Newly Sequenced Genomes Reveal Patterns of Gene Family Expansion in select Dragonflies (Odonata: Anisoptera)

Ethan R. Tolman[*1,2,3], Christopher D. Beatty[1,4], Paul B. Frandsen[5,6,7,] Jonas Bush[8], Or R. Bruchim[9], Ella Simone Driever[9], Kathleen M. Harding[10], Dick Jordan[3], Manpreet K. Kohli[1,3,11], Jiwoo Park[9], Seojun Park[9], Kelly Reyes[12,13], Mira Rosari[9], Jisong L. Ryu[9], Vincent Wade[13,14], Jessica L. Ware[1]

[*]Corresponding author

1. American Museum of Natural History, Department of Invertebrate Zoology, New York, New York
2. The City University of New York Graduate Center, New York, New York
3. Conservation Connection Foundation, Boise, Idaho, USA
4. Program for Conservation Genomics, Department of Biology, Stanford University, Stanford, California, USA
5. Department of Plant and Wildlife Sciences, Brigham Young University, Provo, Utah, USA
6. LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Frankfurt, Germany
7. Data Science Lab, Office of the Chief Information Officer, Smithsonian Institution, Washington, District of Columbia, USA
8. Huck Institute of Life Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA
9. Timberline High School, Boise, Idaho, USA
10. Redwood City, California, USA
11. Baruch College, Department of Natural Sciences, New York, New York, USA
12. The Young Women's Leadership School, Manhattan, New York, USA
13. American Museum of Natural History, Science Research Mentorship Consortium, New York, New York, USA
14. Manhattan Village Academy, Manhattan, NY, USA

Classification: Biological Sciences, Evolution

Keywords: Odonata, genomics, gene family expansion, life history

Abstract
As in many other lineages, the dragonflies, or anisoptera (Insecta: Odonata), contain a myriad of life history and diversification patterns. Two families within anisoptera highlight the diversity of patterns in this sub-order. As the most recently derived anisopteran family, some species of Libellulidae typically live fewer than two years (although generation times vary across the group), and some clades are thought to be tolerant of poor water quality. The family comprises 1500 known species. In contrast, an earlier diverging family of anisoptera, the Petaluridae, live longer than five years, and are some of the only semi-terrestrial dragonflies. Petaluridae contains only 11 extant species, all of which specialize in fen habitats. Here we sequence the genomes of two dragonflies, the petalurid, *Uropetala carovei*, and the libellulid, *Pachydiplax longipennis*. By comparing these genomes to previously published genomes of Petaluridae and Libellulidae we identify potential biological functions related to significantly expanding gene families in both lineages. Notably, the Libellulidae have significant expansions of gene families related to the maintenance of homeostasis and gene expression, which we hypothesize could play a role in the ability of the Libellulidae to thrive in low-quality water bodies. The Petaluridae have significantly expanded gene families related to perception of taste and negative regulation of apoptosis, which can plausibly be tied to their semi-terrestrial lifestyle and longevity. Additionally, we demonstrate that Odonata have relatively high rates of gene turnover, and that the difference in gene turnover between Libellulidae and Petaluridae are comparable.

Significance
This work is the first major analysis of the evolution of gene families in the order Odonata, providing future directions for research in this order, and filling in a major taxonomic gap. Furthermore, this is the first work to tie genomic traits to life history and evolutionary patterns in Odonata. The patterns of gene family evolution identified, and the traits they are linked to, will be of broad interest to evolutionary biologists.

Introduction
In dragonflies, why did some families diversify into many species while others remained species poor? The most diverse family in anisoptera, Libellulidae, likely originated in the late Paleogene [1] and has since radiated into over 1500 known species. Libellulidae inhabit a wide variety of aquatic habitats, and some taxa are very tolerant of poor water quality [2]. Although generation times have not been studied for most of the Libellulidae, they are generally understood to have, on average, a relatively rapid life-cycle compared to other families of Odonata [3]

In contrast to Libellulidae, the least diverse family within anisoptera is Petaluridae, which originated in the mid-to-late Jurassic [1,4,5], and contains only 11 extant species, with little fossil evidence of widespread extinction [5]. The ages of the individual species lineages within Petaluridae range between 30 and 130 million years [5]. While Libellulidae inhabit a broad range of habitats, Petaluridae only inhabit fens with high water quality [5–8]. The lifespan of the Petaluridae varies from 5 to 20 years depending on the species, quite possibly making them the longest lived dragonfly [7].

To date, no effort has been undertaken to associate genomic traits with the differing evolutionary and ecological histories of Libellulidae and Petaluridae. With the exception of micro-chromosomes, the genomes of these two families differ very little on a structural scale, as the genomes of the petalurid *Tanypteryx hageni* and libellulid *Pantala flavescens* share widespread synteny with both each other, and two distantly related species of Zygoptera[9]. If genomic traits can be linked to the divergent evolutionary

and ecological histories of Libellulidae and Petaluridae, differences could be found in gene and regulatory sequences, as opposed to genome architecture. Previous research has shown that gene family expansion and contraction can be linked to ecological and evolutionary diversification. Identifying these gene family expansions can and should lead to further research into the functional genomics of these species, which may be tied to adaptations to their different habitats. With this information we can also understand how these species may respond to shifting habitats during climate change.

Gene family evolution and life history in Insecta

Major morphological, physiological and behavioral adaptations are thought to have influenced the remarkable success of insects: flight, herbivory and complete metamorphosis are all suggested to be dramatic drivers of diversity. However, the genomic adaptations that have driven diversity in this group are less well-known. A better understanding of adaptive gene family expansion may help explain why insects have been so successful [10]; work on a number of insect genomes has already been revealing in this context. For example, the expansion of gene families associated with detoxification can be related to polyphagy in the Lepidoptera family Noctuidae [11], and in species of Thysanoptera [12]. Detoxifying genes have also duplicated in the weevil *Rhynchophorus ferrugineus*, as have genes related to insecticide resistance [13]. Expansion of gene families related to nutrient metabolism, development, stress response, reproduction, and immune processes may play a major role in the adaptability of fall army worm, *Spodoptera frugiperda* [14]. The expansion of sensory related gene families can be associated with host differences in Hymenoptera [15]. The Antarctic midge *Parochlus stenenii* has undergone many gene family expansions tied to processes such as the innate immune system, unfolded protein response, DNA packaging, protein folding, and unsaturated fatty acid biosynthesis that are likely adaptive to the frigid environment it inhabits [16]. The expansion of transcription factors, proteases, odorant receptors, and horizontally transferred genes involved in nutritional symbioses probably play an important role in interactions with plants and bacteria in the hemipteran *Bactericera cockerelli* [17]. Gene family expansion has also been linked to adaptations necessary for life in the ocean in the family Hermatobatidae [18]. Duplications of genes within the cytochrome p450 family, and families of olfactory binding proteins are associated with pollution tolerance and mate finding in the caddisfly S*tenopsyche tienmushanensis* [10].

To better understand the gene family evolution of the dragonflies in Libellulidae and Petaluridae, and how gene expansion is related to the evolutionary and ecological histories of the two families, we sequenced and annotated the genomes of the libellulid *Pachydiplax longipennis* and the petalurid *Uropetala carovei*. We then included the annotations in a gene family expansion analysis of all publicly available genomes of Odonata (two Libellulidae, one Petaluridae, and three Zygoptera) that were generated with PacBio HiFi data. Because it is relevant for understanding gene family expansion, we also briefly discuss the life history of each of the five Anisoptera used in the analysis below.

<u>*Pachydiplax longipennis*</u> (Libellulidae)

Perhaps no dragonfly interacts with humans more than the monotypic species *Pachydiplax longennis*. *P. longipennis* has far more iNaturalist observations (90,304 as of October 2023), than any other anisoptera (*Erythemis simplicicollis* is the second most observed species, with 82,504 observations as of October 2023). *P. longipennis* is found throughout the continental United States, Mexico, and several Caribbean Islands. In recent years, its range has expanded northward with a warming climate [19]. *P. longipennis* can be observed in localities with a highly diverse assemblage of Odonata, as well as bodies of water with few species, suggesting a high tolerance to poor water quality [20]. While it has not been formally

studied, gbif occurrences suggest that *P. longipennis* is well adapted to cities, especially in the deserts of the Western United States, where it is commonly observed in urban ponds, but not in nearby natural areas with high odonate diversity, such as national parks (Ellie Gammett unpublished). Many cohorts of *P. longipennis* develop over the course of a year, but bi-voltinism is also observed in this species [21].

