RESEARCH ARTICLE

Rapid evolution of mitochondrion-related genes in haplodiploid arthropods

Yiyuan Li[1]* , Gregg W. C. Thomas[2,3,4], Stephen Richards[5], Robert M. Waterhouse[6], Xin Zhou[7] and Michael E. Pfrender[8,9]

Abstract

Background Mitochondrial genes and nuclear genes cooperate closely to maintain the functions of mitochondria, especially in the oxidative phosphorylation (OXPHOS) pathway. However, mitochondrial genes among arthropod lineages have dramatic evolutionary rate differences. Haplodiploid arthropods often show fast-evolving mitochondrial genes. One hypothesis predicts that the small effective population size of haplodiploid species could enhance the effect of genetic drift leading to higher substitution rates in mitochondrial and nuclear genes. Alternatively, positive selection or compensatory changes in nuclear OXPHOS genes could lead to the fast-evolving mitochondrial genes. However, due to the limited number of arthropod genomes, the rates of evolution for nuclear genes in haplodiploid species, besides hymenopterans, are largely unknown. To test these hypotheses, we used data from 76 arthropod genomes, including 5 independently evolved haplodiploid lineages, to estimate the evolutionary rates and patterns of gene family turnover of mitochondrial and nuclear genes.

Results We show that five haplodiploid lineages tested here have fast-evolving mitochondrial genes and fast-evolving nuclear genes related to mitochondrial functions, while nuclear genes not related to mitochondrion showed no significant evolutionary rate differences. Among hymenopterans, bees and ants show faster rates of molecular evolution in mitochondrial genes and mitochondrion-related nuclear genes than sawflies and wasps. With genome data, we also find gene family expansions and contractions in mitochondrion-related genes of bees and ants.

Conclusions Our results reject the small population size hypothesis in haplodiploid species. A combination of positive selection and compensatory changes could lead to the observed patterns in haplodiploid species. The elevated evolutionary rates in OXPHOS complex 2 genes of bees and ants suggest a unique evolutionary history of social hymenopterans.

Keywords Gene family evolution, Evolutionary rate, Gene duplication, Oxidative phosphorylation genes, Hymenoptera, Thysanoptera, Phthiraptera, Mesostigmata, Trombidiformes

*Correspondence:
Yiyuan Li
lyy0005@gmail.com
Full list of author information is available at the end of the article

Li *et al. BMC Biology*   (2024) 22:229

Page 2 of 14

Background

Eukaryotes have repeatedly acquired intimate symbiotic microorganisms, including mitochondria, plastids, and microbiome, which provide great benefits to the eukaryote hosts, facilitating adaptation and diversification of the hosts [1, 2]. The most ancient symbionts, mitochondria, have formed obligate interdependent relationships with their eukaryote hosts. Mitochondrial genes and nuclear genes cooperate closely with each other in oxidative phosphorylation, as well as maintaining mitochondrial replication, transcription, and translation [3]. Mitochondrial genomes in animals have higher mutation rates and lower efficacy of selection than nuclear genomes [4–7]. As a result, mitochondrial genomes often have faster molecular evolutionary rates than nuclear genomes [8–10]. These evolutionary rate differences can generate inter-genomic incompatibilities that have been implicated in a variety of contexts including sexual selection, hybrid breakdown, and speciation [11]. Understanding these inter-genomic interactions is a key to addressing fundamental questions in evolutionary biology, such as the evolution of gene–gene interactions, and the processes of adaptation and speciation [12–16].

Despite the functional importance of mito-nuclear interactions, mitochondrial and nuclear genes have shown dramatically different evolutionary rates among arthropods [17–19]. Species in the order Hymenoptera (sawflies, wasps, bees, and ants) have elevated evolutionary rates (amino acid substitutions per site per million years) in both mitochondrial and mitochondrion-related nuclear genes, but not for the single-copy nuclear genes, compared to other arthropod species [19–21].

Different factors could influence the evolutionary rate of mitochondrial genes, including genetic drift, selection, and mutation rate [3]. Multiple hypotheses have been proposed to explain the elevated rate of molecular evolution in hymenopterans. First, the vast majority of hymenopterans have haplodiploid sex-determination systems [22]. Many of them also have parasitic lifestyles and varying degrees of eusociality. These unique features could lead to repeated founder effects and a small effective population size, which can increase the influence of genetic drift leading to the accumulation of slightly deleterious mutations and elevated evolutionary rates in the mitochondrial and nuclear genomes [23–25]. However, the observed similar evolutionary rate of single-copy nuclear genes between hymenopterans and other insects contradicts this hypothesis. Alternatively, positive selection on the oxidative phosphorylation (OXPHOS) pathway could lead to the observed rapid evolution in both mitochondrial and nuclear OXPHOS genes [18, 20, 26]. In addition, the accumulation of deleterious mutations in mitochondrial genes could drive compensatory changes in related nuclear genes. The fixation of beneficial mutations in nuclear genes could, in turn, reduce the cost of subsequent deleterious mutations in the mitochondrial genome promoting the fixation of additional mitochondrial mutations [27, 28]. As previous studies often focus on hymenopterans, the role of haplodiploidy in the fast evolution of mitochondrial genes is not clear. In addition, previous studies [18, 19, 21] were based on the transcriptome dataset from Misof et al. [29] or conserved nuclear genes [20, 28]. In these studies, only the evolutionary rate of extremely conserved single-copy genes was tested. Therefore, it is also not clear whether less conserved nuclear genes in the genomes would show the same or different evolutionary patterns.

To test these hypotheses in other haplodiploid lineages, we use an extensive arthropod whole genome dataset of 76 arthropod species from 17 orders spanning 500 million years of evolution [30–32]. This dataset includes five independent origins of haplodiploidy [33] and two different types of haplodiploid systems, arrhenotoky and parental genome elimination (PGE). In arrhenotoky species, unfertilized eggs develop into male individuals, whereas PGE males develop from fertilized eggs but then the paternal chromosomes are removed or not passed to offspring [22, 33, 34]. Our dataset includes two independent PGE lineages, the western predatory mite (*Metaseiulus occidentalis*) [35–37] and the body louse (*Pediculus humanus*) [38], and three independent arrhenotoky lineages, the two-spotted spider mite (*Tetranychus urticae*) [39], the western flower thrips (*Frankliniella occidentalis*) [34], and hymenopterans. The inclusion of multiple independent evolutionary origins and different types of haplodiploidy greatly enhances our ability to describe the overall molecular evolutionary patterns of mitochondrial and nuclear genes in haplodiploid species. In addition, as we are using complete genome data, we can study not only single-copy nuclear genes but also 5746 orthologous nuclear genes that are found in most of the sequenced arthropod genomes. Based on this genome dataset, we investigate the evolutionary rate of mitochondrial genes and nuclear genes of haplodiploid arthropod species. We show that the five lineages of haplodiploid species all have fast-evolving genes related to the mitochondrion but not for genes unrelated to the mitochondrion. As a small population size would affect both mitochondrion-related and non-related nuclear genes, this finding rejects the hypothesis that the small effective population size of haplodiploid species increases the influence of genetic drift, leading to the fast-evolving mitochondrion-related genes. In Hymenoptera, bees and ants have even faster evolutionary rates in mitochondrion-related genes than basal hymenopterans, including sawflies and wasps, suggesting other processes could affect the evolutionary

Li *et al. BMC Biology*   (2024) 22:229

Page 3 of 14

pattern. Lastly, haplodiploid species as well as bees and ants have unique gene family expansions and contractions that could imply functional differences in these groups.

Results

Data acquisition

To explore the evolution of mitochondrion-related genes, we collected the genomes of 76 arthropod species with haplodiploid and diploid sex determination systems. We assigned species into three categories, Hymenoptera species (including 24 species from 1 order), non-Hymenoptera haplodiploid species (denoted as "other haplodiploid," including 4 species from 4 orders), and diploid species (denoted as "diploid," including 48 species from 12 orders) (Fig. 1 and Additional file 1: Table S1). Nuclear

genes and their corresponding orthologous groups were retrieved from a previous study [32]. Mitochondrial genes were downloaded from the National Center for Biotechnology Information (NCBI) if available. For the species without published mitochondrial genomes, we assembled and annotated mitochondrial genomes from whole genome shotgun sequencing data (Additional file 1: Table S1).

