0.001016	-136236.96	0.003394	-165727.02	0.011548	-300663.3	0.008834	-706285.27	0.013382	-866857.19	0.005702
7	6p_cod_str_merged_Std	-267482.83	0.000030	-102164.16	0.000996	-135920.77	0.001005	-163881.35	0.000687	-298195.03	0.000530	-697821.46	0.000842	-861573.86	-0.000073
8	6p_cod_str_merged_Lie	-267401.93	-0.000227	-102136.92	0.000887	-135911.17	0.001000	-163875.57	0.000731	-298159.84	0.000449	-698597.63	0.001968	-861795.86	0.000196
9	3p_cod_merged_Std	-270549.71	0.010462	-102417.87	0.001999	-136542.97	0.004319	-166112.25	0.012877	-300904.32	0.008613	-708479.68	0.015302	-872187.32	0.011500
10	3p_cod_merged_Lie	-270375.98	0.009855	-102404.97	0.001961	-136549.69	0.004405	-166053.33	0.012585	-301107.9	0.009325	-707979.8	0.014594	-867341.46	0.005897
11	18p_str_dom_cod_Std	-266654.33	0.000972	-101403.75	0.002720	-135346.48	0.001583	-163076.74	0.001134	-296798.01	0.002166	-694410.01	0.001175	-858472.63	0.001052
12	18p_str_dom_cod_Lie	-266552.79	0.000713	-101332.15	0.002812	-135288.11	0.001544	-163139.79	0.001774	-296744.17	0.002177	-695932.36	0.003374	-858174.48	0.000780
13	12p_MAyIM_Std	-267059.53	0.000499	-101624.45	0.000687	-135492.2	0.000288	-163372.84	0.000282	-297248.97	0.000774	-695114.67	-0.000189	-859304.23	-0.000137
14	12p_MAyIM_Lie	-266918.52	0.000059	-101580.8	0.000749	-135456.1	0.000251	-163352.28	0.000334	-297188.96	0.000713	-695321.82	0.000151	-864413.18	0.005721
15	9p_dom_cod_Std	-270085.89	0.010784	-101987.82	0.002067	-136205.26	0.004292	-165581.22	0.012317	-299928.3	0.008366	-702650.05	0.009426	-865605.73	0.006110
16	9p_dom_cod_Lie	-269952.48	0.010390	-101946.86	0.002038	-136102.57	0.003751	-165492.82	0.011973	-300189.74	0.009304	-704903.68	0.012687	-866257.68	0.006886
17	6p_cod_str_Std	-267444.99	-0.000081	-102095.03	0.001063	-135916.02	0.000955	-163881.47	0.000676	-298059.82	0.000582	-696732.27	-0.000283	-860381.13	-0.001059
18	6p_cod_str_Lie	-267310.46	-0.000501	-102075.56	0.001049	-135884.63	0.000842	-163871.57	0.000695	-297993.96	0.000436	-696636.17	-0.000398	-860317.3	-0.001117
19	3p_cod_Std	-270442.75	0.010078	-102417.85	0.001979	-136523.15	0.004188	-166112.31	0.012878	-300874.47	0.008520	-704613.47	0.009800	-867493.75	0.006088
20	3p_cod_Lie	-270402.77	0.009962	-102406.45	0.001975	-136498.36	0.004072	-166053.31	0.012585	-301165.82	0.009512	-705686.33	0.011343	-866653.73	0.005128
21	approach1_single _GTRGFO	-278024.74	0.037655	-106145.62	0.036875	-141712.15	0.041403	-171977.05	0.047633	-309936.39	0.037809	-726260.79	0.039939	-894114.14	0.036168
22	approach2_single _MF_Std_GHOST	-268528.07	0.005001	-104928.98	0.028962	-139975.57	0.031387	-166422.58	0.017580	-297947.55	0.002017	-710623.79	0.020981	-854058.51	-0.007046
23	approach3_single _MF_Lie	-276826.69	0.033240	-105814	0.033726	-141257.98	0.038163	-170525.82	0.038863	-307587.26	0.030003	-716577.93	0.026110	-884055.7	0.024540
24	approach4_single _MF_mixed	-276947.43	0.033960	-107321.18	0.048624	-142540.53	0.047915	-171070.1	0.042433	-306604.23	0.026840	-708875.7	0.015194	-854731.35	0.005040
25	13p_gen_Std	-272777.4	0.022229	-104956.41	0.034179	-139869.51	0.032855	-167659.44	0.026920	-303310.88	0.021795	-706124.03	0.016099	-872293.71	0.015358
26	13p_gen_Lie	-272416.37	0.021042	-104913.94	0.034165	-139794.05	0.032793	-167538.59	0.026428	-303199.67	0.021533	-705736.9	0.015603	-871624.09	0.014644
27	13p_gen_merged_Std	-273304.61	0.022095	-105375.5	0.031652	-140129.8	0.032351	-168027.47	0.025940	-304107.82	0.021357	-709947.52	0.019412	-876434.52	0.018596
28	13p_gen_merged_Lie	-272993.42	0.021016	-105355.62	0.031585	-140067.17	0.032096	-167930.85	0.025455	-303981.57	0.020984	-709082.09	0.018611	-874955.67	0.017646
29	3p_dom_Std	-276792.92	0.033801	-105658.4	0.033681	-141122.42	0.037967	-170371.96	0.038828	-307390.61	0.030356	-716816.47	0.027275	-884088.43	0.025322
30	3p_dom_Lie	-276639.87	0.033278	-105639.84	0.033578	-141094.04	0.037862	-170317.71	0.038564	-307273.56	0.030001	-715936.91	0.026032	-883418.8	0.024557
31	3p_dom_merged_Std	-276926.59	0.033919	-105696.6	0.033312	-141162.08	0.037847	-170434.33	0.038745	-307606.17	0.030538	-719387.77	0.030514	-887472.79	0.028845
32	3p_dom_merged_Lie	-276787.25	0.033425	-105680.07	0.033190	-141145.45	0.037769	-170381.01	0.038463	-307493.63	0.030182	-718654.47	0.029473	-886950.16	0.028247
33	6p_str_dom_Std	-273985.89	0.024350	-105368.36	0.033037	-140553.89	0.035062	-168378.01	0.028087	-305024.95	0.023917	-711212.43	0.020455	-878006.14	0.019362
34	6p_str_dom_Lie	-273669.51	0.023263	-105336.01	0.032916	-140506.83	0.034929	-168300.21	0.027741	-304876.25	0.023490	-710636.03	0.019661	-877426.26	0.018710
35	6p_str_dom_merged_Std	-274392.34	0.024467	-105499.13	0.032107	-140678.65	0.034700	-168482.16	0.027795	-305528	0.023583	-714083.13	0.023320	-881888.79	0.022376
36	6p_str_dom_merged_Lie	-274025.28	0.023471	-105471.58	0.031936	-140643.48	0.034537	-168492.86	0.027452	-305390.3	0.023142	-713610.37	0.022656	-881445.29	0.021870