Pantala flavescens (Libellulidae)
P. flavescens, the wandering glider, is one of two species in *Pantala*. It is among the most well-studied species of anisoptera, and is the most widely distributed species [22]. It makes long transcontinental and trans-oceanic migrations [23,24], and is frequently found in warmer regions, particularly in the tropics, but is found in temperate regions in the summer [25]. Though it is nNot surprisingly—given the migratory nature of the species—evidence suggests widespread gene flow between populations across the globe, and it may be appropriate to refer to many, if not all, of the migratory sub-populations as panmictic [24]. An smc++[26] analysis, using individuals from Guangdong Province, China found that the effective population size of *P. flavescens* has varied between 1 million and 10 million individuals over the last million years, with declines associated with the onset of the agricultural and industrial period [27].

Sympetrum striolatum (Libellulidae)
S. striolatum is a palearctic dragonfly [28] that is found in Europe and other areas in the Mediterranean [29], with bodies of salt-water likely serving as a major barrier to gene flow within this species [30]. A demographic model of *S. striolatum*, generated from a reference assembly of an individual from Wytham Woods, Oxfordshire, UK [31] estimated a relative stable effective population size around 1 million individuals throughout most of the past 10 million years (which may predate the divergence times of this species), with declines to an effective population size of around 100,000 occurring in the last 100,000 years [32].

Tanypteryx hageni (Petaluridae)
T. hageni is one of only two species in *Tanypteryx* and is the only Petaluridae to inhabit the fens of Western North America, ranging from Southern California, USA, to Southern British Columbia, Canada[6]. *T. hageni* burrows in its nymphal stage, and takes five years to emerge as an adult [6]. A recent behavioral analysis demonstrated that *T. hageni* nymphs regularly leave their burrows for extended periods of time during both day and night, navigating a terrestrial environment [33].

Uropetala carovei (Petaluridae)
U. carovei is one of two species of the genus *Uropetala*, and is found in the lowlands of both the Northern and Southern islands of New Zealand [34]. Their lifespan ranges from five to six years, depending upon the length of the growing seasons [35]. Like *T. hageni*, *U. carovei* burrows in its nymphal stage[36]. It has been suggested that *U. carovei* leaves its burrow to forage [37], but it is unclear if *U. carovei* is as active outside of its burrow as *T. hageni*.

Results

Sequencing and Genome Size Estimation
We generated over 6 million high-fidelity reads, totaling more than 93 Gbp, for *U. carovei* (N50=15,101; L50=2,488,241; ave=14,547.95) and over 5.8 million reads totaling more than 76 Gbp for *P. longipennis* (n50 = 13,394; L50 = 2,369,807; ave= 13,034.64). Using GenomeScope 2.0, we estimated a genome size

of 0.88 Gb (87.4% unique sequence, 1.8% heterozygosity, kcov: 40.5, err: 0.0885%, k:21) for *P. longipennis* (supplementary fig. 1A). and 1.00 Gb (67.5% unique sequence, 1.18% heterozygosity, kcov: 36.7, err: 0.224%, k:21) for *U. carovei* (supplementary fig. 1B).

Genome Assembly and QC

The initial assembly of *U. carovei* was highly contiguous when compared to other Odonata whole genome assemblies (length=1.33Gb, N50=8.67Mb, L50=49, ave=1.72Mb, N=773). The assembly of *P. longipennis* was among the most contiguous genomes ever assembled without scaffolding (length=1.0Gb, N50=66.9 1Gb, L50=7, ave=3.71Mb, N=270). The blobplot of *U. carovei* showed potential contamination (supplementary fig. 2A) of contigs with coverage above 100x, or below 10x. Such contigs that were not assigned to Arthropoda were removed. Only one contig of *P. longipennis*, which was assigned by BLAST to tracheophyta (a common laboratory plant) was removed from the assembly (supplementary fig. 2B), as all other contigs were assigned to Arthropoda. The BUSCO results indicated that the genome was highly complete, with nearly 97 percent of assessed Insecta genes present in the assembly of *U. carovei* (C:96.9%[S:95.9%,D:1.0%],F:1.0%,M:2.1%,n:1367), and 98% of the Insecta genes present in *P. longipennis* (C:98.0% [S:95.2%,D:2.8%], F:0.5%, M:1.5%, n:1367). Both are comparable to other HiFi assemblies of Odonata [38,39].

Whole genome alignment

We generated eight pseudo-chromosomes from the cactus alignment (supplementary table 2). The remaining four chromosomes still showed high levels of synteny with a select number of contigs (supplementary table 2, supplementary fig. 3).

Demographic histories

All estimates for the change in effective population size (N_e) for *T. hageni* using the pairwise sequential markovian coalescent (PSMC) showed an initial drop in population size starting around 30 million years ago (Mya), until a brief uptick in population size approximately 3 Mya, followed by another period of decline stabilizing approximately 200,000 years before present (supplementary fig. 4). Regardless of the mutation rate used in the model, the N_e of *T. hageni* never climbed higher than 2 million individuals, and lower estimates of N_e do not show an effective population size above 1 million individuals (supplementary fig. 4).

The PSMC model suggests the N_e of *Uropetala carovei* fluctuated throughout a period from 25-100 Ma depending on the mutation rate (supplementary fig. 5). In the scenario with the lowest genome-wide mutation rate, from 100 Mya to 10 Mya, the N_e of *Uropetala* increases, but continuously decreases until around a million years ago. In the scenario with the fastest mutation rate, the population starts increasing around 25 Mya and peaks around 2.5 Mya (supplementary fig. 5). Among the different scenarios the highest peak was between 7 and 25 million individuals (supplementary fig. 5).

PSMC models showed one climb in N_e of *P. longipennis*, starting between 7 and 20 Mya, and peaking from 1-6 Mya (supplementary figure 6). The peak N_e was estimated to be between 0.5 and 6 million, with declines stabilizing less than 100k years before present at ~100,000 (supplementary fig. 6).

Annotation

All of our annotations recovered more than 97% of BUSCO Insecta db10 genes (fig. 1A) after filtering. We initially recovered 114,332 genes for *H. americana*, 103,410 for *P. pennipes*, 66,745 for *I. elegans*,

45,043 for *P. longipennis*, 53,217 for *S. striolatum*, 53,629 for *U. carovei*, and 75,637 for *T. hageni*. After filtering for genes that had a significant BLAST hit with the proteome of *P. flavescens* we recovered 13,567 genes for *H. americana*, 15,064 from *P. pennipes*, 15,378 from *I. elegans*, 16,663 from *P. longipennis*, 17,672 from *S. striolatum*, 15,122 from *U. carovei*, and 17,447 from *T. hageni* (fig. 1B).

Gene family evolution

Phylogeny

All expected relationships were recovered with posterior probability support of 1 (fig. 1C). *P. flavescens* was recovered as the outgroup within Libellulidae, with posterior probability support of 1.

Gene birth-death rate

All lambda models converged with the exception of each species being assigned an individual lambda (supplementary table 2). The shared lambda for all Odonata was estimated to be 0.19, with differences between the suborders (anisoptera=0.21, zygoptera=0.16) and families (Libellulidae=0.24, Petaluridae=0.19, Platycnemididae=0.30, Coenagrionidae=0.21, Calopterygidae=0.08) identified in their respective models (fig. 1C).

Expansions and contractions

The suborders had differing numbers of significantly expanding (anisoptera=82, zygoptera=21) and contracting (anisoptera=276, zygoptera=17) gene families (fig. 1C). The Petaluridae had 82 significantly expanding, and 276 significantly contracting gene families, compared to 109 significantly expanding and 291 significantly contracting families in the Libellulidae (fig. 1C). Of the species included, *S. striolatum* had the largest number of significant gene family expansions (282) and *P. flavescens* had the fewest (533). *H. americana* had the highest number of significantly contracting gene families (620), and *U. carovei* had the fewest significantly contracting families (182).