In total, we included 38,195 ortholog groups of nuclear genes and 13 mitochondrial genes. Among these orthologous groups, we focused on subsets pertaining to mitochondrial and nuclear interactions, including 13 mitochondrion-encoded OXPHOS genes (mtOXPHOS), 65 nuclear-encoded OXPHOS genes (nucOXPHOS), and 74 nuclear-encoded mitochondrial ribosomal proteins (nucMTRP). As a control, we estimated the evolutionary
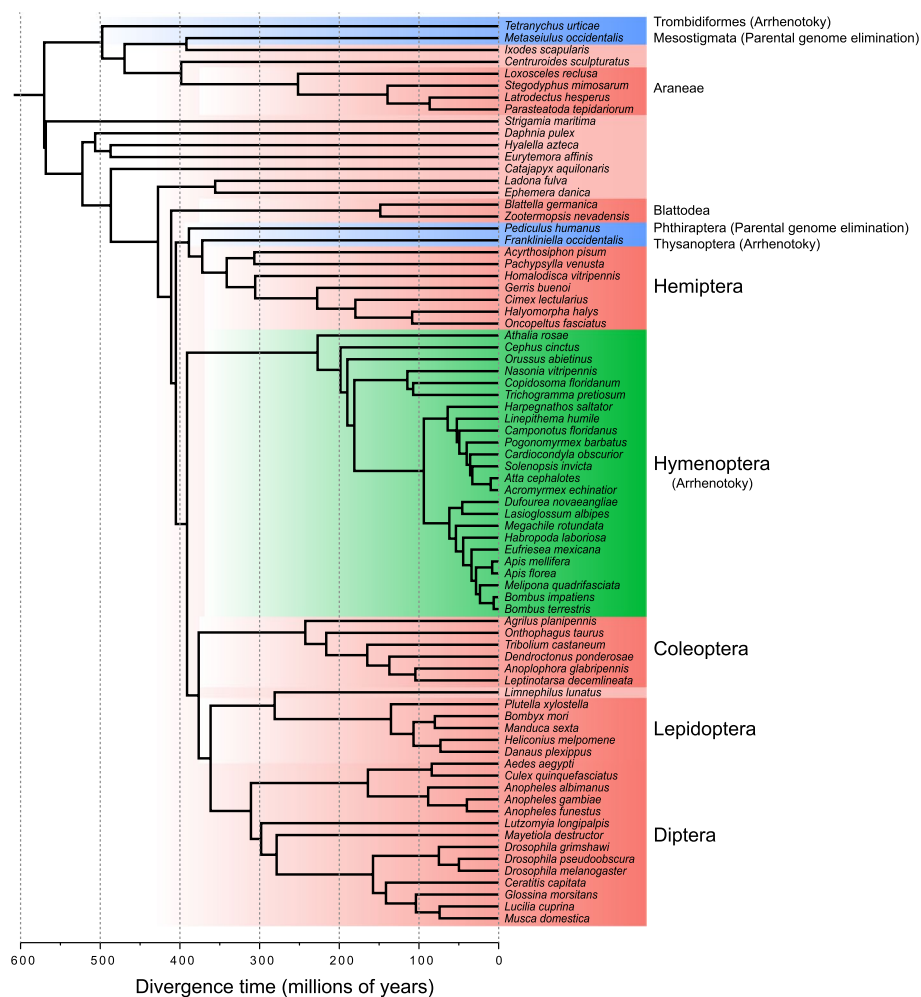
Fig. 1 Distribution of haplodiploidy in the phylogeny of 76 arthropod species. The tree is based on the time tree of Thomas et al. [32]. Hymenopterans are colored in green; non-Hymenoptera species are colored in blue; diploid species are colored in red. Different types of haplodiploidy are in parenthesis. Names of major arthropod orders with more than one species in the tree are next to the corresponding clade

Li *et al. BMC Biology*   (2024) 22:229

Page 4 of 14

rates of genes not related to mitochondrion in the genomes, including 77 nuclear-encoded cytosolic ribosomal proteins (nucCRP), 150 nuclear-encoded single-copy genes (nucControlSingle), and 5746 nuclear-encoded genes that were found in $\geq 80\%$ of the species (i.e., ≥ 61 species) and were not necessarily single copy (nucControl) (Table 1, Additional file 1: Table S2).

Haplodiploid species have fast-evolving mitochondrion-related genes

We first measured the evolutionary rate by concatenating the amino acid sequences of all the genes in the same gene category for each species. We then measured the number of amino acid substitutions per site per million years (AA/site/MY) based on terminal branch length and root-to-tip branch length. In general, terminal branch evolutionary rate and root-to-tip evolutionary rate showed a similar pattern (Fig. 2 and Additional file 2: Fig. S1), except that Hymenoptera nucControl genes showed significantly higher evolutionary rates than diploid species based on terminal branch length. Estimating evolutionary rates using terminal branches is often affected by the divergence time of the branches. Sequences from early diverging lineages could be saturated, leading to an underestimation of the evolutionary rate [40, 41]. At the same time, when lineages are under-sampled, polymorphisms could be treated as recent substitutions leading to the overestimation of evolutionary rates of recent terminal branches [42]. Therefore, more recent terminal branches could have higher evolutionary rates than early diverging branches due to estimation bias. As hymenopterans were well sampled in our study, they often had short divergence times, potentially leading to the observed pattern of fast evolution in their nucControl genes (Additional file 2: Fig. S1). Consistent with this hypothesis, we observed a negative correlation between divergence time and terminal branch evolutionary rate (Additional file 2: Fig. S2). To account for this bias, we estimated the evolutionary rate from the root to each tip in the tree, even though they are not independent of each other.

We first examined the evolutionary rates among different gene groups. Across all species, mtOXPHOS genes (median = 0.00176 AA/site/MY) had the highest evolutionary rates of all gene categories (Additional file 2: Fig. S3). Among nuclear genes, nucCRP (median = 0.00084 AA/site/MY) and nucControlSingle (median = 0.00118 AA/site/MY) were the most conservative gene groups with the lowest evolutionary rate. nucOXPHOS (median = 0.00120 AA/site/MY), nucMTRP (median = 0.00147 AA/site/MY), and nucControl genes showed an intermediate evolutionary rate (median = 0.00136 AA/site/MY). Based on the Kruskal–Wallis test and multiple comparison tests after Kruskal–Wallis, mtOXPHOS and nucMTRP genes evolved significantly faster than nucCRP genes across all species. mtOXPHOS genes in hymenopterans showed a significantly faster evolutionary rate than nucControlSingle and nucControl genes but the pattern was not found in diploid species.

Among different groups of arthropods, mtOXPHOS, nucOXPHOS, and nucMTRP genes evolved significantly faster in hymenopterans than in diploid species, while only mtOXPHOS genes evolved significantly faster in non-Hymenoptera haplodiploid species than diploid species (Fig. 1). As only four non-Hymenoptera species were included in our analysis, the small sample size limited our power to distinguish evolutionary rate differences for the genes. Hymenopterans had fast-evolving mtOXPHOS (median = 0.00282 AA/site/MY), nucOXPHOS (median = 0.00145 AA/site/MY), and nucMTRP (median = 0.00171 AA/site/MY) genes (Fig. 2). Other haplodiploids, including both PGE and arrhenotoky lineages, had fast-evolving mtOXPHOS (median = 0.00413 AA/site/MY) and nucMTRP (median = 0.00189 AA/site/MY) genes. Although the comparison was not significant, other haplodiploid species had higher evolutionary rates (median = 0.00136 AA/site/MY) than diploid species (median = 0.00112 AA/site/MY).

Table 1 Gene categories, functions, and the number of genes within each gene category in the present study

Gene category	Function	Number of orthologous groups
mtOXPHOS	Mitochondrial-encoded OXPHOS protein coding genes	13
nucOXPHOS	Nuclear-encoded OXPHOS protein coding genes	65
nucMTRP	Nuclear-encoded mitochondrial ribosomal protein genes	74
nucCRP	Nuclear-encoded cytosolic ribosomal protein genes	77
nucControlSingle	Single-copy nuclear-encoded genes found in all 76 arthropod species	150
nucControl	Nuclear-encoded genes found in at least 61 arthropod species	5746