Coll = Collembola, Hyal = *Hyalella*, Meta = *Metacrangonyx*, Pseu = *Pseudoniphargus*, Prim = Primates, Mamm = Mammals, Vert = Vertebrates

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

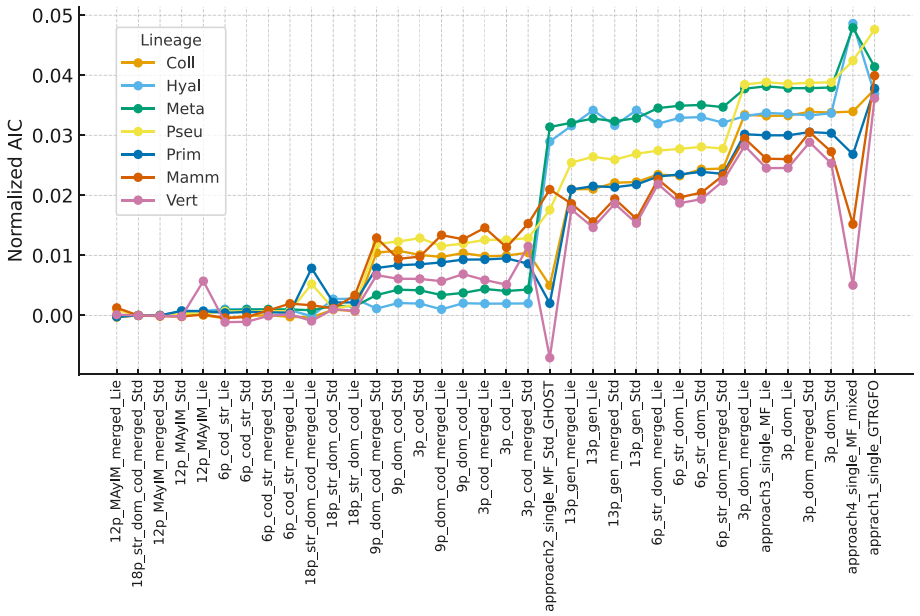


FIGURE 4. Line plot showing normalized Akaike Information Criterion (AIC) values from maximum likelihood analyses conducted across seven lineages, including four concatenated analyses and 32 partitioned analyses—16 before and 16 after merging eight partition schemes (see Table 2 and Materials and Methods). For each lineage, the AIC value of the maximum-likelihood tree inferred under the 18-partition scheme selected by PartitionFinder 2 was taken as the reference ("ref") value. Normalization was then calculated as: AIC_norm(i) = (AIC_i - AIC_ref) / AIC_ref. Coll = Collembola, Hyal = *Hyalella*, Meta = *Metacrangonyx*, Pseu = *Pseudoniphargus*, Prim = Primates, Mamm = Mammals, Vert = Vertebrates. For each of the 36 partition schemes, the median normalized AIC value was calculated across the seven lineages, and schemes are ordered accordingly.

partly in *Metacrangonyx* and *Pseudoniphargus*), suggesting lineage-specific biases in category representation (Fig. 6). These differences likely reflect variation in the number of genes encoded on the negative strand (e.g., in Primates, Mammals, and Vertebrates only *nad6* is encoded on the negative strand). Finally, arthropod lineages also showed enrichment of certain categories in TM helices, with corresponding depletions in MA or IM regions, whereas such domain effects were minimal to absent in Primates, Mammals, and Vertebrates.

Topological Tree Congruence and Tree Space Analyses

We assessed tree congruence with the AU test under standard Markov models, comparing a partition scheme that incorporates transmembrane regions (18p) with one based solely on codon position and strand (9p, without transmembrane regions). Those trees, inferred using models that included or excluded secondary structure, were congruent across all lineages, except in Vertebrates. To explore the differences among tree topologies obtained in the 36 ML analyses, we performed a Principal Coordinates Analysis (PCoA) of quartet distances (Fig. 7). Across lineages, trees were concentrated in a narrow region of tree space, with silhouette scores above 0.8, indicating that the major features of tree space were precisely preserved. In most datasets, trees were tightly grouped around the focal topology—defined as that obtained from the optimal merging of 18 partitions—although broader divergence was observed in Collembola (Fig. 7a) and *Metacrangonyx* (Fig. 7c). Clustering analysis revealed 2–4 groups of topologies in *Hyalella*, *Pseudoniphargus*, and Primates (Fig. 7b, d, e), but 6–8 in Collembola, *Metacrangonyx*, Mammals, and Vertebrates (Fig. 7a, c, f, g). Trees inferred under standard and Lie Markov models typically fell close together, as did those generated before and after merging identical partition schemes. By contrast, unpartitioned trees (models 21 and 24 in Fig. 7) were the most divergent from the focal tree. Previously published trees (tree 37, see Material and Methods) also showed strong divergence in Collembola and *Metacrangonyx*, but not in *Hyalella* or *Pseudoniphargus*.

DISCUSSION

The integration of TM region information into phylogenetic analyses has so far been predominantly restricted to applications in orthology detection within protein families (Bjarnadóttir et al. 2006; Wojtkowska et al. 2012). By contrast, other structural features—such as residue-residue contacts and relative solvent accessibility—have been employed more broadly to improve the accuracy of MSAs (e.g., Deng and Cheng 2011). In this context, we present TRAMPO, a novel pipeline that incorporates secondary structural elements—specifically, transmembrane helices (TM), matrix-facing (MA), and intermembrane-facing (IM)