Figure 1: Gene Family Evolution in Petaluridae and Libellulidae

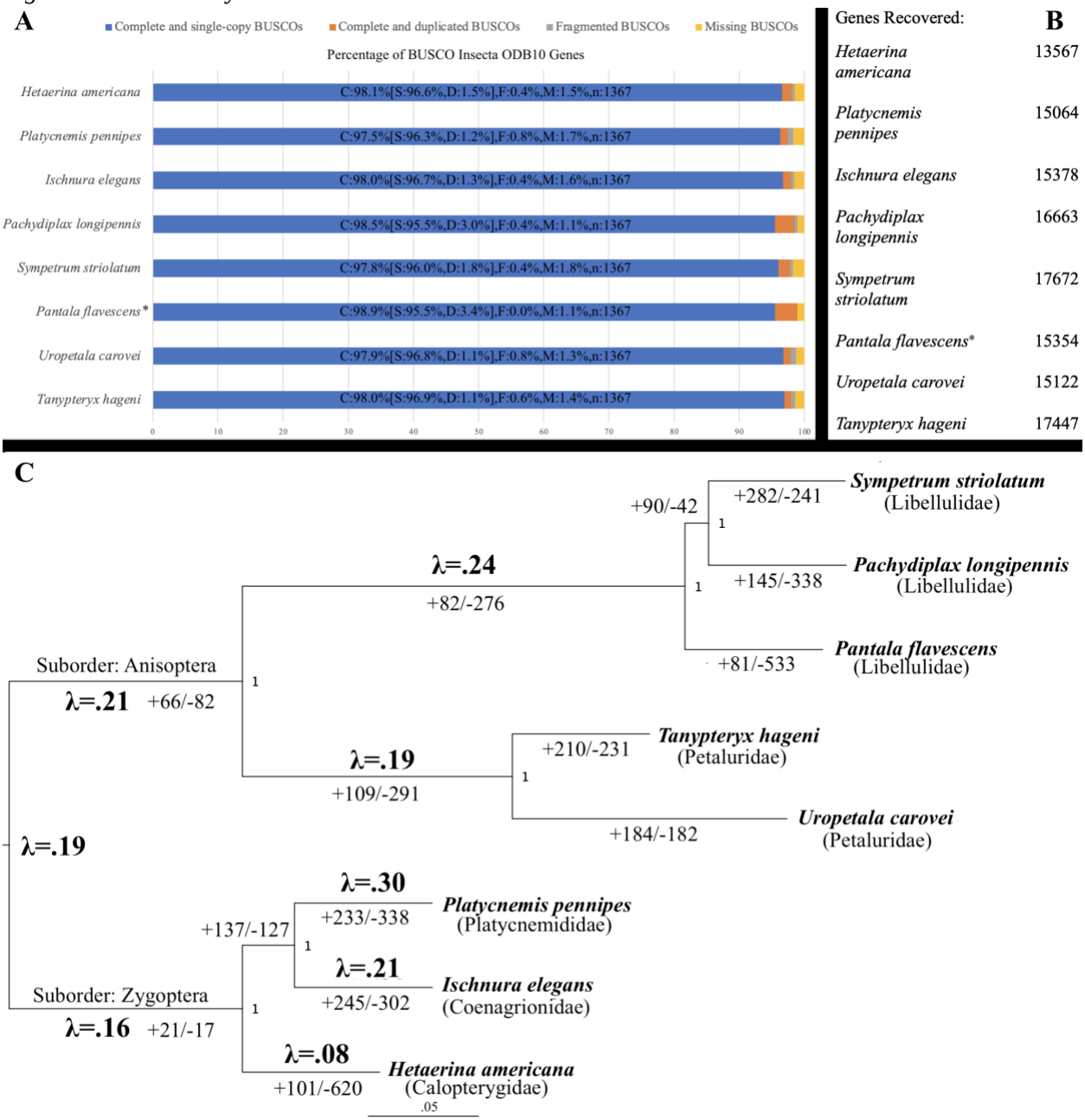

Figure 1: (A) BUSCO results of the eight utilized Odonata genomes. (B) Number of genes recovered with GALBA [40] annotation, after filtering out genes that did not have a BLAST match with the genome annotation of *P. flavescens.* (C) Species tree calculated from BUSCO Insecta genes [41] using Weighted-Astral [42] with posterior probability at each node. The most likely lambda models, and the number of significant (p<.05) gene family expansions and contractions, as identified by CAFE5 [43] are shown on branches and nodes.

*Annotation previously reported [27]

Functional annotation

Figure 2: Enriched GO Terms of Families and Suborders

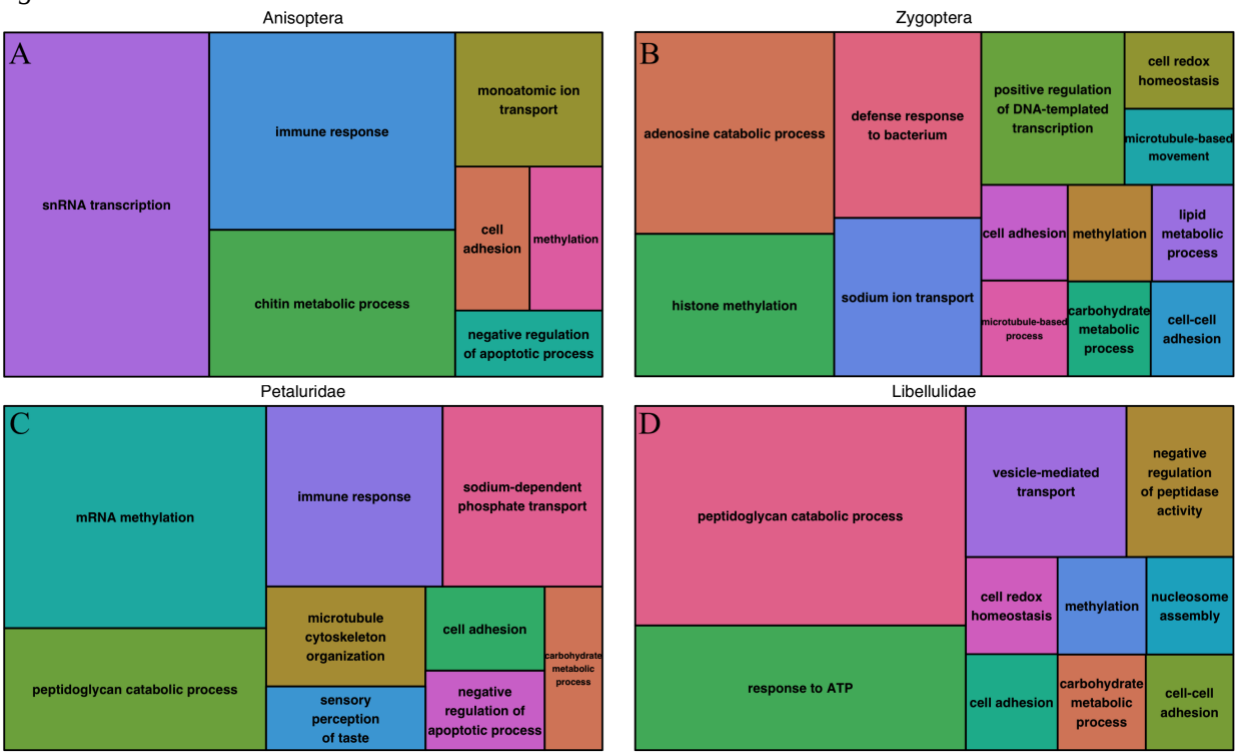

Figure 2: TreeMaps summarizing rapidly expanding gene families in the suborders (A) Anisoptera and (B) Zygoptera, as well as the anisopteran families (C) Petaluridae and (D) Libellulidae. The enriched biological process terms were summarized with REVIGO [44].

Gene Ontology (GO) terms associated with significantly expanding gene families were summarized into 7 biological processes in the suborder anisoptera, and 13 biological processes in the suborder zygoptera. GO terms were summarized into 9 biological processes in significantly expanding gene families in Petaluridae (fig. 2C) and 10 biological processes from the significantly expanding gene families of Libellulidae (fig. 2D).

Figure 3: Enriched GO Terms of Anisoptera

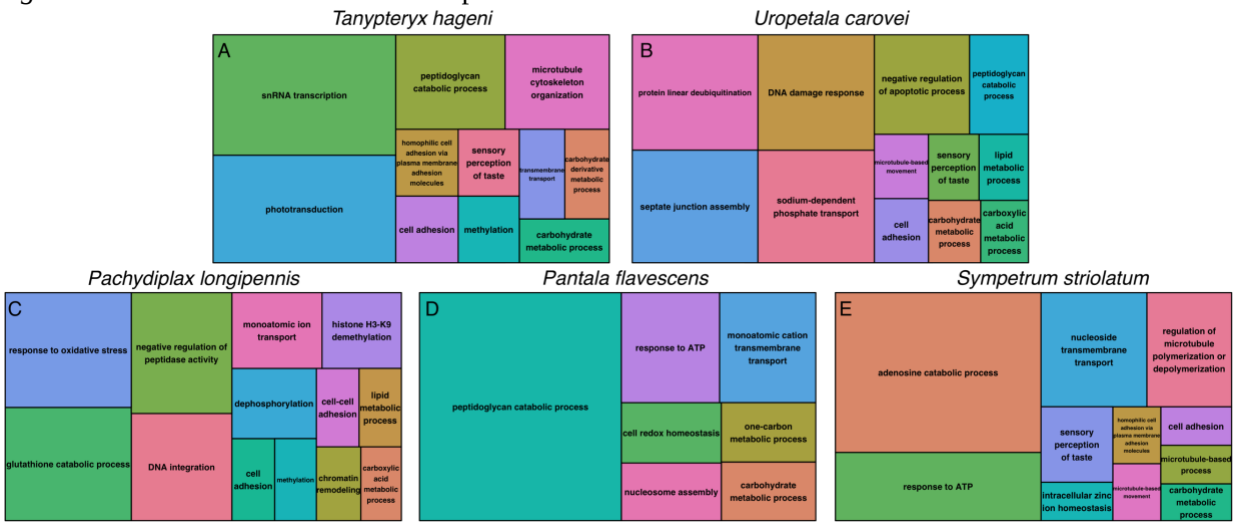

Figure 3: TreeMaps summarizing rapidly expanding gene families in the Petaluridae (A) Tanypteryx hageni and (B) Uropetala carovei, as well as the Libellulidae (C) Pachydiplax longipennis, (D) Pantala flavescens, and (E) Sympetrum striolatum. The enriched biological process terms were summarized with REVIGO [44].

