Evolution of Opsin Genes in Caddisflies (Insecta: Trichoptera)

Ashlyn Powell[1*], Jacqueline Heckenhauer[2,3], Steffen U. Pauls[2,3], Blanca Ríos-Touma[4], Ryoichi B. Kuranishi[5,6], Ralph W. Holzenthal[7], Ernesto Razuri-Gonzales[3], Seth Bybee[8], and Paul B. Frandsen[1*]

Affiliations:

[1]Department of Plant and Wildlife Sciences, Brigham Young University, Provo, Utah, USA

[2]LOEWE Centre for Translational Biodiversity Genomics, Frankfurt, Germany

[3]Senckenberg Research Institute and Natural History Museum Frankfurt, Frankfurt, Germany

[4]Facultad de Ingenierías y Ciencias Aplicadas, Ingeniería Ambiental, Grupo de Investigación en Biodiversidad, Medio Ambiente y Salud, Universidad de Las Américas, Quito, Ecuador

[5]Graduate School of Science, Chiba University, Chiba, Japan

[6]Kanagawa Institute of Technology, Kanagawa, Japan

[7]Department of Entomology, University of Minnesota, St. Paul, Minnesota, USA

[8]Department of Biology, Brigham Young University, Provo, Utah, USA

***Authors for Correspondence**:

Ashlyn Powell, Department of Plant and Wildlife Sciences, Brigham Young University, Provo, UT, 84602, USA; Phone: (469) 441-5058; Email: ashlynpowell913@gmail.com

Paul B. Frandsen, Department of Plant and Wildlife Sciences, Brigham Young University, Provo, UT, 84602, USA; Phone: (801) 422-2283; Email: paul_frandsen@byu.edu

Abstract:

Insects have evolved complex and diverse visual systems controlled by light-sensing protein molecules called opsins, which couple with a chromophore to form photopigments. Insect photopigments group into three major gene families based on wavelength sensitivity: long wavelength (LW), short wavelength (SW), and ultraviolet wavelength (UV). Here, we identified 123 opsin sequences from whole genome assemblies across 25 species of caddisfly (Insecta: Trichoptera). We discovered the LW opsins have the most diversity across species and form two separate clades in the opsin gene tree. Conversely, we observed a loss of the SW opsin in half of the trichopteran species in this study, which might be associated with the fact that caddisflies are active during low-light conditions. Lastly, we found a single copy of the UV opsin in all the species in this study, with one exception: *Athripsodes cinereus* has two copies of the UV opsin and resides within a clade of caddisflies with colorful wing patterns.

Keywords: Caddisflies, evolution, opsins, visual systems, Trichoptera, insects

Significance:

While opsin evolution in some insect groups has been well-characterized, it has never been studied across caddisflies. Our findings provide insight into the diversity of opsins in caddisflies and form a basis for further research into the evolutionary drivers and complex visual systems in Trichoptera.

Introduction:

Within the visual system, the ability to perceive light is critical and plays an essential role in the life histories of insects including finding food, avoiding predators, and selecting a mate (van der Kooi et al., 2021). Light perception occurs within three different types of visual organs in insects: the stemmata of larvae and the ocelli and compound eyes of adults (Guignard et al., 2022; van der Kooi et al., 2021). Upon light absorption, photoreceptors within the eyes—composed of an opsin protein and a chromophore—change their configuration from a resting state to a signaling state, thereby indicating a physiological response (Shichida & Matsuyama, 2009). In insects, visual opsins form three major gene clades based on their peak wavelength sensitivity, including long wavelength opsins (LW, 500–600 nm), short wavelength opsins (SW, 400–500 nm), and ultraviolet wavelength opsins (UV, 300–400 nm) (Feuda et al., 2016; Lord et al., 2016, van der Kooi et al., 2021). In many insect groups, an additional opsin type, Rhodopsin 7 (RH7), is present. RH7 opsins, which are not used in color vision, do not have a known function in most insect groups but were found to be involved in circadian rhythms in *Drosophila* (Ni et al., 2017). Furthermore, in some taxa, opsins have been found to have additional non-canonical or light-independent functions (Feuda et al. 2022).