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

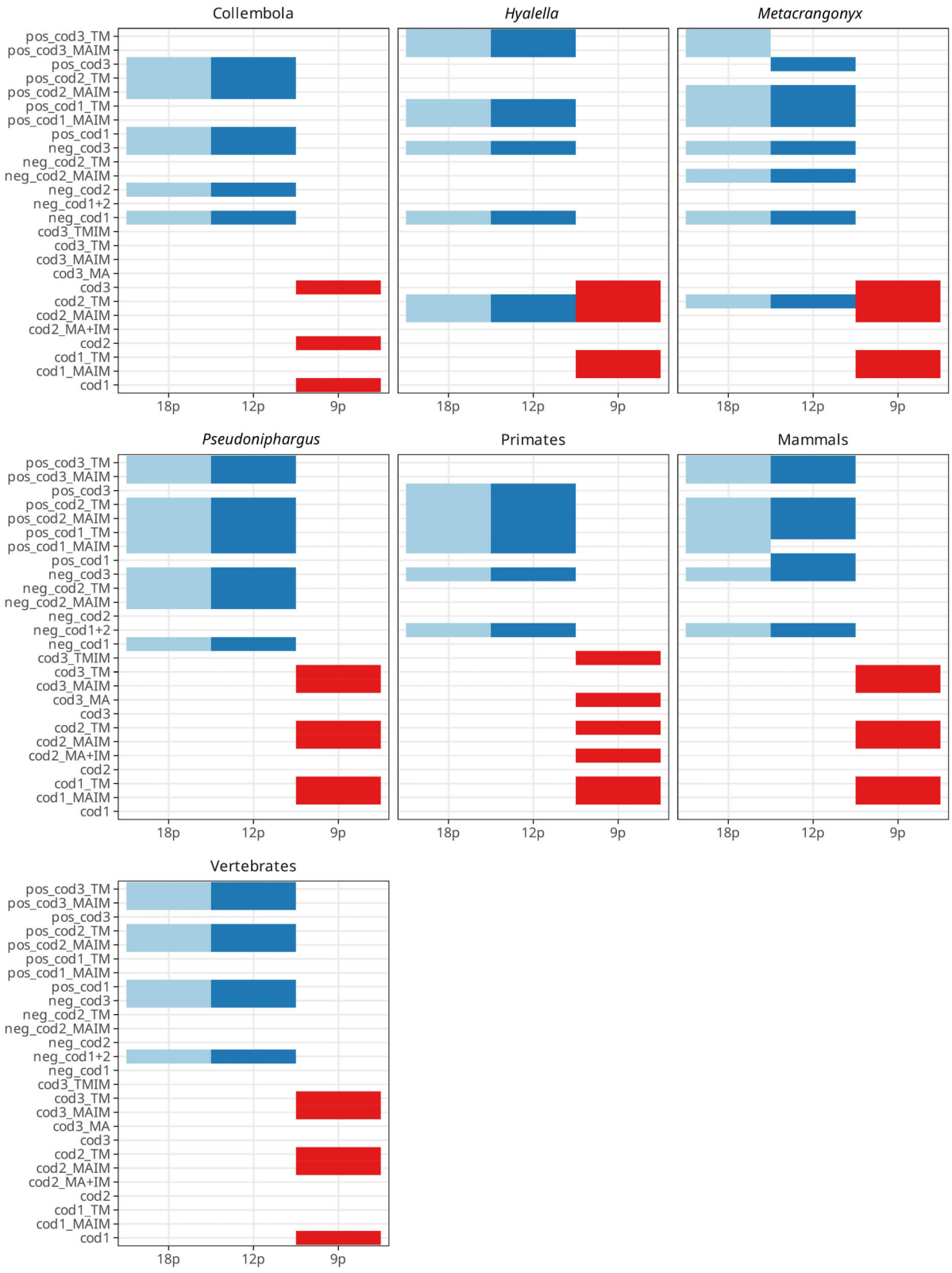


FIGURE 5. Partitions identified by PartitionFinder 2 across seven lineages under three alternative schemes. Light blue indicates site sets retained under the 18-partition analyses (18p), dark blue those from the 12-partition analyses (12p), and maroon those from the 9-partition analyses (9p). See Table 2 and Table 3 for details.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

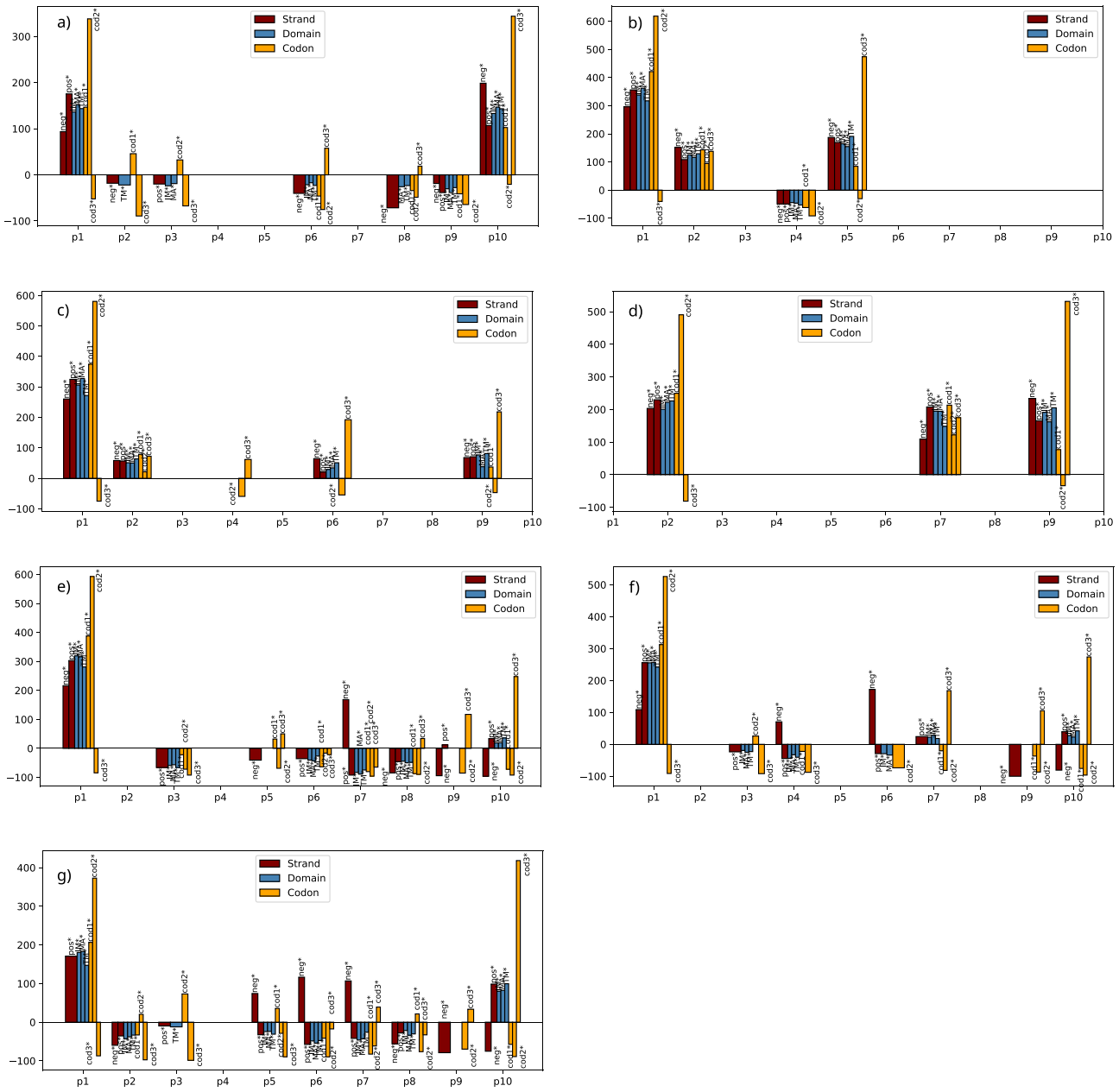


FIGURE 6. Bar plots showing General Heterogeneous evolution On a Single Topology (GHOST) categories (p1-p10) with significant deviations from expected values, based on standardized residuals with Bonferroni correction, across strand orientation, transmembrane domains, and codon positions in the seven studied lineages: (a) Collembola, (b) *Hyalella*, (c) *Metacrangonyx*, (d) *Pseudoniphargus*, (e) Primates, (f) Mammals, and (g) Vertebrates. For categories with significant deviations, the y-axis indicates the percentage by which they are over- or under-represented relative to expectation.

loop domains—into phylogenetic partitioning schemes, aiming to improve model fit and tree inference.