Li *et al. BMC Biology*   (2024) 22:229

Page 5 of 14
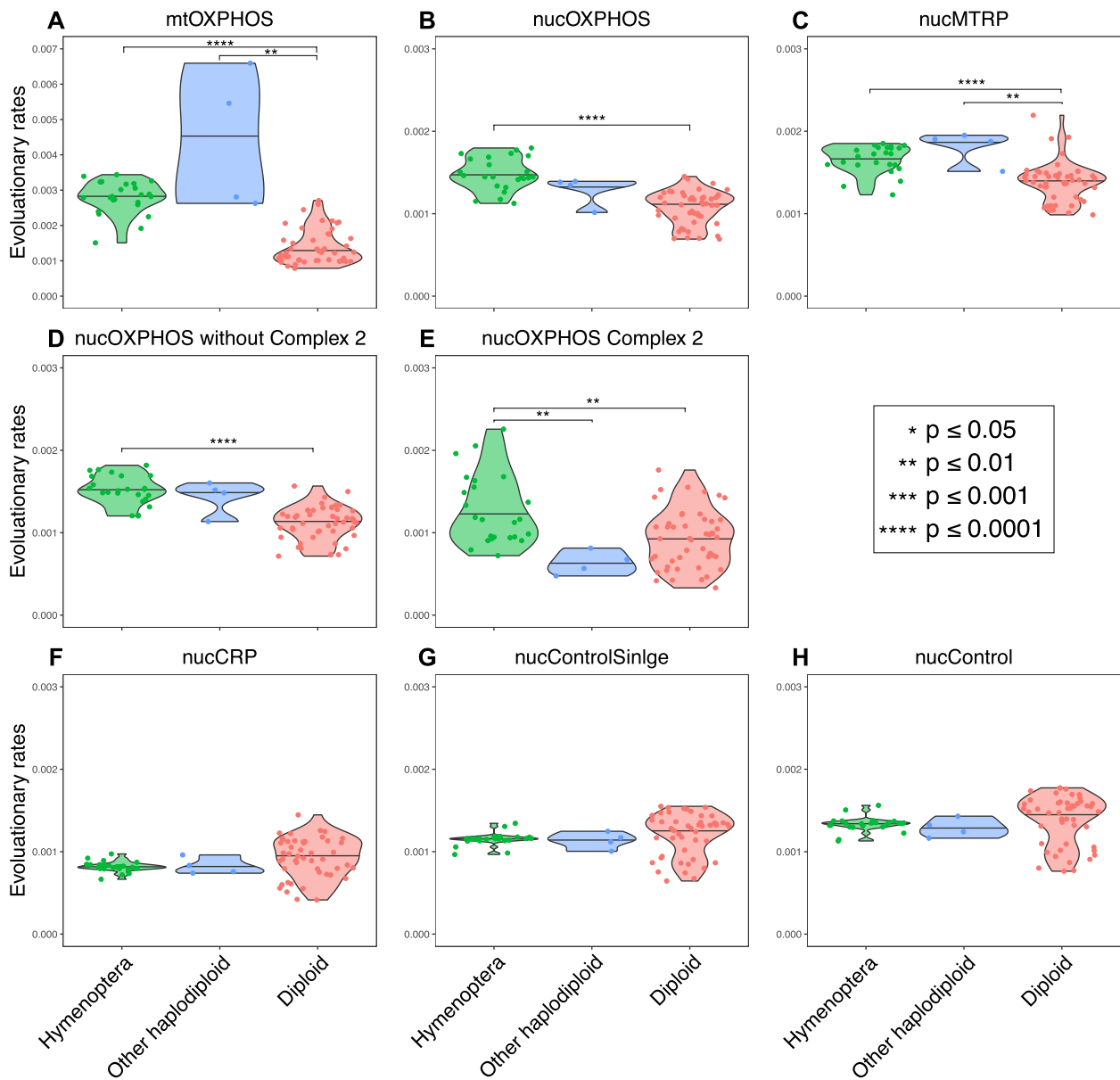
Fig. 2 Root-to-tip evolutionary rates (AA/site/MY) of different gene categories among arthropod groups. Asterisks indicate significant differences among hymenopterans (green), other non-Hymenoptera haplodiploid species (blue), or diploid species (red). Note that the y-axis of mtOXPHOS genes is different from nuclear genes due to the elevated evolutionary rates of mitochondrial genes

Coevolution between mitochondrial and nuclear genes

mtOXPHOS and their nuclear counterparts often coevolve to maintain the function of the OXPHOS pathway and mitochondrial replication, transcription, and translation [19, 43]. To investigate the relationships between mtOXPHOS genes and related nuclear genes, we estimated Spearman's rank correlation coefficients of evolutionary rates between different gene categories. Across all 76 arthropod species, the evolutionary rate of mtOXPHOS genes was significantly correlated with

the rates of nucOXPHOS and nucMTRP genes but not with other nuclear genes (Fig. 3 and Additional file 1: Table S3).

When only considering diploid or non-Hymenoptera haplodiploid species, the evolutionary rates of mtOX-PHOS and nucOXPHOS genes did not show a strong correlation (Fig. 3A). For example, there was variation among the non-Hymenoptera haplodiploid species: the two PGE haplodiploid species (*M. occidentalis* and *P. humanus*) showed an evolutionary rate of 0.0032 AA/
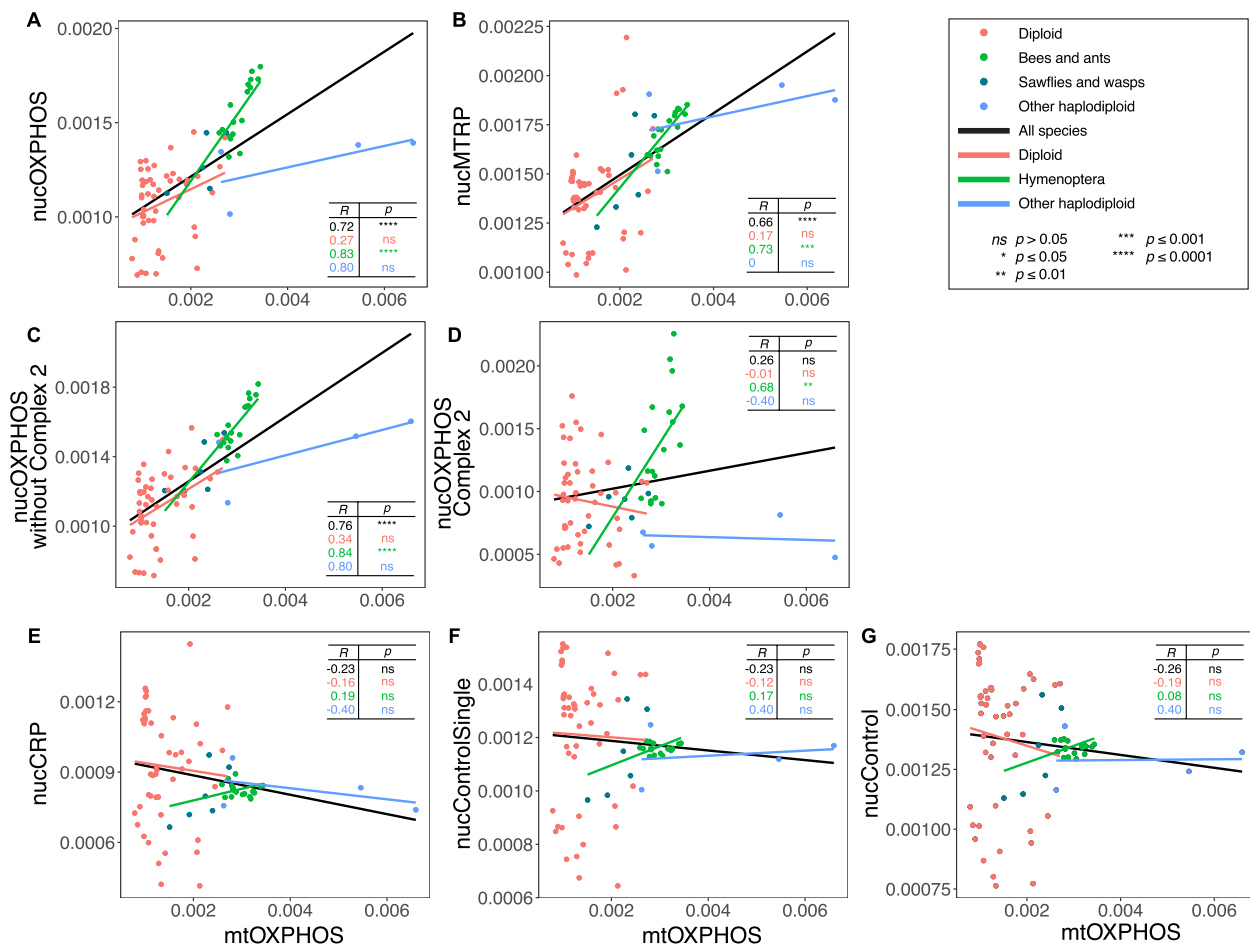
Fig. 3 Spearman's rank correlation of evolutionary rates (AA/site/MY) between mtOXPHOS genes and **A** nucOXPHOS, **B** nucMTRP, **E** nucCRP, **F** nucControlSingle, and **G** nucControl genes based on root-to-tip evolutionary rate. In addition, we also estimated the correlation of evolutionary rates between mtOXPHOS genes and **C** nucOXPHOS genes not from complex 2 and **D** genes only from complex 2. Spearman's correlation coefficient (*R*) and *p* values (*p*) are used to estimate the correlation between evolutionary rates of mtOXPHOS genes and evolutionary rates of nuclear-encoded gene categories based on all four arthropod groups (black), diploid species (red), hymenopterans (bees and ants: green, sawflies and wasps: dark green), and other non-Hymenoptera haplodiploid species (blue). Linear regression represents the general trend of the correlation

site/MY and 0.0070 AA/site/MY, and the two arrhenotoky species (*T. urticae* and *F. occidentalis*) had an evolutionary rate of 0.0074 AA/site/MY and 0.0029 AA/site/MY, respectively, while their nuclear genes had similar evolutionary rates (Fig. 3B). On the other hand, when only considering hymenopterans, there was a significant correlation, possibly because our study included a more comprehensive selection of species from Hymenoptera than for any other arthropod orders.