Insects typically possess one or more copies of each opsin type. Moreover, multiple cases of gene duplications and losses have been observed throughout insect opsin evolution (Briscoe et al., 2010; Feuda et al., 2016; French et al., 2015; Frentiu et al., 2007; Futahashi et al., 2015; Guignard et al., 2022; Lord et al., 2016; Macias-Muñoz et al., 2019; Mulhair et al., 2023; Sharkey et al., 2017; Sison-Mangus et al., 2006; 2008; Sondhi et al., 2021; Spaethe and Briscoe, 2004; Suvorov et al., 2017). These duplications, and subsequent diversification, of visual opsin genes are the primary mechanisms of evolution that allow for greater visual capacity and

flexibility (Frentiu et al., 2007; Suvorov et al., 2017) and are usually linked to a particular life history strategy, living environment, or light condition (Guignard et al., 2022; Sondhi et al., 2021). For example, in dragonflies (Insecta: Odonata), numerous duplication events, resulting in up to 33 visual opsin genes in some species, have been observed. This is likely related to their advanced flying capabilities, prominent eyes, and ecology (Futahashi et al., 2015). Conversely, there has been a loss of the SW opsin gene observed in the American cockroach (*Periplaneta americana*; Insecta: Dictyoptera) (French et al., 2015) and in all studied species of Neuropteroidea, likely due to having nocturnal ancestors (Lord et al., 2016; Sharkey et al., 2017; Sharkey et al., 2023). Additionally, in butterflies and moths (Insecta: Lepidoptera), diurnal species were found to possess significantly more duplications than nocturnal species. Together, these results indicate that light conditions may be drivers of opsin evolution (Sondhi et al., 2021). While the evolution of opsins has been relatively well studied in some insect orders, opsin genes in caddisflies (Insecta: Trichoptera) have only been characterized in a single species, *Limnephilus lunatus*, as part of a broader comparative study across insects (Guignard et al., 2022).

Caddisflies are a highly diverse order of aquatic insects. As eggs, larvae, and pupae, they mainly inhabit benthic habitats of freshwater rivers, streams, ponds, and lakes, but as adults, they occupy terrestrial environments adjacent to freshwater (Morse et al., 2019). There are two suborders of Trichoptera—both monophyletic groups—characterized by differences in habitat, morphology, and silk use: Annulipalpia (retreat-making) and Integripalpia (cocoon- and case-making) (Thomas et al., 2020). Adult caddisflies resemble small moths; however, the wings and bodies of most species are covered in small hairs instead of scales (Holzenthal et al., 2007). In some families of caddisflies, distinct colors and patterns on the wings can be identifying features

to distinguish different species of caddisflies, as in butterflies and moths (Holzenthal et al., 2007). While most caddisflies possess wing hairs and patterns that are somber shades of brown, yellow, or gray, a few species have brightly colored wings with red, orange, green, or silver regions (Fig. 1). The wing color in some of these species is due to the development of hairs into scales (Holzenthal et al., 2007). Additionally, adult caddisflies possess varying eye sizes, some much larger than others (Fig. 1). Presumably, such varied environments—both aquatic and terrestrial—and wing colorations require a plastic and diverse visual system. To gain an understanding of opsin evolution in caddisflies, we analyzed the occurrence and phylogenetic relationships of opsin genes from whole genome assemblies across 25 caddisfly species, including five newly assembled genomes and 20 publicly available genome assemblies (Frandsen et al., 2019; Frandsen et al., 2023; Heckenhauer et al., 2019; Heckenhauer et al., 2022; Heckenhauer et al., 2023; Kawahara et al., 2022; Luo et al., 2018; Olsen et al., 2021; Ríos-Touma et al., 2021, The Darwin Tree of Life Project Consortium, 2022), representing the major evolutionary lineages within the order.