While the approach has broad potential applications, TRAMPO is specifically optimized for the 13 mitochondrial PCGs, whose products are rich in transmembrane helices. Although mitochondrial genes are sometimes dismissed as outdated markers in the era of phylogenomics, the growing accessibility of sequencing, assembly, and annotation tools (Bernt et al. 2013; Dierckxsens et al. 2017; Cucini et al. 2021b) has revitalized their utility (Blair 2023). Thanks to their conserved structure, easily traceable orthology, and compact genome size, mitochondria remain an invaluable phylogenetic resource—particularly in underexplored taxa. Nonetheless, some caution is advised, as their high mutation rates and the risk of substitutional saturation can, in some cases, com-

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

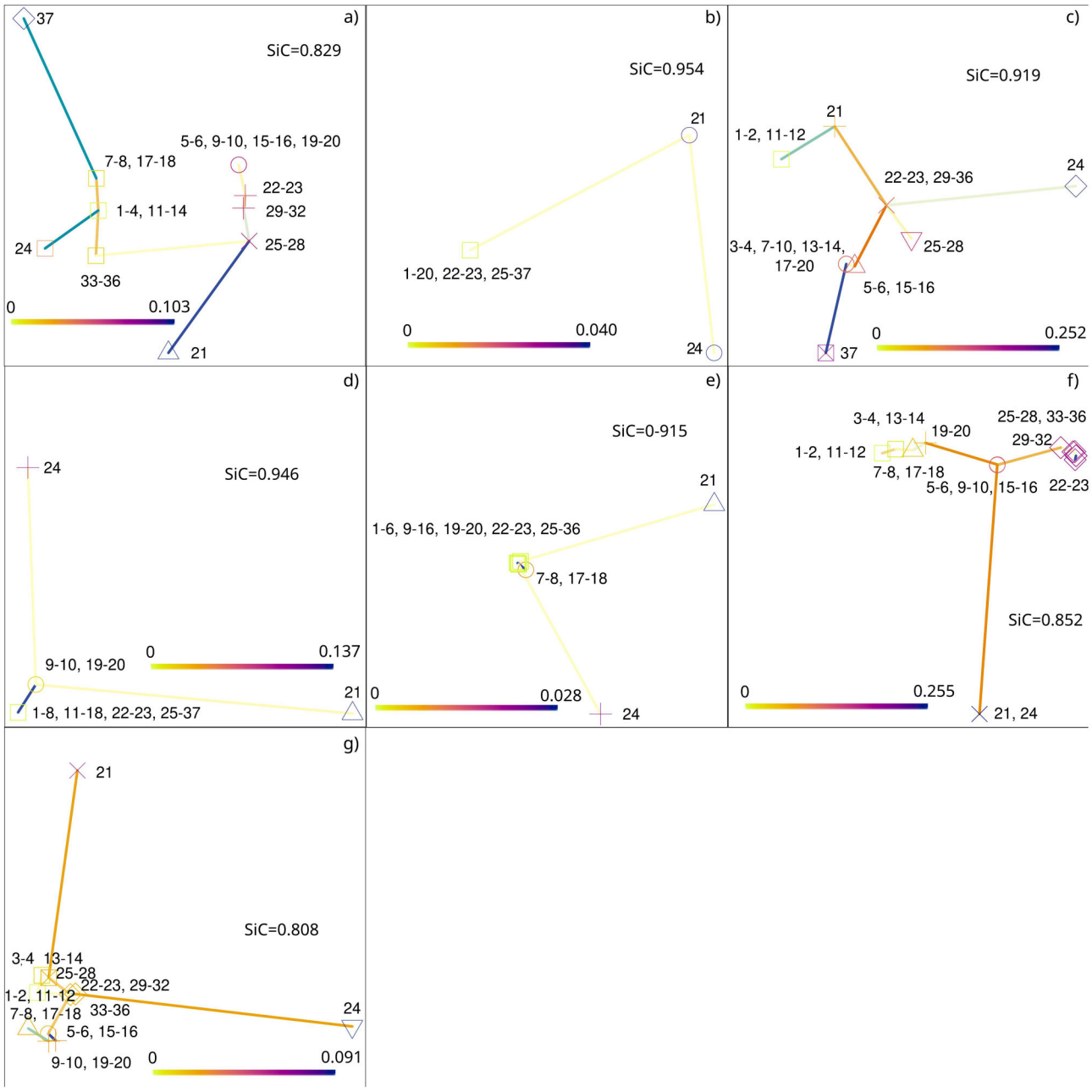


FIGURE 7. Principal Coordinates Analysis (PCoA) of tree space based on quartet distances for seven lineages. Each point represents a tree inferred under a specific partitioning scheme, with tree numbers corresponding to the schemes listed in Table 3. Clusters of trees, identified by k-means, are shown with different symbols. Colors in lines indicate the quartet distance of each tree from the focal topology, defined as that obtained from the optimal merging of 18 partitions for each lineage. (a) Collembola, (b) *Hyalella*, (c) *Metacrangonyx*, (d) *Pseudoniphargus*, (e) Primates, (f) Mammals, and (g) Vertebrates). Previously published topologies (tree 37) are included for *Hyalella*, *Metacrangonyx*, *Pseudoniphargus*, and Collembola (see Materials and Methods).

plicate inference at deep phylogenetic levels, as illustrated in placental mammals (Morgan et al. 2013).

In line with these considerations, our results confirm previously reported patterns of compositional bias and among-site rate heterogeneity across codon positions (Pons et al. 2010; Cucini et al. 2021a) and encoding strands (Hassanin et al. 2005). Crucially, we extend these findings by showing that compositional biases also correlate with secondary structural domains. Specifically, second codon positions of codons that encode TM helices exhibited strong thymine enrichment—a pattern not observed in MA and IM domains (Fig. 2; Supplementary Fig. S2-S3). At the amino-acid level, TM domains were consistently enriched in hydropho-

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

bic residues and depleted in acidic, amide, and basic residues (Fig. 3). These trends were evident across all seven metazoan lineages studied here, suggesting that the association between secondary structure and compositional bias is a conserved feature of mitochondrial genomes.

Consistent with this, integrating structural domain information into partitioning schemes improved model fit, as reflected in consistently lower AIC scores across all datasets. This improvement was robust, though smaller in magnitude compared to traditional partitioning based on codon position or strand orientation. It is noteworthy that structural partitioning also led to reduced computational demands—particularly compared to mixture models—while still outperforming GHOST in all but the Vertebrates dataset. Models employing discrete FreeRates categories outperformed those using gamma distributions or invariant sites, in line with prior findings. Although Lie Markov models produced slightly lower AIC values, they required substantially longer runtimes and had limited influence on tree topology. Despite these differences, the topologies inferred under both standard and Lie Markov models remained fully congruent, clustering within the same region of tree space.