It has been hypothesized that nucOXPHOS genes that do not directly interact with mitochondrial genes tend to have less correlated evolutionary rates with mitochondrial genes [19, 44]. We found that mtOXPHOS genes and nucOXPHOS genes excluding OXPHOS complex 2 (Fig. 3C, Additional file 1: Table S3) have a

better correlation than that including OXPHOS complex 2 (Fig. 3A) in arthropods. Arthropod mtOXPHOS genes and nucOXPHOS complex 2 genes have no significant correlation (Fig. 3D). Only hymenopterans show a strong correlation between mtOXPHOS genes and nucOXPHOS complex 2 genes. This observed correlation in hymenopterans is likely due to the evolutionary rate variations within hymenopterans that bees and ants tend to have higher evolutionary rates than sawflies and wasps in their nucOXPHOS genes (Fig. 3A).

Bees and ants have elevated mitochondrion-related nuclear gene evolutionary rates

To investigate rate variation among hymenopterans, we assigned hymenopterans to two groups: eusocial

Hymenoptera (including bees [superfamily: Apoidea] and ants [superfamily: Formicoidea; family: Formicidae]) and basal lineages (including sawflies [superfamilies: Cephoidea and Tenthredinoidea] and wasps [superfamilies: Chalcidoidea and Orussoidea]). As sawflies and wasps have higher divergence times than bees and wasps, we used tip-to-root evolutionary rates for this analysis. We found that bees and ants had significantly higher evolutionary rates in mtOXPHOS and nucOXPHOS genes than sawflies and wasps (Fig. 4).

To confirm our findings based on root-to-tip evolutionary rates and to consider the unbalanced sampling of bees, ants, wasps, and sawflies, we also estimated the terminal branch evolutionary rate by subsampling within hymenopterans. For each subsampling, we sampled one hymenopteran without replacement with the rest of the arthropods to keep the divergence time consistent among subsamples. Subsampling results were consistent with root-to-tip estimations that bees and ants had significantly higher evolutionary rates in mtOXPHOS and nucOXPHOS genes than sawflies and wasps (Additional file 2: Fig. S4).

Gene family expansion and contraction of haplodiploid and diploid arthropods

Lastly, as complete genomes were used, we investigated gene family expansions and contractions related to haplodiploidy. Based on a feature selection analysis using the R package, Boruta [45], multiple mitochondrion-related genes had significantly different copy numbers among haplodiploid and diploid species. For example, 3 out of 35 genes had significantly different copy numbers between haplodiploid/diploid species (Additional file 1: Table S4). Between bees/ants and sawflies/wasps, 9 out of 26 genes with significantly different copy numbers were mitochondrion-related. There were two copies of the mitochondrial translation elongation factor gene (EOG8Z0DM0, FlyBase IDs mEFTu1 [FBpp0086790] and mEFTu2 [FBpp0088054]) in sawflies and wasps but only one copy in most of the bees and ants. Interestingly, different from mEFTu1, mEFTu2 is overexpressed in the testis of *Drosophila melanogaster* [46, 47]. Changes in the gene family could relate to mitochondrial functions in bees and ants.
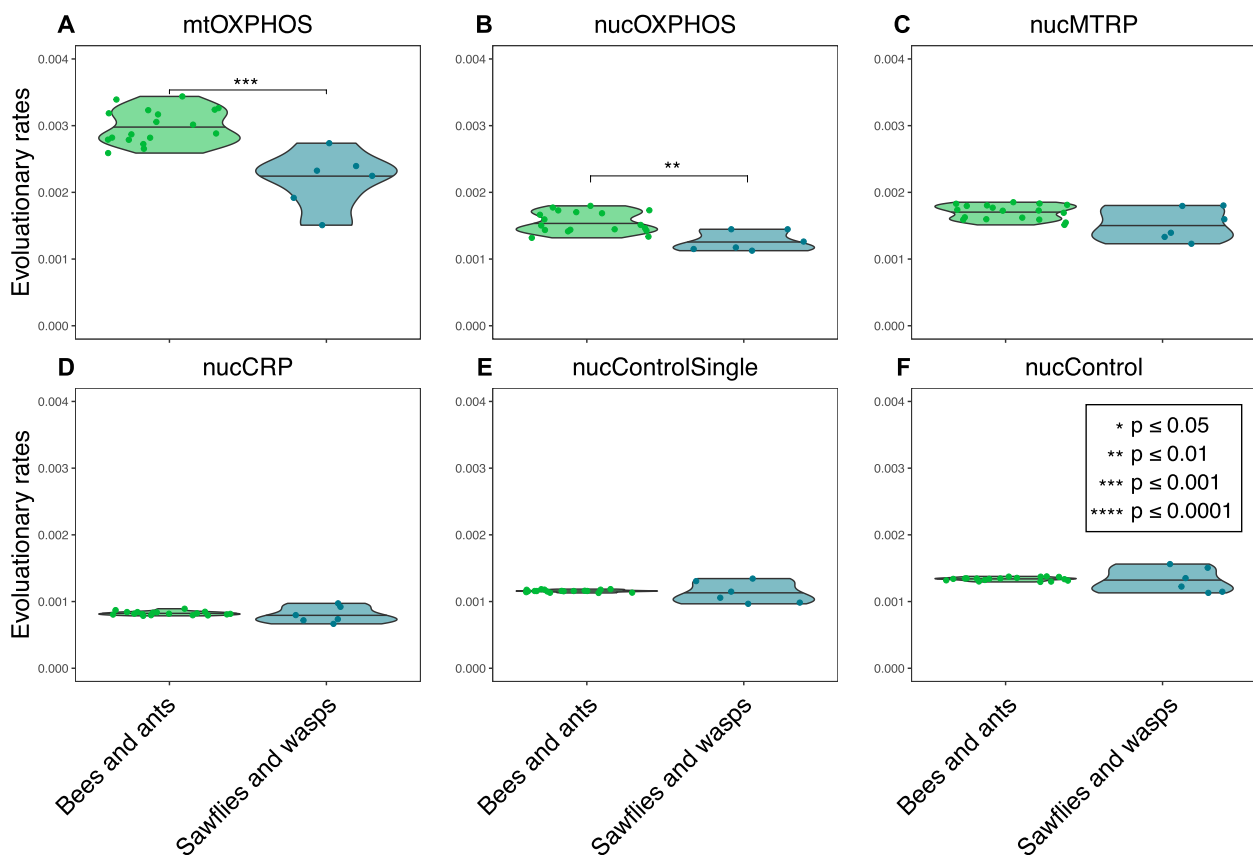
Fig. 4 Root-to-tip evolutionary rates of different gene categories among hymenopterans. Asterisks indicate significant differences between bees and ants (green) and sawflies and wasps (dark green)

Li *et al. BMC Biology*   (2024) 22:229

Page 8 of 14

Discussion

The coevolution of mitochondrial and nuclear genes is critical to mitochondrial functions. The mitochondrion is often uniparentally inherited and effectively haploid. As a result, the mitochondrion has a reduced effective population size relative to diploid nuclear genes and is prone to accumulation of deleterious mutations and, therefore, evolves faster than nuclear genes. The rate of molecular evolution in the mitochondrial genome varies dramatically among arthropods; however, the reasons for the accelerated mitochondrial rate in some arthropods remain unknown. Previous studies hypothesized that the elevated evolutionary rates in mitochondrial genomes of hymenopterans are due to their small effective population size from haplodiploidy, parasitism, or selection in the mitochondrion [28, 48]. To test these hypotheses, we investigated the evolutionary rate of mitochondrial and nuclear genes across five haplodiploid lineages, including three haplodiploid lineages in Insecta and two haplodiploid lineages in Acari (Fig. 1). These five haplodiploid lineages have arisen independently, providing a unique opportunity to comprehensively explore the evolution of mitochondrion-related genes.

Small population size is not the reason for fast-evolving mitochondrion-related genes in haplodiploid species

We showed that all haplodiploid species tested here, across 500 million years of evolution, have fast-evolving mitochondrial genes and mitochondrion-related nuclear genes (Fig. 2). Our observation suggests that haplodiploidy is linked to fast-evolving mitochondrial genes. Our results are consistent with other studies showing elevated evolutionary rates in mitochondrial genes of hymenopterans [19–21] and *P. humanus* [49].