In the Petaluridae, 11 biological processes were summarized from GO terms associated with significant gene family expansion in *T. hageni* (fig. 3A), and 12 biological processes were summarized from GO terms associated with significant gene family expansion in *U. carovei* (fig. 3B). GO terms associated with significantly expanding gene families in the Libellulidae *P. longipennis*, *P. flavescens*, and *S. striolatum* were summarized into 13, 7, and 11 biological processes respectively (fig. 3C-E).

Discussion

Our results provide a number of testable hypotheses relating to genome size, gene familiy expansion, and the diversification of odonates into the range of ecosystems they currently inhabit. We consider here the strengths and implications of the technologies used to generate these results, and then some of the major research areas suggested by our results.

Genome size estimation and quality control

Genome size estimates of Odonata assembled from PacBio high-fidelity sequences have been found to diverge from genome sizes estimated with Genomescope [45]. Although many contigs from the *U. carovei* assembly were not assigned to the Arthropoda, it is likely an artifact of an incomplete NCBI reference database [38]. Another Petaluridae, *T. hageni*, showed a similar profile with contigs not assigned to Arthropoda, even in well supported scaffolds with HiC data, likely due to poor coverage in NCBI databases [38]. This phenomenon has also been observed in horseshoe crabs [38]. With our current understanding, taxonomically mis-assigned contigs with a similar GC content and coverage to contigs not assigned to Arthropoda are not a major source of concern [38].

Implications of PacBio Revio Technology in Insect genomics

PacBio high-fidelity sequencing has become the gold standard for insect genomics [46]. The near chromosome-level assembly of *P. longipennis* presented here, before any scaffolding, further demonstrates the potential of this sequencing technology. As this technology is further developed, complete, chromosome-length genome assemblies will be available without any scaffolding.

GALBA as a tools for genome annotation

The genome annotation tool GALBA has been demonstrated to outperform more established pipelines, such as BRAKER2 [47] when annotating insect genomes, and does not need further sequencing of transcriptomes [40]. While our genome annotations initially had a high (>50,000) number of genes compared to the annotation of *P. flavescens*, by conservatively retaining only genes which had a BLAST hit to a protein sequence from *P. flavescens* [27], we were able to generate highly complete annotations (fig 1. A,B), further supporting the efficacy of this pipeline in insects.

Demographic histories

Similar to *I. elegans* and *S. striolatum* [32] the PSMC model for *P. longipennis* showed a decline in effective population size $(N)_e$ roughly corresponding with the Pleistocene ice ages. As the climate warms this species could become more widely distributed, but more research into the ecological niche of this species is needed to affirm this assumption. The N_e of the two fen-dwelling petalurid species–*U. carovei* and *T. hageni*–were estimated to be smaller than *P. longipennis* or the other libellulids through time [27,32]. If the Petaluridae have consistently been habitat specialists through the period covered by our analysis, we would predict that their population size would be limited by availablity of their habitat.

Gene birth/death rate

While models estimating different lambda values across the tree are not directly comparable in a statistical framework, the values we recover show similar trends. The lambda estimated for each suborder is approximately the mean of the branches in the suborder, and the global lambda is approximately the mean of both suborders, or each family. With the exception of *P. pennipes* (λ=0.08), lambda values across the estimated tree are much higher than those found elsewhere across the tree of life (e.g., Libellulidae λ=0.24). For example, global lambdas of primates (λ<0.01), birds of paradise (λ<0.01), and the insect orders Hymenoptera (λ<0.01) [43], and Hemiptera (λ<0.01) [17] are all less than 5% of the global lambda of Odonata (λ=0.19). This is not solely due to the age of Odonata (anisoptera and zygoptera split some 250 MA [1,4]), as the tree used to study gene family expansion in Hemiptera was estimated to have a root divergence time older than 500 Mya [17]. An analysis of the synteny between chromosome level genome assemblies of Odonata found a high degree of conservation in chromosome structure, even between suborders [9]. In *P. flavescens* and *T. hageni* the relative proportion of CDS content was not retained in a linear model predicting the proportion of synteny in a given partition of the genome assembly, suggesting divergence in gene sequences, even though the chromosome structure was largely conserved [9]. The high gene turnover rates observed in Odonata could explain this dissociation between CDS and synteny. Although caution is necessary when comparing λ values between families (as we do not have complete taxonomic coverage, and CAFE5 does not test for significant differences between the different specified branches) it is notable that the birth/death rate for Libellulidae (λ=0.24) differs from that for Petaluridae (λ=0.19). Our estimates, made with currently available data, provide evidence that high levels of ecological diversification and speciation in Libellulidae, and the habitat specialization and low levels branching of the Petaluridae, could both be supported in systems with high levels of gene turnover.

T. hageni and *U. carovei* represent the two major clades of Petaluridae [48] (Laurasian and Gondwanan) and ~18% of the family. However, the three species of Libellulidae anlayzed here represent only 3 of 29 genera and 3 of >1500 species. Additionally, our analysis includes only 2 of the 39 families of Anisoptera, 3 of the 39 Zygoptera, and only 8 of ~6300 species of Odonata [49], so we acknowledge the birth/death rates calculated could be an artifact of poor taxon sampling, and more genomes are needed to more accurately model gene birth and death. We have included all of the high-quality odonate genome assemblies currently available in this analysis, offering predictions about birth/death rates in this order that can be refined as more data becomes available.

While more research is necessary to better understand the molecular processes tied to the evolutionary and ecological histories of odonates, our functional annotations of expanding gene families can be used to generate testable hypotheses related to the ecology and evolution of Petaluridae and Libellulidae. Here, we discuss hypotheses derived from our GO analyses that warrent further investigation in future studies.

As *T. hageni* and *U. carovei* represent ~18% of all Petaluridae, and both major clades within the family [48], we believe it is highly likely that gene family expansions shared between these two species are shared across Petaluridae. However, as the family Gomphidae is likely sister to Petaluridae [1,4,50], without representative genomes for Gomphidae, we cannot determine whether these expanded gene families would be found in both families, thereby suggesting that they may have originated in a most-recent common ancestor. Given these limitations, the expanding gene families shared between the two Petaluridae still suggest candidate genes to explore for genomic adaptations related to the life history attributes of this family, namely the longevity of these lineages through evolutionary time, the long development times seen in these species, and the unique semi-terrestrial lifestyle of the nymphs.

Hypothesis 1: longer lived odonates have expanded immune response families. One of the exapanded gene families seen in the petalurids is associated with *immune response*. A possibilty is that these are associated with organisms descended from older lineages; another long-lived arthropod lineage, the mangrove horseshoe crab, also has an expansion of genes related to immune response [51]. The two petalurid species analyzed here separated from their closest relatives approximately 70 Mya [5]. To explore this correlation, data on more "ancient" arthropod lineages are needed.

Hypothesis 2: terrestrial nymphs have different sensory perception profiles. Across the family Petaluridae, and in both representative species, we detected an expansion in genes that play a part in *sensory perception of taste* (fig. 2C). While all odonate nymphs are aquatic, the nymphs of these petaltail species have been observed to forage in the terrestrial habitats around their subterranean burrows (which fill with water, providing an aquatic environment). Odonates are generalist predators in their nymph and adults stages. Terrestrial foraging could bring these petalurid nymphs in contact with a range of potential prey very different from those they would encounter in a strictly aquatic environment. Insects generally utilize taste to avoid toxic chemicals, find suitable food sources, select mates and find ovipositon sites [52]. Adult Odonata, in particular, have been observed catching and rejecting toxic prey upon tasting it [52]. It is currently unclear if nymphs exhibit this behavior. With terrestrial foraging, petalurid nymphs may encounter a greater proportion of toxic prey, such as arachnids (which have been observed in and around the burrows of *T. hageni* [33] Further examination of the specific genes involved in taste perception in petalurids needs to be performed, along with greater documentation of their foraging behavior and analysis of their diet to test this hypothesis.

Hypothesis 3: terrestrial nymphs have different phototransduction gene families. The expansion of gene families related to *phototransduction* in *T. hageni* could be due to the differing light

conditions the nymphs face in a terrestrial environment. Opsin genes in particular have undergone a number of expansions and contractions across Odonata, but the reasons for this evolutionary pattern are currently unclear [53]. Transcriptomic work will be needed to confirm whether expanded gene families related to phototransduction are expressed in the nymphs. As the nymphal behavior of the Petaluridae is better understood, and more genomes and transcriptomes are sequenced, gene families associated with taste and vision could be a promising avenue for understanding their interaction with a terrestrial environment as nymphs.