These results connect directly with earlier phylogenetic practice: codon and strand-based partitioning have long been standard in mitochondrial phylogenetics, improving tree reconstruction across diverse taxa including beetles (Timmermans et al. 2010), springtails (Nardi et al. 2020), grasshoppers (Fenn et al. 2008), hymenopterans (Song et al. 2016), crustaceans (Bauzà-Ribot et al. 2012), birds (Powell et al. 2013), and bovids (Bover et al. 2019). Our study builds on this foundation by introducing a third axis of biologically informed partitioning: structural domains. PartitionFinder 2 consistently retained TM second codon positions as separate partitions, highlighting their distinct evolutionary signal (Fig. 5). In vertebrates, however, structural partitioning was constrained by the limited representation of TM regions on the negative strand—since only one protein-coding gene (*nad6*) is transcribed from it. Consequently, PartitionFinder 2 grouped all second codon positions from the negative strand into a single partition, as expected under its greedy algorithm when signal is weak or limited. Nonetheless, methods such as PartitionFinder 2 have limitations, as they do not subdivide initial partitions into smaller subsets, and information-theoretic criteria (e.g., AIC, BIC) used in these approaches do not formally test whether resulting partitions are compositionally homogeneous (Susko and Roger 2020; Crotty and Holland 2022).

Further evidence for the biological meaning of structural domains comes from GHOST models. The distribution of site-level probabilities assigned by GHOST models also supported the biological relevance of codon, encoding strand, and structural partitions. Certain GHOST categories were significantly unevenly distributed with certain categories of sites over- or underrepresented in specific regions, especially at the codon level, with more modest effects across encoding strands and structural domains (Fig. 6). These patterns mirror our compositional and phylogenetic findings, reinforcing the idea that secondary structure contributes meaningfully to mitochondrial genome evolution.

In addition, earlier studies, such as Pandey and Braun (2020), demonstrated that secondary structure can influence phylogenetic inference, especially at deep divergences. Using a structurally partitioned dataset, they found conflicting placements of ctenophores and sponges depending on whether buried or exposed residues were analysed. In our study, structural partitioning influenced topology to varying degrees. For *Hyalella* and *Pseudoniphargus*, tree topologies (Fig. 7) remained stable despite large likelihood differences (Fig. 4). Conversely, Collembola, Mammals, and Primates exhibited minor topological shifts, while *Metacrangonyx* and Vertebrates showed greater variation across partitioning schemes.

Surprisingly, phylogenetic informativeness—measured by AIC score reduction—did not improve in Mammals and Vertebrates after incorporating structural information, despite their strong thymine bias at second codon positions of codons that encode TM helices. TM helices at second positions were retained as independent partitions, suggesting they harbour meaningful phylogenetic signal. However, in these deep-diverging clades (>300 Ma), the impact of structural signal may be obscured by extensive substitution saturation—especially at third and, to a lesser extent, first codon positions. Additionally, the near-universal positive strand coding of vertebrate mitochondrial genes limits the contribution of strand-specific signal.

Furthermore, we observed a relationship between lineage age and the size of tree space. Recent lineages (e.g., *Hyalella*, ∼25 Ma (Zapelloni et al. 2021); *Pseudoniphargus*, ∼50 Ma (Stokkan et al. 2018); Primates, ∼75 Ma (Pozzi et al. 2014)) produced fewer and more tightly clustered topologies, whereas older lineages (*Metacrangonyx*, ∼125 Ma (Pons et al. 2019); Vertebrates, ∼440 Ma (Brazeau and Friedman 2015)) displayed greater topological dispersion, except for Mammals (∼220 Ma; Luo 2007) (Fig. 7). This pattern was not explained by taxon sampling. For example, *Metacrangonyx* (23 taxa) and Vertebrates (150 taxa) had similar tree space breadths, suggesting that phylogenetic ambiguity increases with evolutionary depth. Mammals and Vertebrates groups, particularly the latter, appear to include highly divergent and ancient lineages, in which heterotachy and heterogeneity among species may pose greater challenges to phylogenetic resolution than heterogeneity within or among genes.

These findings support the expected time-dependent roles of codon positions: third codon positions, with their rapid evolution, are informative over short timescales, while second positions—

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

constrained yet compositionally biased—gain importance over longer evolutionary periods. In this context, the thymine enrichment at second codon positions of codons that encode TM helices likely drives much of the signal captured by structural partitioning, particularly in *Metacrangonyx*, which also exhibit elevated substitution rates (Bauzà-Ribot et al. 2012; Pons et al. 2019).

The discordance in Vertebrates, where the best topology arose under codon + strand partitioning despite the lowest AIC being achieved by GHOST in an unpartitioned analysis, suggests potential limitations of AIC for comparing models with different structures. AIC may not be optimal for evaluating complex mixture models like GHOST relative to explicitly partitioned models incorporating rate heterogeneity (e.g., Gamma, FreeRates).

In the broader context of phylogenetic methodology, we recognize that, although information criteria measures can be applied readily across datasets and lineages, they also have several limitations: (i) they assess only relative model fit rather than absolute accuracy, (ii) they do not provide a direct statistical test of tree topology, (iii) they may be sensitive to sample size and model complexity penalties, and (iv) they cannot fully account for biological realism, such as site dependencies or structural constraints. More robust approaches—such as cross-validation, bootstrapping, or other simulation-based testing approaches (Liu et al. 2023) and parametric bootstrapping (Jermiin et al. 2020)—may provide more reliable criteria for selecting among competing hypotheses, particularly when evaluating mixture models. A further development of the analyses presented here could include cross-validation on nuclear single-copy genes, as structural information on nuclear proteins at both the secondary and tertiary levels is becoming increasingly available.

Conclusion

Based on our results, we recommend incorporating secondary structural domain information—particularly transmembrane helices and loop regions—into partitioning schemes alongside codon position and strand orientation when conducting phylogenetic analyses of mitochondrial PCGs. Our conclusions, while grounded in model selection criteria (e.g., AIC), ultimately rely on the assumption that these models accurately reflect the evolutionary process. Given that true evolutionary histories are unknown, simulation-based validation remains ideal but challenging—particularly when modelling realistic features such as transmembrane domains and loop regions. Cross-validation may serve as an alternative, though its application at this scale remains computationally intensive (Liu et al. 2023). Additionally, we advocate for merging MA and IM loop domains into a single partition, as this strategy produced comparable or improved AIC scores while reducing model complexity. Structural partitioning consistently outperformed GHOST, except for Vertebrates, and mixture models in terms of model fit and computational efficiency. The benefits of this approach are particularly evident in lineages older than 50 million years, with intermediate ancient lineages showing substantial advantages, although these benefits may diminish in very ancient lineages exceeding 300 million years due to increased substitution saturation. TRAMPO enables straightforward implementation of these strategies and is freely available at https://github.com/dbajpp0/TRAMPO.