Elevated evolutionary rates also correlate with fast mitochondrial gene rearrangement rates in arthropods [50, 51]. Previous studies have shown that most of the haplodiploid species included in this study have elevated mitochondrial rearrangement rates, including species in Hymenoptera [52], Thysanoptera (including *F. occidentalis*) [53, 54], and Phytoseiidae (including *M. occidentalis*) [55–57]. The most extreme mitochondrial rearrangement case is in *Pediculus* species, where instead of a single circular chromosome, the mitochondrial genome fractured into 20 circular mini-chromosomes [58]. Besides *Pediculus*, fragmented mitochondria and endosymbionts have arisen in multiple lineages, such as mini-chromosomes in *Silene* [59] and nematodes [60, 61], and fragmented genomes of endosymbiont *Hodgkinia* in cicadas [62], which all have elevated evolutionary rates. The fragmented genomes of mitochondria and symbionts are possibly nonadaptive for their hosts [63, 64].

For nuclear genes, our finding that mitochondrion-related genes (nucOXPHOS and nucMTRP) have fast evolutionary rates but that other nuclear genes (nucControlSingle and nucControl) do not (Figs. 2 and 3) rejects the hypothesis that the reduced effective population size caused by haplodiploidy or parasitism is the reason for the fast-evolving mitochondrial genomes. As both haplodiploidy and parasitism lead to a small population size, all nuclear genes should exhibit rapid evolution. Even though we included a larger set of nucControl genes than previous studies, we still did not find a genome-wide pattern of rapid evolution.

The nuclear compensation hypothesis could explain the fast-evolving mitochondrial-related genes in haplodiploid species

Alternatively, both nuclear compensation and selection on the OXPHOS pathway hypotheses could lead to the observed pattern. The nuclear compensation explanation is based on the observation that the mitochondrion often has fast evolutionary rates and accumulates more deleterious mutations than nuclear genes [5, 8], and thus, nuclear genes are hypothesized to be under positive selection to compensate for the deleterious changes in the mitochondrion [27]. In this scenario, mitochondrion-related nuclear genes are under positive selection but not the mitochondrial genes and unrelated nuclear genes. Non-Hymenoptera haplodiploid species show this pattern of rapid evolution in the nucOXPHOS complexes that consist of both mito-nuclear genes but not in the complexes with only nuclear-encoded genes (complex 2) (Fig. 2E). The observed evolutionary rate differences for OXPHOS genes are consistent with the nuclear compensation explanation. For haplodiploid species, haploid males express recessive mutations could lead to more efficient selection to purge deleterious and fix beneficial mutations, expediting the nuclear compensation process.

We were not able to reliably calculate *dN/dS* ratio (the ratio of the number of non-synonymous nucleotide substitutions per non-synonymous site to the number of synonymous nucleotide substitutions per synonymous site) for haplodiploid species, due to the deep divergence among the 76 arthropods. Our previous study, based on a comprehensive sampling of species in Hymenoptera, Coleoptera, Lepidoptera, and Diptera, demonstrated that 17 out of 23 tested nucOXPHOS genes of hymenopterans had an elevated *dN/dS* ratio compared to other holometabolous species, but only 2 out of 13 mitochondrial genes of hymenopterans had an elevated *dN/dS* ratio compared to other holometabolous species [21]. As nucOXPHOS genes under positive selection are important genes for mitochondrial functions, it is likely that the elevated *dN/dS* ratio is due to positive selection to

Li *et al. BMC Biology*   (2024) 22:229

Page 9 of 14

compensate for the deleterious mutations in fast-evolving mitochondrial genes. In another arthropod example, positive selection was detected in nucOXPHOS and nucMTRP but not mitochondrial genes for copepod species with fast-evolving mitochondrial genes [65, 66]. Therefore, it is possible that the fast-evolving nucOXPHOS genes found in haplodiploid species were due to compensation for the fast-evolving mtOXPHOS genes.

The positive selection hypothesis could explain the fast-evolving mitochondrion-related genes in bees and ants

The positive selection on OXPHOS genes hypothesis claims that both mitochondrial and nuclear genes are under positive selection. Although most studies [21, 43, 67] have shown that mitochondrial genes often have a low *dN/dS* ratio supporting the nuclear compensation hypothesis, Piccinini et al. [44] found that mtOXPHOS genes of bivalves had similar *dN/dS* ratios as nucOXPHOS genes, suggesting positive selection on the whole OXPHOS pathway, possibility due to adaptation to environmental stress.

Bees and ants experience different evolutionary patterns than other haplodiploid species. None of the haplodiploid species except bees and ants showed a correlation between mtOXPHOS genes and nucOXPHOS complex 2 genes (Fig. 3D), although genes in complex 2 do not directly interact with mtOXPHOS genes. It is possible that the whole OXPHOS pathway of bees and ants is under positive selection due to their high metabolic demand. For example, bees and ants often maintain their nest temperature using metabolic heat [68, 69]. Bees and ants are also central place foragers. A recent study has found that the OXPHOS genes of pollinating fig wasps (superfamily: Chalcidoidea) evolved significantly faster than those genes of non-pollinating fig wasps [70]. Five of the OXPHOS genes in pollinators showed signals of positive selection possibly due to energetic demand during pollination [70].

In addition, we propose another explanation for the elevated mitochondrial gene evolutionary rate that haplodiploid species could accumulate deleterious mutations on the mitochondrial recombination, replication, and repair genes due to small population size. The disruption of mitochondrial recombination, replication, or repair genes (RRR genes) could cause increased substitution rates in mitochondrial genomes. Such cases have been found in *D. melanogaster* [71], *Arabidopsis thaliana* [72], and African killifishes [73]. The observed elevated evolutionary rates on nucOXPHOS complex 2 of bees and ants (Figs. 3 and 4) also support this explanation, as eusocial hymenopterans have even smaller population sizes than sawflies and wasps [25, 74] potentially leading to a higher

chance of deleterious mutations occurring on mitochondrion-related genes. Recent studies [75, 76] also found that low copy numbers of mitochondrial DNA could lead to high mitochondrial evolutionary rates due to the less efficient homologous recombination for mitochondrial DNA repair. The nuclear compensation, positive selection, and less efficient mitochondrial repair system are not exclusive to each other and could work together affecting the mito-nuclear evolution of bees and ants.

Gene family expansions and contractions in different haplodiploid groups

Besides evolutionary rates, we also investigated gene gains and losses in haplodiploid species. Overall, arthropods have conserved copy number for mitochondrion-related genes due to the functional importance of mitochondrion. Only a few mitochondrion-related gene families were found to have significant expansions or contractions in haplodiploid species. We also found several gene losses with mitochondrion-related GO terms, which could affect the function of mitochondrial genes. Many bees and ants have lost one of the mitochondrial translation elongation factor gene copies, *mEFTu2*. The gains and losses of mitochondrion-related genes could affect mitochondrion evolution. For example, the loss of the mitochondrial single-stranded binding protein has been associated with the mitochondrial mini-chromosomes in the *P. humanus* genome [77].

Biased sampling efforts of genomes could affect the estimation of evolutionary rates

One caveat to our study is the biased genome sequencing efforts devoted to hymenopterans due to their agricultural importance [78, 79]. Of the 76 studied arthropod species, 26 were hymenopterans, and only four species represented four non-Hymenoptera haplodiploid orders. This biased sampling effort limited our interpretation of the global effects of haplodiploidy on evolutionary rates. Although we performed subsampling and simulation to take this sampling bias into account, more genomes from other haplodiploid orders could vastly improve our understanding of haplodiploidy and its effects on genome evolution in the future.

Conclusions
We studied the evolutionary and co-evolutionary dynamics of mitochondrial and nuclear genes across major haplodiploid arthropod lineages. We found elevated mitochondrion-related gene evolution in all five haplodiploid lineages. Our findings suggest that the small population size of haplodiploid species is not likely to be the reason for the fast-evolving mitochondrial genes. But investing different gene groups,

we propose three alternative explanations, including nuclear compensation, positive selection on the OXPHOS pathway, and deleterious substitutions on the mitochondrial RRR genes. We also showed several gene family gains and losses in haplodiploid species, suggesting the unique evolution history of their mitochondrion-related genes.

Methods

Taxon sampling and sequence acquisition

The genome sequences of 28 arthropod species from the Arthropod 5000 Genomes Initiative (i5K) project [30] and 48 previously sequenced arthropods spanning 21 arthropod orders were used in this study (Additional file 1: Table S1). Within the 76 species, 28 species were haplodiploid, including 24 species from Hymenoptera and 4 species from Trombidiformes, Mesostigmata, Phthiraptera, and Thysanoptera (Fig. 1).