Results:

We found a total of 123 caddisfly opsin sequences in this study. Of these, 60 were LW opsins, 12 were SW opsins, 26 were UV opsins, and 25 were RH7 opsins (Fig. 2). These opsin sequences were found across 25 species with representatives in all major lineages of Trichoptera (Fig. 3). The number of opsins found within each species ranged from three opsin paralogs to as many as nine opsin paralogs. These sequences formed four distinct clades in the maximum likelihood gene tree corresponding to the LW, SW, UV, and RH7 opsin groups (Fig. 2).

We observed slight differences in the topology between the gene tree and the species tree. Within each opsin clade, the arrangements of suborders were maintained, but occasionally interspecific relationships of opsin sequences did not match those in the established species tree (Frandsen et al., 2023). We also found that there were differences between the opsin gene trees constructed from the CDS alignment and the peptide alignment, primarily within the LW and UV opsin groups (Fig. S2). However, none of these areas of incongruence occurred across opsin classes, i.e. the sequences in each of the LW, SW, UV, and RH7 clades were consistent between the trees. In addition, we found that many of the branches that differed between data types occurred in areas of the trees with lower bootstrap support values (Fig. S2). While we present the CDS tree in Fig. 2, none of our overall interpretations are affected by incongruence between the amino acid and CDS analyses.

The LW opsins formed two distinct clades in the opsin gene tree, a phenomenon observed in other insect orders (Feuda et al., 2016; Futahashi et al., 2015; Guignard et al., 2022; Lord et al., 2016; Mulhair et al. 2023; Sharkey et al., 2017; Sharkey et al., 2023; Sondhi et al., 2021, Suvorov et al., 2017). We labeled the larger group LW1 and the smaller one LW2, with letters (e.g. a, b, c, etc.) distinguishing multiple copies within a species. All species had at least one LW opsin in the LW1 clade, while only 14 species had an additional LW2 opsin. The 14 species with the LW2 opsin were distributed across both suborders (Fig. S1). In species with multiple copies of the LW1 opsin, the paralogs were adjacent on the same contig in the genome assembly (Table S1). In addition, in all but two of the species with a LW2 opsin, both the LW1 and LW2 opsins were on the same contig (Table S1). In the other two species, *Atopsyche davidsoni* and *Stenopsyche tienmushanensis*, the LW2 opsin was on a separate contig than that of the LW1 opsins, but the sequences were full length (Table S2a).

In the opsin gene tree, we observed a few instances of clusters of opsins from the same species (Fig. 2). This was true for copies of the LW1 opsin in many species including *Stenopsyche tienmushanensis*, *Leptonema lineaticorne*, *Plectrocnemia conspersa*, *Hesperophylax magnus*, *Eubasilissa regina*, *Odontocerum albicorne*, and *Athripsodes cinereus* (Fig. 2). In addition, *Athripsodes cinereus* was the only species with two copies of the UV opsin and we found that the paralogs were paired in the opsin gene tree and adjacent to each other on the same contig in the genome assembly (Table S1). Upon testing the branches of these duplicate opsins for selection with the Branch-Site Unrestricted Statistical Test for Episodic Diversification (BUSTED, Kosakovsky Pond et al., 2020; Murrell et al., 2015), we found evidence of episodic diversifying selection for three opsin paralogs within species: Leptonema lineaticorne LW1c, Stenopsyche tienmushanensis LW1b, and Eubasilissa regina LW1b (Table S3).

We also found instances of gene duplication in the common ancestor of closely related species. For example, Clades A and B in Fig. 2 each have one or more LW1 opsin from the species *Leptonema lineaticorne*, *Hydropsyche tenuis*, *Parapsyche elsis*, *Arctopsyche grandis*, and *Plectrocnemia conspersa*, all but the latter of which belong to the family Hydropsychidae. We tested opsin sequences from each of these duplication events separately for episodic diversifying selection using BUSTED (Kosakovsky Pond et al., 2020; Murrell et al., 2015) and found evidence of selection in the sequences in Clade A but not Clade B (Fig, 2, Table S3). Similarly, multiple duplication events occurred in the LW1 opsins in the lineage leading to *Nectopsyche paramo* and *Athripsodes cinereus*, both from the family Leptoceridae (Fig. 2: Clades C, D, and E). Among these duplications, we found evidence of selection in the opsin sequences belonging to Clade E (Table S3).