Acknowledgments

Joan Pons is indebted to Professor Carlos Juan for the long discussions we had, more than ten years ago, about the idea of using secondary structures of mitochondrial proteins in a phylogenetic framework. Trempó (pronounced trəm'po in Catalan), incorrectly written trampó in Spanish, is a simple summer salad typical from Mallorca made with tomatoes, pale green peppers and onions, dressed with salt and extra virgin olive oil, that blends the colours of the Italian flag. We are grateful to the editor and the three anonymous reviewers for their insightful and constructive comments, which substantially enhanced the quality of the manuscript.

Supplementary Material

Supplementary material is available at *Systematic Biology* online.

Conflict of Interest

Authors declare no conflicts of interest.

Funding

This work was funded to J.P. by the Conselleria de Fons Europeus, Universitat i Cultura, Direcció General de Política Universitària i Recerca of Govern Balear (Balearic Islands, Spain) under the call "Projectes de Recerca Científica i Tecnològica 2020-2024". The project, titled "Evolutionary signatures of past climate change and island isolation in the genome of *Myotragus balearicus*", was awarded with grant number PDR2020/45. C.C. was supported by the PhD program in Life Sciences at the University of Siena and by the ERASMUS + 2021/2022 fellowship.

Data Availability

All the original data and scripts necessary to reproduce the analyses reported in this study can be accessed through the Dryad link: https://doi.org/10.5061/dryad.wh70rxx29.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

References

Abadi S., Azouri D., Pupko T., Mayrose I. 2019. Model selection may not be a mandatory step for phylogeny reconstruction. Nat. Commun. 10(1): 934. https://doi.org/10.1038/s41467-019-08822-w.

Akaike H. 1973. Information theory and an extension of the maximum likelihood principle. In Petrov B. N., Csáki F., editors. 2nd International Symposium on Information Theory Tsahkadsor Armenia USSR, September 2-8, 1971. Budapest: Akadémiai Kiadó, p.267–281.

Barbera S., Cucini C. 2022. A. glimpse into the past: phylogenesis and protein domain analysis of the group XIV of C.-type lectins in vertebrates. Bmc Genomics [Electronic Resource] 23(1):420. https://doi.org/10.1186/s12864-022-08659-6.

Bauzà-Ribot M.M., Juan C., Nardi F., Oromí P., Pons J., Jaume D. 2012. Mitogenomic phylogenetic analysis supports continental-scale vicariance in subterranean thalassoid crustaceans. Curr. Biol. 22(21):2069–2074. https://doi.org/10.1016/j.cub.2012.09.012.

Bernt M., Donath A., Jühling F., Externbrink F., Florentz C., Fritzsch G., Pütz J., Middendorf M., Stadler P.F. 2013. MITOS: improved *de novo* metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69(2):313–319. https://doi.org/10.1016/j.ympev.2012.08.023.

Bjarnadóttir T.K., Gloriam D.E., Hellstrand S.H., Kristiansson H., Fredriksson R., Schiöth H.B. 2006. Comprehensive repertoire and phylogenetic analysis of the G. protein-coupled receptors in human and mouse. Genomics 88(3):263–273. https://doi.org/10.1016/j.ygeno.2006.04.001.

Blair C. 2023. Organellar DNA continues to provide a rich source of information in the genomics era. Mol. Ecol. 32(9):2144–2150. https://doi.org/10.1111/mec.16872.

Boutet E., Lieberherr D., Tognolli M., Schneider M., Bairoch A. 2007. UniProtKB /Swiss-Prot. Methods Mol. Biol. 406:89–112. https://doi.org/10.1007/978-1-59745-535-0_4.

Bover P., Llamas B., Mitchell K.J., Thomson V.A., Alcover J.A., Lalueza-Fox C., Cooper A., Pons J. 2019. Unraveling the phylogenetic relationships of the extinct bovid *Myotragus balearicus* Bate 1909 from the Balearic Islands. Quat. Sci. Rev. 215:185–195. https://doi.org/10.1016/j.quascirev.2019.05.005.

Brazeau M.D., Friedman M. 2015. The origin and early phylogenetic history of jawed vertebrates. Nature 520(7548):490–497. https://doi.org/10.1038/nature14438.

Crotty S.M., Holland B.R. 2022. Comparing partitioned models to mixture models: do information criteria apply? Syst. Biol. 71(6):1541–1548. https://doi.org/10.1093/sysbio/syac003.

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. https://doi.org/10.1093/sysbio/syz051.

Cucini C., Fanciulli P.P., Frati F., Convey P., Nardi F., Carapelli A. 2021a. Re-evaluating the internal phylogenetic relationships of Collembola by means of mitogenome data. Genes 12(1):44. https://doi.org/10.3390/genes12010044.

Cucini C., Leo C., Iannotti N., Boschi S., Brunetti C., Pons J., Fanciulli P.P., Frati F., Carapelli A., Nardi F. 2021b. EZmito: a simple and fast tool for multiple mitogenome analyses. Mitochondrial DNA Part B. 6(3):1101–1109. https://doi.org/10.1080/23802359.2021.1899865.

Dayhoff M.O., Schwartz R.M., Orcutt B.C. 1978. A model of evolutionary change in proteins. In: Dayhoff M.O., editor. Edition Atlas of Protein Sequence and Structure. Natl. Biomed. Res. Found. Washington DC, p. 345–352.

Deng X., Cheng J. 2011. MSACompro: protein multiple sequence alignment using predicted secondary structure, solvent accessibility, and residue-residue contacts. BMC Bioinf. 12(1):472. https://doi.org/10.1186/1471-2105-12-472.

Dennerlein S., Rehling P. 2015. Human mitochondrial COX1 assembly into cytochrome c oxidase at a glance. J. Cell Sci. 128:833–837. https://doi.org/10.1242/jcs.161729.

Dierckxsens N., Mardulyn P., Smits G. 2017. NOVO plasty: de novo assembly of organelle genomes from whole genome data. Nucleic. Acids. Res. 45:gkw955. https://doi.org/10.1093/nar/gkw955.

Dobson L., Reményi I., Tusnády G.E. 2015. CCTOP: a Consensus Constrained TOPology prediction web server. Nucleic. Acids. Res. 43(W1):W408–W412. https://doi.org/10.1093/nar/gkv451.

Erpenbeck D., Nichols S.A., Voigt O., Dohrmann M., Degnan B.M., Hooper J.N., Wörheide G. 2007. Phylogenetic analyses under secondary structure-specific substitution models outperform traditional approaches: case studies with diploblast LSU. J. Mol. Evol. 64(5):543–557. https://doi.org/10.1007/s00239-006-0146-3.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading Syst. Biol. 27(4):401–410. https://doi.org/10.1093/sysbio/27.4.401.

Fenn J.D., Song H., Cameron S.L., Whiting M.F. 2008. A. preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. Mol. Phylogenet. Evol. 49(1):59–68. https://doi.org/10.1016/j.ympev.2008.07.004.