A set of 38,195 ortholog groups was obtained from Thomas et al. [32] based on the orthologous gene prediction of OrthoDB8 [80]. Two categories of nuclear genes related to mitochondrial function were used in our study: (i) nuclear-encoded OXPHOS genes (nucOXPHOS) and (ii) nuclear-encoded mitochondrial ribosomal protein genes (nucMTRP) [19, 81]. To estimate the background evolutionary rate of genes in the nuclear genome, three gene categories with no relationship to mitochondrial genes were used, including (i) nuclear-encoded cytosolic rRNA protein genes (nucCRP), (ii) 150 single copy nuclear-encoded genes (nucControlSingle) (Thomas et al. [32]), and (iii) all the nuclear-encoded genes (nucControl).

Multiple approaches were used to extract groups of orthologous genes in the different categories described above from the complete set of orthologs. For nucOXPHOS genes, the OXPHOS gene annotations of *Drosophila melanogaster* [82, 83] were used to search the *D. melanogaster* gene identifiers (IDs) in each ortholog group (Additional file 1: Table S2). For nucMTRP genes, 75 genes under the FlyBase ID "FBgg0000059" were used to search the *Drosophila melanogaster* gene IDs in ortholog groups. For nucCRP genes, 93 genes under the FlyBase ID "FBgg0000141" were used to match the genes in ortholog groups [19, 81]. Of the 38,195 ortholog groups, 150 ortholog groups had only 1 copy in any of the species. These 150 ortholog groups were used as nucControlSingle genes. As RAxML [84] treated gaps as undetermined base pairs, missing genes could inflate the evolutionary rate estimation. Therefore, 5746 ortholog groups that could be found in ≥ 80% of the species (at least 61 species) were used as background genes (nucControl gene category).

Assembly and annotation of mitochondrial genomes

Reference mitochondrial genomes of 32 arthropod species are available on NCBI. Mitochondrial genes of 12 species have already been annotated from transcriptome data in our previous study (Li et al. [21]). For the rest of the species, we first assembled the mitochondrial genome from the available genome sequence data using MITObim version 1.9 [85]. To minimize the computational requirements, we extracted between 12,500,000 and 100,000,000 reads for mitochondrial genome assembly (Additional file 1: Table S1). If a mitochondrial reference genome from the same genus was available, "-quick" option of MITObim was used where mitochondrial genomes were assembled based on the reads with a certain kmer overlap to the mitochondrial reference. If a mitochondrial genome from the same genus was not available, the COI barcode of the same species was used as a seed for assembly. Assemblies from MITObim were annotated online using MITOS version 2 [86] based on NCBI RefSeq 81 and the invertebrate genetic code. We were able to assemble and annotate 23 arthropod species using this approach. For the remaining 9 species without mitochondrial genomes and their mitochondrial genomes could not be assembled from the genomic data, a reference mitochondrial genome from a closely related species was used in the analysis (Additional file 1: Table S1). The protein-coding genes of the mitochondrial genomes were used for subsequent mitochondrial OXPHOS gene (mtOXPHOS) evolutionary rate analysis.

Estimating evolutionary rates

Due to the deep divergence, we used amino acid sequences of orthologous groups to estimate evolutionary rates. In cases where a species had multiple paralogs within an orthologous group, a random paralog was selected for subsequent analysis. The amino acid sequences of each orthologous group were first aligned using the L-INS-i algorithm in MAFFT version 7.475 [87]. GBlocks version 0.91b [88] was then used to remove the low-quality regions of the alignment with "-t = p -b4 = 5 -b5 = a -e = -gb" meaning the input aligned sequences were amino acid alignment, the minimum length of a block was 6 amino acids, and gaps were allowed. Lastly, genes in the same gene category (i.e., mtOXPHOS, nucOXPHOS, nucMTRP, nucCRP, nucControlSingle, and nucControl) were concatenated.

To estimate evolutionary rates, branch lengths of gene trees were estimated using the concatenated amino acid sequence alignments with RAxML version 8.2.3 and the same fixed topology from Thomas et al. [32]. RAxML version 8.2.3 [84] was used with the PROTGAMMAAUTO model option, which automatically selected the

Li *et al. BMC Biology*   (2024) 22:229

Page 11 of 14

best-fitting amino acid substitution model based on the log-likelihood value and approximated across-site rate heterogeneity with a gamma distribution. In RAxML, the "-t" option was used to estimate branch lengths based on the given topology. We estimated both the terminal branch length and the root-to-tip branch length, which included the terminal branch length and all the internal nodes leading to the common ancestor of arthropods, using phytools version 0.6–44 [89]. We calculated the evolutionary rate as the branch length estimation divided by the divergence time (MYA). The divergence time of each species was retrieved from the time tree of Thomas et al. [32].

To evaluate whether the gene family size of control genes (5746 genes) biased the estimation of evolutionary rate, 1000 random subsamplings were performed based on the gene number and gene family size of nucOX-PHOS and nucMTRP genes. The evolutionary rate was estimated in the same way above. Spearman's rank correlation coefficient (cor.test function in R) was used to estimate the differences between the evolutionary rate of subsamplings and 5746 control genes.

Statistical analysis of gene evolutionary rates

R version 3.4.4 [90] was used to perform statistical tests. Kruskal–Wallis test (kruskal.test) and multiple comparison tests after Kruskal–Wallis (kruskalmc) from pgirmess package version 2.0.3 [91] were used to test the significance of gene evolutionary rate differences between Hymenoptera species, other non-Hymenoptera haplodiploid species, and diploid species. Spearman's rank correlation coefficient (cor.test function in R) was used to calculate the correlation of rates between gene groups. The PCA package in R was used to find the principal components that explain the pattern of evolutionary rates within arthropod species. ggplot2 version 2.2.1.9 [92] and ggtree version 1.10.5 [93] were used to visualize data and phylogenetic trees.

Mitochondrion-related gene family expansion and contraction in arthropods

Gene family expansions and contractions were estimated based on the number of genes in each ortholog group and the time tree using Dupliphy version 1.0 [94]. To explore the unique gene family expansions and contractions in haplodiploid species and Hymenoptera species, a random forest algorithm was used to select significant features using the R package Boruta version 5.3.0 [45]. Three different sets of species were used for feature selection analysis: (a) haplodiploid species versus diploid species, (b) hymenopterans versus diploid species, and (c) haplodiploid species except hymenopterans versus diploid species. Gene copy numbers were used as input attributes to separate the two groups. For gene families that separate different groups of insects based on the Boruta approach, we used a rough fix of tentative attributes (TentativeRoughFix) to judge whether these gene families were significantly important.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-024-02027-4.

Additional file 1: Tables S1–S4. Table S1 Arthropod species, genomes, and mitochondrial genomes used in the study. Table S2 Nuclear-encoded OXPHOS, mitochondrial ribosomal protein genes, and cytosolic ribosomal proteins used in the study. Table S3 Spearman's rank correlation analysis of root-to-tip evolutionary rates between mitochondrial genes and other gene categories. Table S4 Genes that separate different arthropod groups

Additional file 2: Figures S1–S4. Fig. S1 Terminal branch evolutionary rates of different gene categories among arthropod groups. Asterisks indicate significant differences among hymenopterans, other non-Hymenoptera haplodiploid species, or diploid species. Fig. S2 Correlation between divergence time and terminal branch evolutionary rates based on Spearman's rank correlation. Spearman's rank correlation coefficient is estimated based on all species from hymenopterans, other non-Hymenoptera haplodiploid species, or diploid species. Fig. S3 Root-to-tip evolutionary rate of different gene categories among all species, hymenopterans, other haplodiploid species, and diploid species. Significance was denoted based on letters. Letters on top of the violin plots denote significant differences based on the Kruskal-Wallis test and multiple comparison tests after Kruskal-Wallis. Letters were ordered alphabetically based on the median of the distributions from the highest to the lowest. Fig. S4 Terminal branch evolutionary rate of different gene categories among hymenopterans with subsampling strategy. In each subsampling, one hymenopteran with the rest of the arthropods to keep the divergence time consistent among subsamples

---

Authors' contributions

 YL and MEP designed and performed research. GWCT, SR, and RMW contributed data. GWCT, SR, RMW, and XZ advised on data analysis. All authors contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

The amino acid sequences of all the gene categories and scripts used in the current study are available at https://github.com/lyy005/Mitochondrial_nuclear_gene_evolution_in_arthropods (last accessed Sept 26, 2024). The information of the 76 arthropod species used in this study is available at https://arthrofam.org. Complete mitochondrial genomes assembled in this study are under the NCBI accession numbers: MN842135–MN842145.

Declarations

Ethics approval and consent to participate
Not applicable.