Hypothesis 4 : highly tolerant dragonflies have expanded gene families to deal with dessication and bactierial exposure. While it is certainly plausible that shared expansions in the three Libellulidae analyzed here are shared across the family, we will be more certain about gene family expansions among libellulids when more genomes are available. Several biological processes related to the shared expansions of these three taxa may be associated with the generalist life-history of libellulids. A major overrepresented GO term shared between these three species is the *peptidoglycan catabolic process* (fig. 2D). Peptidoglycan is an essential element in the walls of most bacteria [54]. Several hypotheses could explain why gene families related to catabolizing peptidoglycan would be present. First: genes involved in the peptidoglycan catabolic process were upregulated when the Antarctic midge *Belgica antarctica* was placed under desiccation conditions [55]. In the case of odonates, expansions could be adaptive to deal with more arid environments, or in ephemeral aquatic environments. Secondly, Libellulidae often inhabit warm, eutrophic waters, where harmful bacteria are more likely to be found [56]. Expansion of gene families involved in the peptidoglycan catabolic process could enable Libellulidae to tolerate exposure to high levels of bacteria.

Hypothesis 5 : highly tolerant dragonflies have expanded gene families to deal with exposure to free radicals. Another GO term associated with expanded gene families is *cell redox homeostasis*, a dynamic but essential process ensuring balance between oxidizing and reducing reactions in cells that is also involved in many biological responses [57]. These elements have been shown to maintain homeostasis in fluid conditions [57]. An ability to maintain homeostasis would be critical for Libellulidae, as the shallow bodies of water they inhabit can often undergo rapid changes in temperature and water chemistry. GO terms associated with expanding gene families of *P. longipennis* could offer some explanation for how this species has been so successful in colonizing urban environments. One of the most overrepresented GO terms for this species was *response to oxidative stress* (fig. 3D). Birds in urban or polluted environments have increased antioxidant protection, and face higher levels of oxidative stress [58], so it is likely that Odonata face these same pressures. It follows that a greater capacity to respond to oxidative stress in urban environments would allow *P. longipennis* to outcompete other species that do not have this capacity. Additionally, the expansion of gene families related to *chromatin remodeling* and *methylation* (fig. 3D) could allow *P. longipennis* sufficient plasticity to thrive in an urban environment.

Hypothesis 6: highly tolerant dragonflies have expanded gene families to allow for greater phenotypic plasticity. Expansions related to both *methylation*, and *response to ATP* could result in increased phenotypic and molecular plasticity, which are a necessity for evolutionary stability in fluctuating environments [59], and could be another molecular adaptation to fluctuating environments.

Hypothesis 7: migratory dragonflies can more efficiently process carbohydrates. In *P. flavescens*, the expansion of gene families related to the *carbohydrate metabolic process* are a plausible mechanism for allowing this species to undertake trans-oceanic and trans-continental migrations in a single generation [24]. It is likely that an especially efficient process for obtaining energy from

carbohydrates would be selected for on such long migrations. The genome of closely related non-migratory species to *P. flavescens* are needed to explore the genomic adaptations to a migratory lifestyle.

Although they represent only 2 of the 39 families of Odonata, the Petaluridae and Libellulidae are phylogenetically distant [1,4,50,60], so shared gene family expansions and contractions could be shared throughout the anisoptera, but the sequencing of more families will be necessary to better understand gene family expansion in this suborder. Of note are the shared expansion of gene families related to the *chitin metabolic process* (fig. 3a). This could be due to differing diets between the suborders (as anisoptera are capable of capturing larger prey), but further research is needed to determine the cause.

In our current analysis only 3 of the 39 families in the order zygoptera are represented; these three families are relatively phylogenetically close to one another [1,4], and as such broad interpretation of the results of significantly expanding gene families is difficult. We included these genomes in our analysis as they at least offer a useful comparion to our findings in the anisoptera.

Conclusions

Our findings are the first comprehensive analysis of gene family expansion in Odonata. The results suggest that gene turnover is much higher in Odonata than in other insect orders, and we identify a number of gene families that may be associated with the differing evolutionary and ecological histories of the Petaluridae and the Libellulidae. The role of these gene families, and the relatively high levels of gene turnover will be an exciting avenue of future research as more genomes from across the Odonata, along with more transcriptome data, become available.

Materials and Methods

Data selection and genome sequencing

Selection of publicly available genomes

In addition to the genomes of *P. longipennis* and *U. carovei* we included publicly available genomes of Odonata from GenBank (October 1, 2023). To minimize bias due to differing genome sequencing methods we only included genomes that had been generated from PacBio high-fidelity technology, the gold standard for insect genomics [46]. We identified four anisopteran genomes sequenced with this technology, comprising two Libellulidae (*Pantala flavescens* [27] and *Sympetrum striolatum* [31]) and one Petaluridae (*Tanypteryx hageni* [38]) . We also chose to include the three publicly available genome assemblies of Zygoptera, the Platycnemididae *Platycnemis pennipes* [61], the Coenagrionidae *Ischnura elegans* [62], and the Calopterygidae *Hetaerina americana* [63] to use as an outgroup in the gene family expansion analysis.

Specimen collection

We sequenced the genome of *Uropetala carovei*, which diverged from the lineage leading to *T. hageni* in the late Jurassic [5], to represent the previously unsampled gondwanan clade of Petaluridae. *U. carovei* was collected as a nymph near Aorere goldfields in New Zealand, and was injected with and stored in 100% ethanol prior to DNA extraction and sequencing. To expand the taxon coverage of the Libellulidae we sequenced the genome of *Pachydiplax longipennis*. An adult specimen was collected from Parkcenter Park Pond in Boise, Idaho, and was stored in 100% ethanol prior to sequencing.

Sequencing and Genome Size Estimation

High molecular weight DNA was extracted from both the *U. carovei* and *P. longipennis* specimens using the Qiagen Genomic-tip kit. High molecular weight DNA was sheared to 18 kbp using a Diagenode Megaruptor and separate libraries were prepared for each specimen using the PacBio HiFi SMRTbell Express Template Kit 2.0. Each library was sequenced on a single PacBio Revio SMRT cell in CCS mode at the BYU DNA sequencing center.

Genome Assembly and QC

For both *U. carovei* and *P. longipennis* we used HiFi reads to estimate genome size with Genomescope 2.0 [45]. We first generated read histograms with a histogram length of 21 using Jellyfish [64], and ran Genomescope 2.0 with default settings. We generated both whole genome assemblies with hifiasm v.0.16.0 [65].

We used taxon-annotated GC-coverage plots generated with BlobTools v1.1.1 [66] to screen for contamination. We mapped all HiFi reads against the final assembly for each species using minimap2 v2.1 [67] and sorted the resulting bam files with samtools v1.13 [68]. We then assigned taxonomy with megablast [69] using the parameters: task megablast and -e-value 1e-25. We calculated coverage using the blobtools function map2cov, created the blobtools database for each species using the command blobdb, and generated the blobplot with the command "blobtools plot". After examining the blobplots we removed one contig from the assembly of *P. longipennis* which was assigned to the phylum tracheophyta, and removed all contigs from the assembly of *U. carovei* with coverage above 100x, or below 10x, which did not blast to Arthropoda.

Additional contaminants were automatically removed from the assembly of *U. carovei* as a part of the NCBI genome upload (supplementary table 1). No additional contaminants were removed from the assembly of *P. longipennis*.

We generated all contiguity stats for both assemblies with the command line tool assembly-stats (v0.1.4) [70]. We assessed the completeness of both assemblies with BUSCO (v.4.1.4) [41] using the Insecta ODB10 database, in genome mode with the flag –long to retrain the initial BUSCO results for a more accurate identification of BUSCO genes.

Whole genome alignment

Given the contiguity of the assembly of *P. longipennis*, we calculated synteny between the draft genome assembly and the chromosome length genome assembly of *S. striolatum* to identify any complete chromosomes. We identified potentially homologous genes using BLASTp [71], restricting output to 5 hits per gene and filtering for hits with an e-value below 1e-5. We combined the BLAST output and gff annotation files for *P. longipennis* and *S. striolatum* and used MCscanX [72] to identify areas of synteny between the genomes with a minimum of 5 genes. We visualized the syntenic relationships in a Circos (v.0.69-18)[73] plot to identify chromosome-length contigs.

To assemble the remaining contigs into chromosomes, we aligned the chromosome length assemblies of *Pantala flavescens*, *Sympetrum striolatum*, *Tanypteryx hageni* and the Platycnemididae *Platycnemis pennipes* with the assembly of *P. longipennis* using progressive cactus v.2.6.7 [74]. We then used ragout v.2.3 [75,76], with the option --solid-scaffolds, to generate a scaffolded assembly from the alignment. As the karyotype of *U. carovei* is currently unclear [77], it would not be possible to assess the quality of the scaffolding, so we did not attempt to scaffold this genome. This would not impact the results of the gene family expansion analysis.