Foster P.G. 2004. Modeling compositional heterogeneity. Syst. Biol. 53(3): 485–495. https://doi.org/10.1080/10635150490445779.

Gowri-Shankar V., Rattray M. 2006. On the correlation between composition and site-specific evolutionary rate: implications for phylogenetic inference. Mol. Biol. Evol. 23(2):352–364. https://doi.org/10.1093/molbev/msj040.

Hallgren J., Tsirigos K.D., Pedersen M.D., Almagro-Armenteros J.J., Marcatili P., Nielsen H., Krogh A., Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. BioRxiv 2022.04.08.487609. https://doi.org/10.1101/2022.04.08.487609. bioRxiv

Hassanin A., Léger N., Deutsch J. 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoa, and consequences for phylogenetic inferences. Syst. Biol. 54(2):277–298. https://doi.org/10.1080/10635150590947843.

Ho S.Y.W., Jermiin L.S. 2004. Tracing the decay of the historical signal in biological sequence data. Syst. Biol. 53,623–637. https://doi.org/10.1080/10635150490468607.

Hunter J.D. 2007. Matplotlib: a 2D graphics environment. Comput. Sci. Eng. 9(3):90–95. https://doi.org/10.1109/MCSE.2007.55.

Jayaswal V., Ababneh F., Jermiin L.S., Robinson J. 2014. Mixture models of nucleotide sequence evolution that account for heterogeneity in the substitution process across sites and across lineages. Syst. Biol. 63:726–742. https://doi.org/10.1093/sysbio/syu03.

Jermiin L.S., Catullo R.A., Holland B.R. 2020. A new phylogenetic protocol: dealing with model misspecification and confirmation bias in molecular phylogenetics, NAR Genom. Bioinform. 2(2):qaa041. https://doi.org/10.1093/nargab/lqaa041

Jermiin L.S., Graur D., Lowe R.M., Crozier R.H. 1994. Analysis of directional mutation pressure and nucleotide content in mitochondrial cytochrome b genes. J. Mol. Evol. 39(2):160–173. https://doi.org/10.1007/BF00163805.

Jermiin L.S., Ho S.Y.W., Ababneh F., Robinson J., Larkum A.W.D. 2004. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst. Biol. 53,(4): 638–643. https://doi.org/10.1080/10635150490468648.

Jermiin L.S., Jayaswal V., Ababneh F., Robinson J. 2017. Phylogenetic Model Evaluation. In Keith J. M. editor. Bioinformatics: Volume I.: Data Sequence Analysis and Evolution(2nd ed.). New York: Humana Press, p.379–420. https://doi.org/10.1007/978-1-4939-6622-6_15.

Kainer D., Lanfear R. 2015. The effects of partitioning on phylogenetic inference Mol. Biol. Evol. 32(6):1611–1627https://doi.org/10.1093/molbev/msv026.

Kalyaanamoorthy S., Minh B.Q., Wong T.K., von Haeseler A., Jermiin L.S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14(6):587–589. https://doi.org/10.1038/nmeth.4285.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30(4):772–780. https://doi.org/10.1093/molbev/mst010.

Kruskal J.B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika. 29(1):1–27.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

Lanfear R., Calcott B., Ho S.Y., Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29(6):1695–1701. https://doi.org/10.1093/molbev/mss020.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34:772–773. https://doi.org/10.1093/molbev/msw260.

Le S.Q., Gascuel O. 2010. Accounting for solvent accessibility and secondary structure in protein phylogenetics is clearly beneficial. Syst. Biol. 59(3):277–287. https://doi.org/10.1093/sysbio/syq002.

Le S.Q., Lartillot N., Gascuel O. 2008. Phylogenetic mixture models for proteins. Phil. Trans. R. Soc. B 363(1512):3965–3976. https://doi.org/10.1098/rstb.2008.0180.

Leavitt J.R., Hiatt K.D., Whiting M.F., Song H. 2013. Searching for the optimal data partitioning strategy in mitochondrial phylogenomics: a phylogeny of Acridoidea (Insecta: orthoptera: caelifera) as a case study. Mol. Phylogenet. Evol. 67(2):494–508. https://doi.org/10.1016/j.ympev.2013.02.019.

Liu Q., Charleston M.A., Richards S.A., Holland B.R. 2023. Performance of AIC and BIC in selecting partition models and mixture models. Syst. Biol. 72(1):92–105. https://doi.org/10.1093/sysbio/syac081.

Lloyd R.E., McGeehan J.E. 2013. Structural analysis of mitochondrial mutations reveals a role for bigenomic protein interactions in human disease. PLoS One 8(7):e69003. https://doi.org/10.1371/journal.pone.0069003.

Luo Z.X. 2007. Transformation and diversification in early mammal evolution. Nature 450,(7172): 1011–1019. https://doi.org/10.1038/nature06277.

Manuel C., Sakalli E., Schmidt H.A., Viñas C., von Haeseler A., Elgert C. 2025. When the past fades: detecting phylogenetic signal with SatuTe. Mol. Biol. Evol. 42(5):msaf090. https://doi.org/10.1093/molbev/msaf090.

Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A., Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015.

Minin V., Abdo Z., Joyce P., Sullivan J. 2003. Performance-based selection of likelihood models for phylogeny estimation. Syst. Biol. 52(5):674–683. https://doi.org/10.1080/10635150390235494.

Möller S., Croning M.D.R., Apweiler R. 2001. Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics 17(7):646–653. https://doi.org/10.1093/bioinformatics/17.7.646.

Morgan C.C., Foster P.G., Webb A.E., Pisani D., McInerney J.O., O'Connell M.J. 2013. Heterogeneous models place the root of the placental mammal phylogeny. Mol. Biol. Evol. 30(9):2145–2156. https://doi.org/10.1093/molbev/mst117.

Muse S.V., Gaut B.S. 1994. A. likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol. Biol. Evol. 11:715–724. https://doi.org/10.1093/oxfordjournals.molbev.a040152.

Narakusumo R.P., Riedel A., Pons J. 2020. Mitochondrial genomes of twelve species of hyperdiverse *Trigonopterus* weevils. PeerJ. 8:e10017. https://doi.org/10.7717/peerj.10017.

Nardi F., Cucini C., Leo C., Frati F., Fanciulli P.P., Carapelli A. 2020. The complete mitochondrial genome of the springtail *Allacma fusca*, the internal phylogenetic relationships and gene order of Symphypleona. Mitochondrial DNA Part B 5(3):3103–3105. https://doi.org/10.1080/23802359.2020.1800425.

Pandey A., Braun E.L. 2020. Phylogenetic analyses of sites in different protein structural environments result in distinct placements of the metazoan root. Biology 9(4):64. https://doi.org/10.3390/biology9040064.

Pons J., Bauzà-Ribot M.M., Jaume D., Juan C. 2014. Next-generation sequencing, phylogenetic signal and comparative mitogenomic analyses in Metacrangonyctidae (Amphipoda: crustacea) Pons Bmc Genomics [Electronic Resource] 15(1):566. https://doi.org/10.1186/1471-2164-15-566.