The number of opsin copies in each species was highly variable across the phylogeny and between suborders. Within the seven species of fixed-retreat makers (Fig. 3: Clade 1), opsin diversity was relatively invariable; all species had one copy of the UV and RH7 opsins, two or three copies of the LW opsin, and a single SW opsin was found in all except *Plectrocnemia conspersa.* The number of opsins recovered within Integripalpia (Fig. 3: Clade 2), however, was more variable. For example, the number of LW opsins among species of Integripalpia ranged from one to five copies. Furthermore, we recovered all three visual opsins in four of the seven species within "basal Integripalpia" (Fig. 3: Clade 2-a). The other three species, all from the family Rhyacophilidae, were found to have lost the SW opsin. Conversely, in the tube-case making Integripalpia (Fig. 3: Clade 2-b), the SW opsin was lost in all but two species, *Nectopsyche paramo* and *Athripsodes cinereus,* which both belong to the family Leptoceridae.

Discussion:

We searched across 25 caddisfly genome assemblies to determine the number and phylogenetic relationships of opsins in Trichoptera. We discovered variety in the number of opsin genes across species, particularly in the LW and SW opsins. Our results suggest that caddisflies have evolved a diverse range of opsins, likely driven by life history strategies and ambient light conditions as has been hypothesized in other insect orders (Briscoe et al., 2010; Feuda et al., 2016; French et al., 2015, Futahashi et al., 2015; Guignard et al., 2022; Lord et al., 2016; Mulhair et al., 2023; Sharkey et al., 2017; Sondhi et al., 2021; Suvorov et al., 2017).

It is important to note that we relied on previous research to verify that the genes we studied correspond to insect opsins (Guignard et al., 2022). Given the conservation of gene structure and their placement in a gene tree with known opsin sequences, we have confidence

that the genes found in this study encode for opsin proteins. However, we cannot confidently say when, where, or if these genes are expressed. Further studies using transcriptomic analysis in caddisfly opsins can address this limitation.

We found that there were some incongruencies between the gene tree and the species tree, as mentioned in the Results section. Given the relatively smaller number of characters compared to the reference species tree, which was generated from genome-wide data (Frandsen et al. 2023), this is not unexpected and could be due to stochastic error from under-sampling of characters. This is also evidenced by lower bootstrap values in areas of the tree that were incongruent with the species tree (Fig. 2).

The distribution of opsins within Annulipalpia, the fixed-retreat makers, was relatively invariable. Interestingly, their ecological distributions are also less varied; most species inhabit fast-moving streams as larvae and are short-lived in the riparian zone as adults. There were three annulipalpian species, *Stenopsyche tienmushanensis*, *Leptonema lineaticorne*, and *Plectrocnemia conspersa*, with a higher number of LW opsins than the other species due to LW1 paralogs that were species-specific in the gene tree and adjacent to each other on the same contig, suggesting three independent tandem duplication events (Fig. 2). The only loss we observed in this suborder was the SW opsin gene in *Plectrocnemia conspersa (Fig. 3)*. The genome assembly of *Plectrocnemia conspersa* was high quality, indicating that this is likely a true loss (Table 1). Given that *Plectrocnemia conspersa* also had a LW duplication, we suspect more dynamic opsin evolution in this lineage compared with other retreat makers.

The SW opsin gene was also lost in Integripalpia within the basal cocoon-making family, Rhyacophilidae, and within the majority of the clade of tube-case makers (Fig. 3 Clade 2-b). Loss of the SW opsin in other groups such as the American cockroach (French et al., 2015) and

Neuropteroidea (Lord et al., 2016; Sharkey et al., 2017; Sharkey et al., 2023) is hypothesized to be associated with the low-light environments in which the ancestors of these insects lived. Caddisflies are primarily crepuscular and thus are most active during low-light conditions, which may be related to the loss of the SW opsin in Rhyacophilidae and in many tube-case making caddisflies.