Li *et al. BMC Biology*   (2024) 22:229

Page 12 of 14

Author details
[1]State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Key Laboratory of Biotechnology in Plant Protection of Ministry of Agriculture and Rural Affairs, Key Laboratory of Green Plant Protection of Zhejiang Province, Institute of Plant Virology, Ningbo University, Ningbo 315211, China. [2]Department of Biology, Indiana University, Bloomington, IN, USA. [3]Department of Computer Science, Indiana University, Bloomington, IN, USA. [4]Current Address: Informatics Group, Harvard University, Cambridge, MA, USA. [5]Human Genome Sequencing Center, Department of Human and Molecular Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX, USA. [6]Department of Ecology & Evolution and Swiss Institute of Bioinformatics, University of Lausanne, 1015 Lausanne, Switzerland. [7]Department of Entomology, College of Plant Protection, China Agricultural University, Beijing 100193, China. [8]Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA. [9]Environmental Change Initiative, Notre Dame, IN, USA.

References
1. Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 2004;5:123–35.
2. Bennett GM, Moran NA. Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. Proc Natl Acad Sci USA. 2015;112:10169–76.
3. Sloan DB, Warren JM, Williams AM, Wu Z, Abdel-Ghany SE, Chicco AJ, et al. Cytonuclear integration and co-evolution. Nat Rev Genet. 2018;19:635–48.
4. Moran NA. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc Natl Acad Sci U S A. 1996;93:2873–8.
5. Lynch M, Blanchard JL. Deleterious mutation accumulation in organelle genomes. Genetica. 1998;102–103:29–39.
6. Barr CM, Neiman M, Taylor DR. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. New Phytol. 2005;168:39–50.
7. Neiman M, Taylor DR. The causes of mutation accumulation in mitochondrial genomes. Proc Biol Sci. 2009;276:1201–9.
8. Brown WM, George M, Wilson AC. Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A. 1979;76:1967–71.
9. Vawter L, Brown WM. Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. Science. 1986;234:194–6.
10. Lynch M. Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. Mol Biol Evol. 1996;13:209–20.
11. Havird JC, Forsythe ES, Williams AM, Werren JH, Dowling DK, Sloan DB. Selfish mitonuclear conflict. Curr Biol. 2019;29:R496–511.
12. Tranah GJ. Mitochondrial–nuclear epistasis: implications for human aging and longevity. Ageing Res Rev. 2011;10:238–52.
13. Burton RS, Barreto FS. A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? Mol Ecol. 2012;21:4942–57.
14. Eyre-Walker A. Mitochondrial replacement therapy: are mito-nuclear interactions likely to be a problem? Genetics. 2017;205:1365–72.
15. Dobler R, Dowling DK, Morrow EH, Reinhardt K. A systematic review and meta-analysis reveals pervasive effects of germline mitochondrial replacement on components of health. Hum Reprod Update. 2018;24:519–34.
16. Moran BM, Payne CY, Powell DL, Iverson ENK, Donny AE, Banerjee SM, et al. A lethal mitonuclear incompatibility in complex I of natural hybrids. Nature. 2024;626:119–27.
17. Barreto FS, Burton RS. Elevated oxidative damage is correlated with reduced fitness in interpopulation hybrids of a marine copepod. Proc Biol Sci. 2013;280:20131521.
18. Mitterboeck TF, Liu S, Adamowicz SJ, Fu J, Zhang R, Song W, et al. Positive and relaxed selection associated with flight evolution and loss in insect transcriptomes. Gigascience. 2017;6:1–14.
19. Yan Z, Ye G, Werren JH. Evolutionary rate correlation between mitochondrial-encoded and mitochondria-associated nuclear-encoded proteins in insects. Mol Biol Evol. 2019;36:1022–36.
20. Kaltenpoth M, Showers Corneli P, Dunn DM, Weiss RB, Strohm E, Seger J. Accelerated evolution of mitochondrial but not nuclear genomes of Hymenoptera: new evidence from crabronid wasps. PLoS ONE. 2012;7: e32826.
21. Li Y, Zhang R, Liu S, Donath A, Peters RS, Ware J, et al. The molecular evolutionary dynamics of oxidative phosphorylation (OXPHOS) genes in Hymenoptera. BMC Evol Biol. 2017;17:269.
22. Normark BB. The evolution of alternative genetic systems in insects. Annu Rev Entomol. 2003;48:397–423.
23. Dowton M, Austin AD. Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the Hymenoptera. J Mol Evol. 1995;41:958–65.
24. Bromham L, Leys R. Sociality and the rate of molecular evolution. Mol Biol Evol. 2005;22:1393–402.
25. Romiguier J, Lourenco J, Gayral P, Faivre N, Weinert LA, Ravel S, et al. Population genomics of eusocial insects: the costs of a vertebrate-like effective population size. J Evol Biol. 2014;27:593–603.
26. Consuegra S, John E, Verspoor E, de Leaniz CG. Patterns of natural selection acting on the mitochondrial genome of a locally adapted fish species. Genet Sel Evol. 2015;47:58.
27. Rand DM, Haney RA, Fry AJ. Cytonuclear coevolution: the genomics of cooperation. Trends Ecol Evol. 2004;19:645–53.
28. Oliveira DCSG, Raychoudhury R, Lavrov DV, Werren JH. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp Nasonia (Hymenoptera: Pteromalidae). Mol Biol Evol. 2008;25:2167–80.
29. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, et al. Phylogenomics resolves the timing and pattern of insect evolution. Science. 2014;346:763–7.
30. i5K Consortium. The i5K Initiative: advancing arthropod genomics for knowledge, human health, agriculture, and the environment. J Hered. 2013;104:595–600.
31. Poelchau M, Childers C, Moore G, Tsavatapalli V, Evans J, Lee C-Y, et al. The i5k Workspace@NAL—enabling genomic data access, visualization and curation of arthropod genomes. Nucleic Acids Res. 2015;43(Database issue):D714–719.
32. Thomas GWC, Dohmen E, Hughes DST, Murali SC, Poelchau M, Glastad K, et al. Gene content evolution in the arthropods. Genome Biol. 2020;21:15.
33. de la Filia AG, Bain SA, Ross L. Haplodiploidy and the reproductive ecology of arthropods. Curr Opin Insect Sci. 2015;9:36–43.
34. Blackmon H, Ross L, Bachtrog D. Sex determination, sex chromosomes, and karyotype evolution in insects. J Hered. 2017;108:78–93.
35. Hoy MA. Parahaploidy of the "arrhenotokous" predator, Metaseiulus occidentalis (Acarina: Phytoseiidae) demonstrated by x-irradiation of males. Entomol Exp Appl. 1979;26:97–104.
36. Cruickshank RH, Thomas RH. Evolution of haplodiploidy in dermanyssine mites (Acari: Mesostigmata). Evolution. 1999;53:1796–803.
37. Hoy MA, Waterhouse RM, Wu K, Estep AS, Ioannidis P, Palmer WJ, et al. Genome sequencing of the phytoseiid predatory mite Metaseiulus occidentalis reveals completely atomized hox genes and superdynamic intron evolution. Genome Biol Evol. 2016;8:1762–75.
38. de la Filia A, Andrewes S, Clark J, Ross L. The unusual reproductive system of head and body lice (Pediculus humanus). Med Vet Entomol. 2018;32:226–34.
39. Macke E, Magalhães S, Khan HD-T, Luciano A, Frantz A, Facon B, et al. Sex allocation in haplodiploids is mediated by egg size: evidence in the spider mite Tetranychus urticae Koch. Proc Biol Sci. 2011;278:1054–63.
40. Ho SYW, Phillips MJ, Cooper A, Drummond AJ. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Mol Biol Evol. 2005;22:1561–8.
41. Ho SYW, Lanfear R, Bromham L, Phillips MJ, Soubrier J, Rodrigo AG, et al. Time-dependent rates of molecular evolution. Mol Ecol. 2011;20:3087–101.
42. Emerson BC, Hickerson MJ. Lack of support for the time-dependent molecular evolution hypothesis. Mol Ecol. 2015;24:702–9.

Li *et al. BMC Biology*   (2024) 22:229

Page 13 of 14

43. Havird JC, Sloan DB. The roles of mutation, selection, and expression in determining relative rates of evolution in mitochondrial versus nuclear genomes. Mol Biol Evol. 2016;33:3042–53.

44. Piccinini G, Iannello M, Puccio G, Plazzi F, Havird JC, Ghiselli F. Mitonuclear coevolution, but not nuclear compensation, drives evolution of OXPHOS complexes in bivalves. Mol Biol Evol. 2021;38:2597–614.

45. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Softw. 2010;36:1–13.

46. Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, et al. Diversity and dynamics of the Drosophila transcriptome. Nature. 2014;512:393–9.