Gene family expansion analysis

Data selection

In addition to the genomes of *P. longipennis* and *U. carovei* we included publicly available genomes of Odonata from GenBank (October 1, 2023). To minimize bias due to differing genome sequencing methods we only included genomes that had been generated from PacBio high-fidelity technology, which has been shown to generate the highest quality insect genomes to date [46]. We identified four anisopteran genomes sequenced with this technology, comprising two Libellulidae (*Pantala flavescens* [27] and *Sympetrum striolatum* [31]) and one Petaluridae (*Tanypteryx hageni* [38]) . We also chose to include the three publicly available genome assemblies of Zygoptera, the Platycnemididae *Platycnemis pennipes* [61], the Coenagrionidae *Ischnura elegans* [62], and the Calopterygidae *Hetaerina americana* [63] to use as an outgroup in the gene family expansion analysis.

Demographic histories

As the demographic history of a species may be important when studying genome evolution, particularly in relic lineages [78,79], we estimated the demographic histories of *U. carovei*, *P. longipennis* and *T. hageni* (estimates for all other species have previously been described [27,32]). As with the blobtools analysis, HiFi reads were mapped to the genome assemblies with minimap2 v2.1.1 [67], and the bam files were sorted with samtools v1.16.1 [68]. Samtools v1.16.1 and bcftools v1.6 [68], with a minimum depth of 5 and a maximum depth of 100, were used to make base calls from the sorted bam files. We then used psmc v0.6.5-r67 [26] to convert the fastq basecalling output to psmcfa format and estimate the pairwise sequential markovian coalescent (PSMC) with 100 bootstrap replicates for each species. For all three species we visualized the PSMC with the programs psmc [26] and gnuplot v5.2 [80] using four different mutation rates (1e-9, 2e-9, 3e-9, and 4e-9), encompassing known genome-wide mutation rates in Insecta [81], and values previously used in demographic analyses of Odonata [27,32]. As *P. longipennis* cohorts are typically uni-voltine, but occasionally exhibit bi-voltinism [21] , we used a generation time of 0.8 to reflect the occasional bi-voltine generations. *U. carovei* has a generation time of 5-6 years depending on the length of growth seasons [7,82] so we used a generation time of 5.5 years. *T. hageni* has a life cycle of 5 years [6,7], and this was the value used in the coalescent analysis.

Annotation

To ensure consistency in the genome annotations we annotated with genomes of *H. americana*, *P. pennipes*, *I. elegans*, *P. longipennis*, *S. striolatum*, *U. carovei*, and *T. hageni* in GALBA [40] using the annotated protein set of *P. flavescens* [27]. GALBA aligns the reference protein set to the soft-masked target genome using miniprot [83] and uses the alignments to train Augustus [84,85], and has produced more accurate annotations of insect genomes than common annotation pipelines such as BRAKER2 [40,47]. Of the published genomes of Odonata, the annotated protein set of *P. flavescens* [27] is by far the most complete [86], with nearly 99% of BUSCO genes present (fig 1.A), justifying its use as a reference. Prior to annotation with GALBA, repetitive elements of all target species were modeled and soft-masked with RepeatModeler2 (v2.0.1) and RepeatMasker (v.4.1.2-p1)[87]. Because the initial GALBA annotations found an unreasonably high (>50,000) number of genes, we only kept annotated genes with a significant BLAST[71] hit (with an e-value cutoff of 1e-10) to the annotated protein set of *P. flavescens* [27]. Given the completeness of this annotated protein set [27],this allowed us to be conservative in our identification of expanding gene families, although we acknowledge we could potentially lose gene families that are completely lost in *P. flavescens*. However, the new filtered annations are much more complete than the original genome annotations [38]. We generated all contiguity stats for both assemblies

with the command line tool assembly-stats (v0.1.4) [70]. We assessed the completeness of all annotations with BUSCO (v4.1.4) [41] using the Insecta ODB10 database, in protein mode.

As part of the annotation process, we also identified the mitochondrial genomes from the de novo assembles of *P. longipennis* and *U. carovei* using mitohifi (v. 3.0.0) [88]. We used the mitochondrial genome of *T. hageni* [38]] as a reference for *U. carovei* and the mitochondrial genome of *S. striolatum* [31] as a reference for *P. longipennis*. The mitochondrial genome sequences have been uploaded, as part of a the reference genome sequences, to GenBank.

Gene family expansion

While the phylogenetic relationships between the families of Odonata are well established [1,4,50], the relationships between the three Libelullidae (*P. longipennis*, *S. striolatum*, and *P. flavescens*) have not previously been studied with genomic data. To resolve this relationship, we generated a species tree using the amino acid sequences of 1102 BUSCO Insecta odb10 genes present in all eight species. We calculated a maximum likelihood tree for each gene with iqtree2 [89] using the option "-m test" for model selection. We then generated a species tree using the estimated gene trees with weighted ASTRAL [42].

We clustered the amino sequences from the longest isoforms of each gene for each species with Orthofinder [90]. We then used CAFE5 [43] to generate four different lambda (gene birth/death rate)models.

1. A global lambda for the whole tree
2. Separate lambdas for anisoptera and zygoptera
3. Separate lambdas for each family
4. Separate lambdas for each species

As models generated with CAFE5 are not likely to converge when using a large number of parameters [91], we only tested the base and error models. For each lambda model we ran the base and error models three times to check for convergence, and selected the model with the highest likelihood. We then identified significantly expanding and contracting gene families from the highest likelihood model (either base or error).

Functional annotation

We obtained GO terms for all of the protein sequences (with the exception of *P. flavescens* which were previously reported [27]) using interproscan [92–94]. We summarized the GO terms of significantly expanding gene families into TreeMaps with REVIGO [44].

Data availability

The genome assemblies and rawreads of *U. carovei* (PRJNA1043152 and *P. longipennis* (PRJNA999334) are available on GenBank. RepeatMasker output (10.6084/m9.figshare.24714909), unfiltered annotations (10.6084/m9.figshare.24714918) and filtered proteins sets used for gene family expansion analysis (10.6084/m9.figshare) are all available on figshare.

Competing interests

The authors declare no competing interests

Acknowledgements
We would like to thank Kristin Gnojewski and Cindy Busche for their support of this project. We acknowledge funding from the Conservation Connection Foundation, the Society of Systematic Biologists Graduate Student Research Award, and the 2023 Sydney Anderson Travel Award from the American Museum of Natural History.

References

1. Suvorov A, Scornavacca C, Fujimoto MS, Bodily P, Clement M, Crandall KA, et al.. Deep Ancestral Introgression Shapes Evolutionary History of Dragonflies and Damselflies. *Systematic Biology*. 2021; doi: 10.1093/sysbio/syab063.

2. Šigutová H, Dolný A, Samways MJ, Sönke H, José Max Oliveira- J, Juen L, et al.. Odonata as indicators of pollution, habitat quality, and landscape disturbance. In: Cordoba-Aguilar A, Beatty C, Bried J, editors. *Dragonflies and Damselflies: Model Organisms for Ecological and Evolutionary Research*. Oxford University Press;

3. Tennessen K. Libellulidae. In: Tennessen KJ, editor. *Dragonfly Nymphs of North America: An Identification Guide*. Cham: Springer International Publishing;

4. Kohli M, Letsch H, Greve C, Béthoux O, Deregnaucourt I, Liu S, et al.. Evolutionary history and divergence times of Odonata (dragonflies and damselflies) revealed through transcriptomics. *iScience*. 2021; doi: 10.1016/j.isci.2021.103324.

5. Ware JL, Beatty CD, Herrera MS, Valley S, Johnson J, Kerst C, et al.. The petaltail dragonflies (Odonata: Petaluridae): Mesozoic habitat specialists that survive to the modern day. *Journal of Biogeography*. 2014; doi: 10.1111/jbi.12273.

6. Turner PE. Allusive Dragons: Functional Integration and Structural Integrity among Natural Populations of the Relict Dragonfly Tanyptery hageni. [PhD Thesis]. [Berkeley, California]: University of California, Berkeley;

7. Baird I. The wetland habitats, biogeography and population dynamics of Petalura gigantea (Odonata: Petaluridae) in the Blue Mountains of New South Wales [PhD]. The University of Western Sydney;

8. Baird IRC. Establishment of larval pits by Tachopteryx thoreyi (Odonata: Petaluridae): habitat modification by a non-burrowing petalurid. *International Journal of Odonatology*. Taylor & Francis; 2019; doi: 10.1080/13887890.2019.1636889.

9. Tolman ER, Beatty CD, Bush J, Kohli MK, Frandsen PB, Gosnell JS, et al.. Exploring chromosome evolution in 250 million year old groups of dragonflies and damselflies (Insecta:Odonata). *Mol Ecol*. 2023; doi: 10.1111/mec.17147.