Interestingly, in contrast to most species within the tube-case makers, we were able to find a copy of the SW opsin in both *Nectopsyche paramo* and *Athripsodes cinereus.* The latter was also the only species with two copies of the UV opsin and the highest number of LW opsins. The LW opsins in *Athripsodes cinereus* all occurred on the same contig and three out of five of them were clustered together in the gene tree. This indicates that some of the copies could be the result of lineage-specific tandem duplication events. This was also true for both copies of the UV opsin in *Athripsodes cinereus* as well as for LW opsins in three additional species: *Hesperophylax magnus*, *Eubasilissa regina*, and *Odontocerum albicorne*. These could be due to species-specific duplications; however, a denser taxon sampling is necessary to pinpoint the exact evolutionary timing and placement of these duplications. *Nectopsyche paramo* and *Athripsodes cinereus* both belong to the family Leptoceridae, a family with long adult antennae and many genera with intricate wing patterns and brightly colored or iridescent wing hairs and scales (Holzenthal et al., 2007). This is true for *Athripsodes cinereus*, which has wings with patches of cream-colored hairs (Fig. 1A), and *Nectopsyche paramo*, which has spotted wings (Fig. 1B). The high number of opsins and presence of the SW opsins could be related to the variety in wing coloration and patterns in some species of Leptoceridae (Fig. 1). Future work should combine more sampling within this interesting area of the caddisfly phylogeny with gene

expression and physiological data to better model the visual system of Trichoptera and to test hypotheses related to color vision and opsin diversity.

To further investigate the role of the duplication events we observed in the LW and UV opsin groups, we tested paralogs in these areas of the tree for positive selection. Using the Branch-Site Unrestricted Statistical Test for Episodic Diversification (BUSTED, Kosakovsky Pond et al., 2020; Murrell et al., 2015), we found evidence of episodic diversifying selection in LW1 opsin paralogs but did not detect evidence of selection in the UV opsin paralogs (Table S3). In each instance when a paralog was found to be under selection, the duplicate paralog was not found to be under selection, possibly suggesting a route to neofunctionalization in those copies undergoing diversifying selection. Denser taxon sampling in future work can help clarify the evolutionary timing of duplication events and the mechanisms and role of selection that we uncovered.

Opsin evolution has been more thoroughly studied in Lepidoptera, the sister order to Trichoptera (Briscoe et al., 2010; Feuda et al., 2016; Finkbeiner & Briscoe, 2021; Frentiu et al., 2007; Guignard et al., 2022; Macias-Muñoz et al., 2019; Mulhair et al., 2023; Sison-Mangus et al., 2006; 2008; Sondhi et al., 2021; Spaethe & Briscoe, 2004). Recent work has identified several instances of family- and species-specific duplications of visual opsin genes, many of which have resulted in adaptations that extend visual capacity (Frentiu et al., 2007; Mulhair et al., 2023; Sison-Mangus et al., 2006; 2008). For example, butterflies in the family Lycaenidae use both a green-shifted SW opsin and a red-shifted LW opsin to better discriminate different wavelengths (Liénard et al., 2021; Sison-Mangus et al., 2006; 2008). Similarly, species within the genus *Heliconius* have evolved an additional UV opsin through a recent duplication followed by positive selection resulting in the ability to see from the UV to the red part of the visual

spectrum as well as, in females, the capacity to differentiate different UV wavelengths (Briscoe et al., 2010; Finkbeiner & Briscoe, 2021). Here, we also found duplication events in the LW opsin group in some species of Trichoptera and one instance of a UV duplication, but we observed no evidence of duplications in the SW opsins, rather we observed loss of this opsin in many species. In Lepidoptera, it has also been found that opsin diversity and selection differ across different diel-niches. Specifically, diurnal species have more duplications and higher selection rates than nocturnal and crepuscular ones (Sondhi et al., 2021). While the majority of caddisflies are crepuscular, across the order there are species that occupy all diel-niches. Therefore, future work in caddisfly opsin evolution could focus on possible associations between diel-niche and opsin diversity and selection rates as found in Lepidoptera.