Pons J., Jurado-Rivera J.A., Jaume D., Vonk R., Bauzà-Ribot M.M., Juan C. 2019. The age and diversification of Metacrangonyctid subterranean amphipod crustaceans revisited. Mol. Phylogenet. Evol. 140:106599. https://doi.org/10.1016/j.ympev.2019.106599.

Pons J., Ribera I., Bertranpetit J., Balke M. 2010. Nucleotide substitution rates for the full set of mitochondrial protein-coding genes in Coleoptera. Mol. Phylogenet. Evol. 56(2):796–807. https://doi.org/10.1016/j.ympev.2010.02.007.

Posada D., Buckley T.R. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53(5):793–808. https://doi.org/10.1080/10635150490522304.

Posada D., Crandall K.A. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14(9):817–818. https://doi.org/10.1093/Bioinformatics/14.9.817.

Powell A.F.L.A., Barker F.K., Lanyon S.M. 2013. Empirical evaluation of partitioning schemes for phylogenetic analyses of mitogenomic data: an avian case study. Mol. Phylogenet. Evol. 66(1):69–79. https://doi.org/10.1016/j.ympev.2012.09.006.

Pozzi L., Hodgson J.A., Burrell A.S., Sterner K.N., Raaum R.L., Disotell T.R. 2014. Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. Mol. Phylogenet. Evol. 75:165–183. https://doi.org/10.1016/j.ympev.2014.02.023.

Puente-Lelievre C., Malik A.J., Douglas J., Ascher D., Baker M., Allison J., Poole A., Lundin D., Fullmer M., Bouckert R., Kim H., Steinegger M., Matzke N. 2023. Tertiary-interaction characters enable fast, model-based structural phylogenetics beyond the twilight zone. BioRxiv 2023.12.12.571181. https://doi.org/https://doi.org/10.1101/2023.12.12.571181. bioRxiv

Ren H., Wong T.K.F., Minh B.Q., Lanfear R. 2025. MixtureFinder: estimating DNA mixture models for phylogenetic analyses. Mol. Biol. Evol. 42(1):msae264, https://doi.org/10.1093/molbev/msae264.

Sammon J.W. 1969. A nonlinear mapping for data structure analysis. IEEE Trans. Comput. C-18(5):401–409.

Schwarz G.E. 1978. Estimating the dimension of a model. Ann. Statist. 6(2): 461–464, https://doi.org/10.1214/aos/1176344136, MR 0468014.

Shapiro B., Rambaut A., Drummond A.J. 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol. Biol. Evol. 23(1):7–9. https://doi.org/10.1093/molbev/msj021.

Sharp P.M., Li W.H. 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. Nucl Acids Res. 15(3):1281–1295. https://doi.org/10.1093/nar/15.3.1281.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51(3):492–508. https://doi.org/10.1080/10635150290069913.

Smith D.R. 2012. Updating our view of organelle genome nucleotide landscape. Front. Gene. 3:175. https://doi.org/10.3389/fgene.2012.00175.

Smith M.R. 2022. Robust analysis of phylogenetic tree space. Syst. Biol. 71(5): 1255–1270. https://doi.org/10.1093/sysbio/syab100.

Song S.N., Tang P., Wei S.J., Chen X.X. 2016. Comparative and phylogenetic analysis of the mitochondrial genomes in basal hymenopterans. Sci. Rep. 6(1):20972. https://doi.org/10.1038/srep20972.

Soubrier J., Steel M., Lee M.S.Y., Der Sarkissian C., Guindon S., Ho S.Y.W., Cooper A. 2012. The influence of rate heterogeneity among sites on the time dependence of molecular rates. Mol. Biol. Evol. 29(11):3345–3358. https://doi.org/10.1093/molbev/mss140.

Steiner F.M., Pautasso M., Zettel H., Moder K., Arthofer W., Schlick-Steiner B.C. 2015. A. new hierarchy of phylogenetic models consistent with heterogeneous substitution rates. Syst. Biol. 64(5):860–868. https://doi.org/10.1093/sysbio/syv026.

Stokkan M., Jurado-Rivera J.A., Oromí P., Juan C., Jaume D., Pons J. 2018. Species delimitation and mitogenome phylogenetics in the subterranean genus *Pseudoniphargus* (Crustacea: amphipoda). Mol. Phylogenet. Evol. 127:988–999. https://doi.org/10.1016/j.ympev.2018.07.002.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

Sueoka N. 1995. Intrastrand parity rules of DNA base composition and usage biases of synonymous codons. J. Mol. Evol. 40(3): 318–325. https://doi.org/10.1007/BF00163236.
Sullivan J., Joyce P. 2005. Model selection in phylogenetics. Annu. Rev. Ecol. Evol. Syst. 36:445–466. https://doi.org/10.1146/annurev.ecolsys.36.102003.152645.
Sullivan J., Swofford D.L., Naylor G.J.P. 1999. The effect of taxon sampling on estimating rate heterogeneity parameters of maximum-likelihood models. Mol. Biol. Evol. 16(10): 1347–1356. https://doi.org/10.1093/oxfordjournals.molbev.a026045.
Susko E., Roger A.J. 2020. On the use of information criteria for model selection in phylogenetics. Mol. Biol. Evol. 37(2): 549–562. https://doi.org/10.1093/molbev/msz228.
Timmermans M., Dodsworth S., Culverwell L., Bocak L., Ahrens D., Littlewood D.T., Pons J., Vogler A.P. 2010. Why barcode? High-throughput multiplex sequencing of mitochondrial genomes for molecular systematics. Nucleic. Acids. Res. 38(21):e197. https://doi.org/10.1093/nar/gkq807.
Tusnády G.E., Simon I. 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics 17(9):849–850. https://doi.org/10.1093/bioinformatics/17.9.849.
Wojtkowska M., Jąkalski M., Pieńkowska J.R., Stobienia O., Karachitos A., Przytycka T.M., Weiner J., Kmita H., Makałowski W. 2012. Phylogenetic analysis of mitochondrial outer membrane β-barrel channels. Genome Biol. Evol. 4(2):110–125. https://doi.org/10.1093/gbe/evr130.
Yang Z. 1995. A space-time process model for the evolution of DNA sequences. Genetics 139(2):993–1005. https://doi.org/10.1093/genetics/139.2.993.
Zapelloni F., Pons J., Jurado-Rivera J.A., Jaume D., Juan C. 2021. Phylogenomics of the *Hyalella* amphipod species-flock of the Andean Altiplano. Sci. Rep. 11(1):366. https://doi.org/10.1038/s41598-020-79620-4.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syag027/8512886 by Biblioteca centrale Ingegneria Siena user on 01 April 2026

Syst. Biol. 0(0):1–19, 2026
© The Author(s) 2026. Published by Oxford University Press on behalf of the Society of Systematic Biologists. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
https://doi.org/10.1093/sysbio/syag027
Advance Access publication March 10, 2026