10. Luo S, Tang M, Frandsen PB, Stewart RJ, Zhou X. The genome of an underwater architect, the caddisfly Stenopsyche tienmushanensis Hwang (Insecta: Trichoptera). *GigaScience*. 2018; doi: 10.1093/gigascience/giy143.

11. Breeschoten T, van der Linden CFH, Ros VID, Schranz ME, Simon S. Expanding the Menu: Are Polyphagy and Gene Family Expansions Linked across Lepidoptera? *Genome Biology and Evolution*. 2022; doi: 10.1093/gbe/evab283.

12. Hu Q-L, Ye Z-X, Zhuo J-C, Li J-M, Zhang C-X. A chromosome-level genome assembly of Stenchaetothrips biformis and comparative genomic analysis highlights distinct host adaptations among thrips. *Commun Biol*. Nature Publishing Group; 2023; doi: 10.1038/s42003-023-05187-1.

13. Hazzouri KM, Sudalaimuthuasari N, Kundu B, Nelson D, Al-Deeb MA, Le Mansour A, et al.. The genome of pest Rhynchophorus ferrugineus reveals gene families important at the plant-beetle interface. *Commun Biol*. Nature Publishing Group; 2020; doi: 10.1038/s42003-020-1060-8.

14. Gao H, Li Y, Tian Y, Zhang H, Kim K, Li B. Gene family expansion analysis and identification of the histone family in Spodoptera frugiperda. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*. 2023; doi: 10.1016/j.cbd.2023.101142.

15. Kuang J-G, Han Z-T, Kang M-H, Ju T, Tso S, Chen X-X, et al.. Chromosome-level de novo genome assembly of two conifer-parasitic wasps, Megastigmus duclouxiana and Megastigmus sabinae, reveals genomic imprints of adaptation to hosts. *Molecular Ecology Resources*. 2023; doi: 10.1111/1755-0998.13785.

16. Kim H, Kim H-W, Lee JH, Park J, Lee H, Kim S, et al.. Gene family expansions in Antarctic winged midge as a strategy for adaptation to cold environments. *Sci Rep*. Nature Publishing Group; 2022; doi: 10.1038/s41598-022-23268-9.

17. Kwak Y, Argandona JA, Degnan PH, Hansen AK. Chromosomal-level assembly of Bactericera cockerelli reveals rampant gene family expansions impacting genome structure, function and insect-microbe-plant-interactions. *Molecular Ecology Resources.* 2023; doi: 10.1111/1755-0998.13693.

18. Wang Y, Luan Y, Luo J, Men Y, Engel MS, Damgaard J, et al.. 300 Million years of coral treaders (Insecta: Heteroptera: Hermatobatidae) back to the ocean in the phylogenetic context of Arthropoda. *Proceedings of the Royal Society B: Biological Sciences*. Royal Society; 2023; doi: 10.1098/rspb.2023.0855.

19. Lis C, Moore MP, Martin RA. Warm developmental temperatures induce non-adaptive plasticity in the intrasexually selected colouration of a dragonfly. *Ecological Entomology*. 2020; doi: 10.1111/een.12839.

20. Tollett VD, Benvenutti EL, Deer LA, Rice TM. Differential Toxicity to Cd, Pb, and Cu in Dragonfly Larvae (Insecta: Odonata). *Arch Environ Contam Toxicol.* 2009; doi: 10.1007/s00244-008-9170-1.

21. Wissinger SA. Life History and Size Structure of Larval Dragonfly Populations. *Journal of the North American Benthological Society*. The University of Chicago Press; 1988; doi: 10.2307/1467827.

22. Russell RW, May ML, Soltesz KL, Fitzpatrick JW. Massive Swarm Migrations of Dragonflies (Odonata) in Eastern North America. *amid*. University of Notre Dame; 1998; doi: 10.1674/0003-0031(1998)140[0325:MSMODO]2.0.CO;2.

23. Rowe RJ. Conservation of Odonata in the South Pacific and Australasia. *International Journal of Odonatology*. 2004; doi: 10.1080/13887890.2004.9748206.

24. Ware J, Kohli MK, Mendoza CM, Troast D, Jinguji H, Hobson KA, et al.. Evidence for widespread gene flow and migration in the Globe Skimmer dragonfly Pantala flavescens. *International Journal of Odonatology*. Wachholtz Verlag; 25:43–552022;

25. May ML. A critical overview of progress in studies of migration of dragonflies (Odonata: Anisoptera), with emphasis on North America. *J Insect Conserv*. 2013; doi: 10.1007/s10841-012-9540-x.

26. Liu S, Hansen MM. PSMC (pairwise sequentially Markovian coalescent) analysis of RAD (restriction site associated DNA) sequencing data. *Molecular Ecology Resources*. 2017; doi: https://doi.org/10.1111/1755-0998.12606.

27. Liu H, Jiang F, Wang S, Wang H, Wang A, Zhao H, et al.. Chromosome-level genome of the globe skimmer dragonfly (Pantala flavescens). *GigaScience*. 2022; doi: 10.1093/gigascience/giac009.

28. John Horne. Emergence, maturation time and oviposition in the Common Darter Sympetrum striolatum (Charpentier). *Journal of the British Dragonfly Society*. 282012;

29. Borkenstein A, Jödicke R. Thermoregulatory behaviour of Sympetrum striolatum at low temperatures with special reference to the role of direct sunlight (Odonata: Libellulidae). 51:83–1092022;

30. Parkes K, Amos W, Moore N, Hoffman J, Moore J. Population structure and speciation in the dragonfly Sympetrum striolatum/nigrescens (Odonata: Libellulidae): An analysis using AFLP markers. *European Journal of Entomology*. 2009; doi: 10.14411/eje.2009.021.

31. Crowley LM, Price BW, Allan EL, Eagles M, University of Oxford and Wytham Woods Genome Acquisition Lab, Natural History Museum Genome Acquisition Lab, et al.. The genome sequence of the Common Darter, Sympetrum striolatum (Charpentier, 1840). *Wellcome Open Res*. 2023; doi: 10.12688/wellcomeopenres.19937.1.

32. Ethan R. Tolman, Or Bruchim, Ella Driever, Manpreet K. Kohli, Dick Jordan, Lilly Montague, et al.. Demographic Histories of Odonata Revealed Through Genomics. *International Journal of Odonatology*. In Review;

33. Christopher D. Beatty, Aaron Goodman, Ethan R. Tolman, Jessica L. Ware. The Larval Behavior of Tanypteryx Hageni. In Preparation;

34. Rowe RJ. AN ANNOTATED KEY TO THE NEW ZEALAND ODONATA. 91981;

35. L. S. Wolfe. A study of the genus Uropetala Selys (Order Odonata) from New Zealand. *Transactions of the Royal Society of New Zealand*. 80:245–75.

36. Green LFB. Aspects of the respiratory and excretory physiology of the nymph of Uropetala carovei (Odonata: Petaluridae). *New Zealand Journal of Zoology*. Taylor & Francis; 1977; doi: 10.1080/03014223.1977.9517934.

37. G. Theischinger, I. Endersby. Identification Guide to the Australian Odonata. Hurtsville, New South Wales, Australia: Department of Environment, Climate Change and Water NSW; 2009.

38. Tolman ER, Beatty CD, Bush J, Kohli M, Moreno CM, Ware JL, et al.. A Chromosome-length Assembly of the Black Petaltail (Tanypteryx hageni) Dragonfly. *Genome Biology and Evolution*. 2023; doi: 10.1093/gbe/evad024.

39. Newton L, Tolman E, Kohli M, Ware JL. Evolution of Odonata: Genomic insights. *Current Opinion in Insect Science*. 2023; doi: 10.1016/j.cois.2023.101073.

40. Brůna T, Li H, Guhlin J, Honsel D, Herbold S, Stanke M, et al.. Galba: genome annotation with miniprot and AUGUSTUS. *BMC Bioinformatics*. 2023; doi: 10.1186/s12859-023-05449-z.

41. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution*. 2021; doi: 10.1093/molbev/msab199.

42. Zhang C, Mirarab S. Weighting by Gene Tree Uncertainty Improves Accuracy of Quartet-based Species Trees. *Molecular Biology and Evolution*. 2022; doi: 10.1093/molbev/msac215.

43. Mendes FK, Vanderpool D, Fulton B, Hahn MW. CAFE 5 models variation in evolutionary rates among gene families. *Bioinformatics*. 2021; doi: 10.1093/bioinformatics/btaa1022.

44. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. Gibas C, editor. *PLoS ONE*. 2011; doi: 10.1371/journal.pone.0021800.

45. Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun*. Nature Publishing Group; 2020; doi: 10.1038/s41467-020-14998-3.

46. 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 Biology and Evolution*. 2021; doi: 10.1093/gbe/evab138.

47. Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics and Bioinformatics*. 2021; doi: 10.1093/nargab/lqaa108.