47. Marygold SJ, Attrill H, Lasko P. The translation factors of Drosophila melanogaster. Fly (Austin). 2017;11:65–74.

48. Dowton M, Cameron SL, Austin AD, Whiting MF. Phylogenetic approaches for the analysis of mitochondrial genome sequence data in the Hymenoptera–a lineage with both rapidly and slowly evolving mitochondrial genomes. Mol Phylogenet Evol. 2009;52:512–9.

49. Ashfaq M, Prosser S, Nasir S, Masood M, Ratnasingham S, Hebert PDN. High diversity and rapid diversification in the head louse, Pediculus humanus (Pediculidae: Phthiraptera). Sci Rep. 2015;5: 14188.

50. Shao R, Dowton M, Murrell A, Barker SC. Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. Mol Biol Evol. 2003;20:1612–9.

51. Xu W, Jameson D, Tang B, Higgs PG. The relationship between the rate of molecular evolution and the rate of genome rearrangement in animal mitochondrial genomes. J Mol Evol. 2006;63:375–92.

52. Dowton M, Austin AD. Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. Mol Biol Evol. 1999;16:298–309.

53. Yan D, Tang Y, Xue X, Wang M, Liu F, Fan J. The complete mitochondrial genome sequence of the western flower thrips Frankliniella occidentalis (Thysanoptera: Thripidae) contains triplicate putative control regions. Gene. 2012;506:117–24.

54. Tyagi K, Chakraborty R, Cameron SL, Sweet AD, Chandra K, Kumar V. Rearrangement and evolution of mitochondrial genomes in Thysanoptera (Insecta). Sci Rep. 2020;10:695.

55. Jeyaprakash A, Hoy MA. The mitochondrial genome of the predatory mite Metaseiulus occidentalis (Arthropoda: Chelicerata: Acari: Phytoseiidae) is unexpectedly large and contains several novel features. Gene. 2007;391:264–74.

56. Dermauw W, Vanholme B, Tirry L, Van Leeuwen T. Mitochondrial genome analysis of the predatory mite Phytoseiulus persimilis and a revisit of the Metaseiulus occidentalis mitochondrial genome. Genome. 2010;53:285–301.

57. Zhang B, Havird JC, Wang E, Lv J, Xu X. Massive gene rearrangement in mitogenomes of phytoseiid mites. Int J Biol Macromol. 2021;186:33–9.

58. Shao R, Zhu X-Q, Barker SC, Herd K. Evolution of extensively fragmented mitochondrial genomes in the lice of humans. Genome Biol Evol. 2012;4:1088–101.

59. Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, et al. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. 2012;10: e1001241.

60. Armstrong MR, Blok VC, Phillips MS. A multipartite mitochondrial genome in the potato cyst nematode Globodera pallida. Genetics. 2000;154:181–92.

61. Kern EMA, Kim T, Park J-K. The mitochondrial genome in nematode phylogenetics. Front Ecol Evol. 2020;8: 250.

62. Campbell MA, Łukasik P, Simon C, McCutcheon JP. Idiosyncratic genome degradation in a bacterial endosymbiont of periodical cicadas. Curr Biol. 2017;27:3568–3575.e3.

63. Campbell MA, Van Leuven JT, Meister RC, Carey KM, Simon C, McCutcheon JP. Genome expansion via lineage splitting and genome reduction in the cicada endosymbiont Hodgkinia. Proc Natl Acad Sci U S A. 2015;112:10192–9.

64. Wu Z, Cuthbert JM, Taylor DR, Sloan DB. The massive mitochondrial genome of the angiosperm Silene noctiflora is evolving by gain or loss of entire chromosomes. Proc Natl Acad Sci U S A. 2015;112:10185–91.

65. Barreto FS, Burton RS. Evidence for compensatory evolution of ribosomal proteins in response to rapid divergence of mitochondrial rRNA. Mol Biol Evol. 2013;30:310–4.

66. Barreto FS, Watson ET, Lima TG, Willett CS, Edmands S, Li W, et al. Genomic signatures of mitonuclear coevolution across populations of Tigriopus californicus. Nat Ecol Evol. 2018;2:1250–7.

67. Nabholz B, Ellegren H, Wolf JBW. High levels of gene expression explain the strong evolutionary constraint of mitochondrial protein-coding genes. Mol Biol Evol. 2013;30:272–84.

68. Frouz J. The effect of nest moisture on daily temperature regime in the nests of Formica polyctena wood ants. Insectes Soc. 2000;47:229–35.

69. Jones JC, Myerscough MR, Graham S, Oldroyd BP. Honey bee nest thermoregulation: diversity promotes stability. Science. 2004;305:402–4.

70. Zhou Y, Huang D, Xin Z, Xiao J. Evolution of oxidative phosphorylation (OXPHOS) genes reflecting the evolutionary and life histories of fig wasps (Hymenoptera, Chalcidoidea). Genes (Basel). 2020;11: 1353.

71. Samstag CL, Hoekstra JG, Huang C-H, Chaisson MJ, Youle RJ, Kennedy SR, et al. Deleterious mitochondrial DNA point mutations are overrepresented in Drosophila expressing a proofreading-defective DNA polymerase γ. PLoS Genet. 2018;14: e1007805.

72. Wu Z, Waneka G, Broz AK, King CR, Sloan DB. MSH1 is required for maintenance of the low mutation rates in plant mitochondrial and plastid genomes. Proc Natl Acad Sci U S A. 2020;117:16448–55.

73. Cui R, Medeiros T, Willemsen D, Iasi LNM, Collier GE, Graef M, et al. Relaxed selection limits lifespan by increasing mutation load. Cell. 2019;178:385–399.e20.

74. Graur D. Gene Diversity in Hymenoptera. Evolution. 1985;39:190–9.

75. Davison A, Chowdhury M, Johansen M, Wellcome Sanger Institute Tree of Life programme, Uliano-Silva M, Blaxter M. High heteroplasmy is associated with low mitochondrial copy number and selection against non-synonymous mutations in the snail Cepaea nemoralis. BMC Genomics. 2024;25:596.

76. Zwonitzer KD, Tressel LG, Wu Z, Kan S, Broz AK, Mower JP, et al. Genome copy number predicts extreme evolutionary rate variation in plant mitochondrial DNA. Proc Natl Acad Sci U S A. 2024;121: e2317240121.

77. Kirkness EF, Haas BJ, Sun W, Braig HR, Perotti MA, Clark JM, et al. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. Proc Natl Acad Sci U S A. 2010;107:12168–73.

78. Branstetter MG, Childers AK, Cox-Foster D, Hopper KR, Kapheim KM, Toth AL, et al. Genomes of the Hymenoptera. Curr Opin Insect Sci. 2018;25:65–75.

79. Hotaling S, Sproul JS, Heckenhauer J, Powell A, Larracuente AM, Pauls SU, et al. Long reads are revolutionizing 20 years of insect genome sequencing. Genome Biol Evol. 2021;13: evab138.

80. Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simão FA, Pozdnyakov IA, et al. OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. Nucleic Acids Res. 2015;43 Database issue:D250–256.

81. Havird JC, Trapp P, Miller CM, Bazos I, Sloan DB. Causes and consequences of rapidly evolving mtDNA in a plant lineage. Genome Biol Evol. 2017;9:323–36.

82. Tripoli G, D'Elia D, Barsanti P, Caggese C. Comparison of the oxidative phosphorylation (OXPHOS) nuclear genes in the genomes of Drosophila melanogaster, Drosophila pseudoobscura and Anopheles gambiae. Genome Biol. 2005;6:R11.

83. Porcelli D, Barsanti P, Pesole G, Caggese C. The nuclear OXPHOS genes in insecta: a common evolutionary origin, a common cis-regulatory motif, a common destiny for gene duplicates. BMC Evol Biol. 2007;7: 215.

84. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and postanalysis of large phylogenies. Bioinformatics. 2014;30:1312–3.

85. Hahn C, Bachmann L, Chevreux B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids Res. 2013;41: e129.

86. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69:313–9.

87. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.

88. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17:540–52.

Li *et al. BMC Biology*   (2024) 22:229

Page 14 of 14

89. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012;3:217–23.

90. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2016.

91. Giraudoux P. pgirmess: Data analysis in ecology. https://cran.r-project.org/web/packages/pgirmess/index.html.

92. Wickham H. Programming with ggplot2. In: Wickham H, editor. ggplot2: elegant graphics for data analysis. Cham: Springer International Publishing; 2016. p. 241–53.

93. Yu G, Smith D, Zhu H, Guan Y, Lam TT-Y. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol. 2017;8:28–36.

94. Ames RM, Money D, Ghatge VP, Whelan S, Lovell SC. Determining the evolutionary history of gene families. Bioinformatics. 2012;28:48–55.

Publisher's Note