Here, we conducted the first comprehensive study of visual opsins in Trichoptera. We found that caddisflies have evolved a diverse range of opsins across species. The species with the highest diversity of opsins were derived from the group with the most colorful and intricately patterned wings, the Leptoceridae. The evolution of opsins in caddisflies also may have been driven by life-history strategies and the low light conditions during which caddisflies are active. The findings of this study provide a basis for future research on the diverse and complex visual systems in Trichoptera.

Materials and Methods:

Genome sequences for 25 species, covering all the major lineages of Trichoptera, were assessed. Among these, five were newly assembled genomes, and 20 were publicly available genome assemblies downloaded from GenBank (Benson et al., 2013). Information on genome sequencing and assembly for the five new genomes can be found in the supplementary

information (Note S1). Using the 1000 opsin sequences found by Guignard et al. (2022), we performed a tBLASTn search of opsin sequences against each caddisfly genome. BLAST hits with an e-value < 10^{-40} were kept. The resulting opsin sequences were extracted from their corresponding genomes using the Python package Pyfaidx (Shirley et al., 2015). We noticed many instances in which multiple queries of different opsin types hit to the same locations in the genomes. In this case, when extracting the sequence from the genome, the widest window from each contig was extracted, and the gene was later classified phylogenetically.

To verify the sequences we extracted, we performed gene prediction using AUGUSTUS v3.4.0 (Stanke, et al. 2006) followed by a BLASTp search against the online NCBI databases, keeping only annotations similar to other insect opsins. We then manually checked the annotations, using outgroup sequences as a guide, to ensure that the entire gene was correctly annotated. We included opsin sequences from a variety of insect orders for outgroup comparison: Lepidoptera (*Danaus plexippus*, *Heliconius melpomene*, *Manduca sexta*, *Papilio xuthus*), Odonata (*Sympetrum frequens)*, Diptera (*Drosophila melanogaster*), and Coleoptera (*Tribolium castaneum*), all accessed through GenBank (Benson et al., 2013).

To manually verify the annotation for each opsin sequence, we first aligned the translation—from the AUGUSTUS annotation—with the outgroup sequences of the same opsin type using MUSCLE (Edgar, 2004). In instances where the caddisfly opsin sequence did not fully align with the outgroup sequences, the annotation was manually modified in Geneious Prime v2023.0.4 (https://www.geneious.com) until the entire gene sequence aligned with the outgroups. This was done by first removing or shortening CDS regions that did not align with or varied greatly from the outgroup sequences, causing large gaps in the alignment. Then, for regions of the gene that were present in the outgroup sequences but missing from the annotation,

we used the search tool in Geneious to find conserved motifs from within those regions and added or extended CDS annotations accordingly. When extending and shortening the CDS regions of the annotation, we ensured that every intron region occurred at canonical splice sites. After we made these manual adjustments to the annotation, we re-aligned the translation with the outgroup sequences, and the alignment was reassessed for completeness, i.e. that all exons and core motif regions were present in the alignment for each sequence from each species. If additional changes were necessary, the same process of adjusting the annotation was followed until the gene aligned with the outgroup sequences.

To further verify the absence of the SW and LW2 opsins in many species, we performed an additional tBLASTn search against the genomes using the orthologous sequences that were successfully obtained from the other caddisfly species. We maintained BLAST hits that were in regions of the genomes we had not previously checked, and we followed the same steps for annotating the gene as described above.

We provide supplementary tables (Table S2a,b,c) with information on the completeness of each visual opsin sequence. We used the number and identity of exons in Lepidoptera as a guide in making the tables, i.e. the column labels “Exon 1, Exon 2, …” refer to the exons present in Lepidoptera. Within the SW and UV opsin groups, the first exon in the lepidopteran outgroup sequences was not highly conserved, making it difficult to recover this exon in the caddisfly opsins despite searching for it both in Geneious and with BLAST. Therefore, these sequences are labeled as partial sequences in the opsin gene trees. In addition, there were a few instances where two exons were combined into a single exon in the caddisfly sequences. For example, the fifth and sixth exons in the lepidopteran UV opsin formed a single continuous exon, with no intron, in the caddisfly UV opsin.