48. Ware JL, Beatty CD, Herrera MS, Valley S, Johnson J, Kerst C, et al.. The petaltail dragonflies (Odonata: Petaluridae): Mesozoic habitat specialists that survive to the modern day. *Journal of Biogeography*. 2014; doi: 10.1111/jbi.12273.

49. May ML. Odonata: Who They Are and What They Have Done for Us Lately: Classification and Ecosystem Services of Dragonflies. *Insects*. 2019; doi: 10.3390/insects10030062.

50. Bybee SM, Kalkman VJ, Erickson RJ, Frandsen PB, Breinholt JW, Suvorov A, et al.. Phylogeny and classification of Odonata using targeted genomics. *Molecular Phylogenetics and Evolution*. 2021; doi: 10.1016/j.ympev.2021.107115.

51. Shingate P, Ravi V, Prasad A, Tay B-H, Garg KM, Chattopadhyay B, et al.. Chromosome-level assembly of the horseshoe crab genome provides insights into its genome evolution. *Nature Communications*. Nature Publishing Group; 2020; doi: 10.1038/s41467-020-16180-1.

52. Rebora M, Salerno G, Piersanti S. Odonata perception is more than vision. In: Cordoba-Aguilar A, Beatty C, Bried J, editors. *Dragonflies and Damselflies: Model Organisms for Ecological and Evolutionary Research*. Oxford University Press;

53. Bybee SM, Futahashi R, Renoult JP, Sharkey C, Simon S, Suvorov A, et al.. Transcriptomic insights into Odonata ecology and evolution. In: Cordoba-Aguilar A, Beatty C, Bried J, editors. *Dragonflies and Damselflies: Model Organisms for Ecological and Evolutionary Research*. Oxford University Press;

54. Vollmer W, Blanot D, De Pedro MA. Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*. 2008; doi: 10.1111/j.1574-6976.2007.00094.x.

55. Teets NM, Peyton JT, Colinet H, Renault D, Kelley JL, Kawarasaki Y, et al.. Gene expression changes governing extreme dehydration tolerance in an Antarctic insect. *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences; 2012; doi: 10.1073/pnas.1218661109.

56. Sharip Z, Fauzi M. Microbial Contamination in Urban Tropical Lentic Waterbodies and Ponds along Urbanization Gradient. *Pertanika Journal of Tropical Agricultural Science*. 422019;

57. Le Gal K, Schmidt EE, Sayin VI. Cellular Redox Homeostasis. *Antioxidants (Basel)*. 2021; doi: 10.3390/antiox10091377.

58. Salmón P, Watson H, Nord A, Isaksson C. Effects of the Urban Environment on Oxidative Stress in Early Life: Insights from a Cross-fostering Experiment. *Integr Comp Biol*. 2018; doi: 10.1093/icb/icy099.

59. Lalejini A, Ferguson AJ, Grant NA, Ofria C. Adaptive Phenotypic Plasticity Stabilizes Evolution in Fluctuating Environments. *Frontiers in Ecology and Evolution*. 92021;

60. Goodman A, Tolman E, Uche-Dike R, Abbott J, Breinholt JW, Bybee S, et al.. Assessment of targeted enrichment locus capture across time and museums using odonate specimens. *Insect Systematics and Diversity*. 2023; doi: 10.1093/isd/ixad011.

61. Benjamin W. Price, E. Louise Allan. The genome sequence of the White-legged damselfly, Platycnemis pennipes (Pallas, 1771). Wellcome Open Research;

62. Price BW, Winter M, Brooks SJ. The genome sequence of the blue-tailed damselfly, Ischnura elegans (Vander Linden, 1820). *Wellcome Open Res*. 2022; doi: 10.12688/wellcomeopenres.17691.1.

63. Grether GF, Beninde J, Beraut E, Chumchim N, Escalona M, MacDonald ZG, et al.. Reference genome for the American rubyspot damselfly, Hetaerina americana. *Journal of Heredity*. 2023; doi: 10.1093/jhered/esad031.

64. Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*. 2011; doi: 10.1093/bioinformatics/btr011.

65. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods*. 2021; doi: 10.1038/s41592-020-01056-5.

66. Laetsch DR, Blaxter ML. BlobTools: Interrogation of genome assemblies. *F1000Res*. 2017; doi: 10.12688/f1000research.12232.1.

67. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018; doi: 10.1093/bioinformatics/bty191.

68. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al.. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021; doi: 10.1093/gigascience/giab008.

69. Shiryev SA, Papadopoulos JS, Schäffer AA, Agarwala R. Improved BLAST searches using longer words for protein seeding. *Bioinformatics*. 2007; doi: 10.1093/bioinformatics/btm479.

70. Trizna M. assembly_stats 0.1.4. Zenodo;

71. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al.. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009; doi: 10.1186/1471-2105-10-421.

72. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al.. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*. 2012; doi: 10.1093/nar/gkr1293.

73. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al.. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009; doi: 10.1101/gr.092759.109.

74. Armstrong J, Hickey G, Diekhans M, Fiddes IT, Novak AM, Deran A, et al.. Progressive Cactus is a multiple-genome aligner for the thousand-genome era. *Nature*. Nature Publishing Group; 2020; doi: 10.1038/s41586-020-2871-y.

75. Kolmogorov M, Raney B, Paten B, Pham S. Ragout—a reference-assisted assembly tool for bacterial genomes. *Bioinformatics*. 2014; doi: 10.1093/bioinformatics/btu280.

76. Kolmogorov M, Armstrong J, Raney BJ, Streeter I, Dunn M, Yang F, et al.. Chromosome assembly of large and complex genomes using multiple references. *Genome Res*. 2018; doi: 10.1101/gr.236273.118.

77. Kuznetsova VG, Golub NV. A checklist of chromosome numbers and a review of karyotype variation in Odonata of the world. *Comp Cytogenet*. 2020; doi: 10.3897/CompCytogen.v14i4.57062.

78. Liu H-L, Harris AJ, Wang Z-F, Chen H-F, Li Z-A, Wei X. The genome of the Paleogene relic tree Bretschneidera sinensis: insights into trade-offs in gene family evolution, demographic history, and adaptive SNPs. *DNA Research*. 2022; doi: 10.1093/dnares/dsac003.

79. Bentley BP, Carrasco-Valenzuela T, Ramos EKS, Pawar H, Souza Arantes L, Alexander A, et al.. Divergent sensory and immune gene evolution in sea turtles with contrasting demographic and life histories. *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences; 2023; doi: 10.1073/pnas.2201076120.

80. Phillips L. gnuplot Cookbook. Birmingham: Packt Publishing;

81. Liu H, Jia Y, Sun X, Tian D, Hurst LD, Yang S. Direct Determination of the Mutation Rate in the Bumblebee Reveals Evidence for Weak Recombination-Associated Mutation and an Approximate Rate Constancy in Insects. *Molecular Biology and Evolution*. 2017; doi: 10.1093/molbev/msw226.

82. Winstanley WJ, Rowe RJ. The larval habitat of Uropetala carovei carovei (Odonata: Petaluridae) in the North Island of New Zealand, and the geographical limits of the subspecies. *New Zealand Journal of Zoology*. Taylor & Francis; 1980; doi: 10.1080/03014223.1980.10423769.

83. Li H. Protein-to-genome alignment with miniprot. *Bioinformatics*. 2023; doi: 10.1093/bioinformatics/btad014.

84. Hoff KJ, Stanke M. Predicting Genes in Single Genomes with AUGUSTUS. *Curr Protoc Bioinformatics*. 2019; doi: 10.1002/cpbi.57.

85. Stanke M, Tzvetkova A, Morgenstern B. AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome. *Genome Biology*. 2006; doi: 10.1186/gb-2006-7-s1-s11.

86. Tolman ER, Beatty CD, Bush J, Kohli M, Moreno CM, Ware JL, et al.. A Chromosome-length Assembly of the Black Petaltail (Tanypteryx hageni) Dragonfly. *Genome Biology and Evolution*. 2022; doi: 10.1093/gbe/evad024.

87. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, et al.. RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences; 2020; doi: 10.1073/pnas.1921046117.

88. Uliano-Silva M, Nunes JGF, Krasheninnikova K, McCarthy SA. marcelauliano/MitoHiFi: mitohifi_v2.0. Zenodo;

89. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al.. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*. 2020; doi: 10.1093/molbev/msaa015.

90. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*. 2019; doi: 10.1186/s13059-019-1832-y.

91. . CAFE. Hahn Lab;

92. Mariani V, Biasini M, Barbato A, Schwede T. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*. 2013; doi: 10.1093/bioinformatics/btt473.

93. Hiranuma N, Park H, Baek M, Anishchenko I, Dauparas J, Baker D. Improved protein structure refinement guided by deep learning based accuracy estimation. *Nat Commun*. Nature Publishing Group; 2021; doi: 10.1038/s41467-021-21511-x.

94. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al.. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. American Association for the Advancement of Science; 2021; doi: 10.1126/science.abj8754.