Opsin Gene Tree Reconstruction

To determine phylogenetic relationships among opsin sequences, we first aligned the caddisfly and outgroup opsin peptide sequences using MAFFT v7.487 (Katoh & Standley, 2013) and created a codon alignment with PAL2NAL v14.1 (Suyama, 2009). We performed phylogenetic reconstruction on both the CDS and peptide alignments. To do this, we first selected the best substitution model using ModelFinder (Kalyaanamoorthy et al., 2017) as implemented in IQ-Tree v2.0.3 (Minh et al., 2020). We then performed a maximum likelihood tree search with 1000 UltraFast bootstrap replicates corrected with the bootstrap nearest neighbor interchange option enabled to guard against the risk of overestimating bootstrap support (-bb 1000 -bnni). We viewed the resulting trees in FigTree v1.4.4 (Rambaut, 2018). To highlight differences between the CDS and peptide trees, we created a face-to-face comparison in R with ggtree v3.2.1 (Yu et al., 2016) and ggplot v3.3.5 (Wickham, 2016) (Fig. S2). To classify which opsin group each sequence belonged to, we labeled them based on their phylogenetic placement in relation to the outgroup sequences, which was consistent between the two trees (Fig S2).

Species Tree and Bar Plot

We created the species phylogeny using the relationships inferred by Frandsen et al. (2023). To summarize the number of opsins in each species, we wrote a custom R script that calculated the totals and plotted them using ggplot2 v3.3.5 (Wickham, 2016).

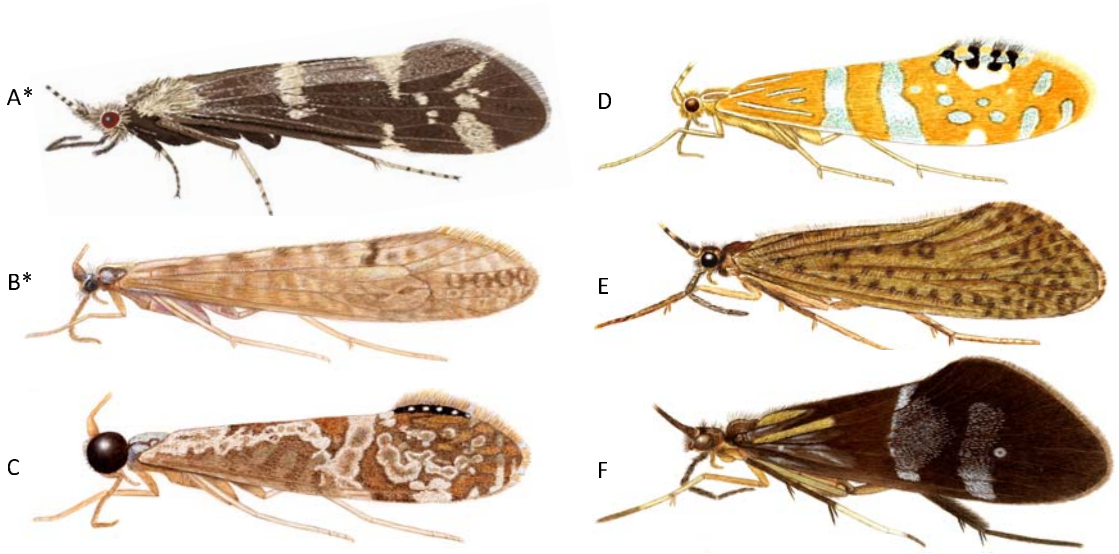
Selection Analysis

We performed tests for selection to assess the duplications in the LW and UV opsin groups. We created separate codon alignments for both the LW and UV opsin groups with

MAFFT v7.487 (Katoh & Standley, 2013) and PAL2NAL v14.1 (Suyama, 2009). We tested for episodic diversifying selection in areas of the tree where duplications had occurred using BUSTED as implemented in HyPhy (Kosakovsky Pond et al., 2020; Murrell et al., 2015). We tested branches of species-specific duplicated opsins individually as well as five deeper duplication events (Fig. 2: Clades A-E) and reported the resulting p-values in Table S3.

Figure 1

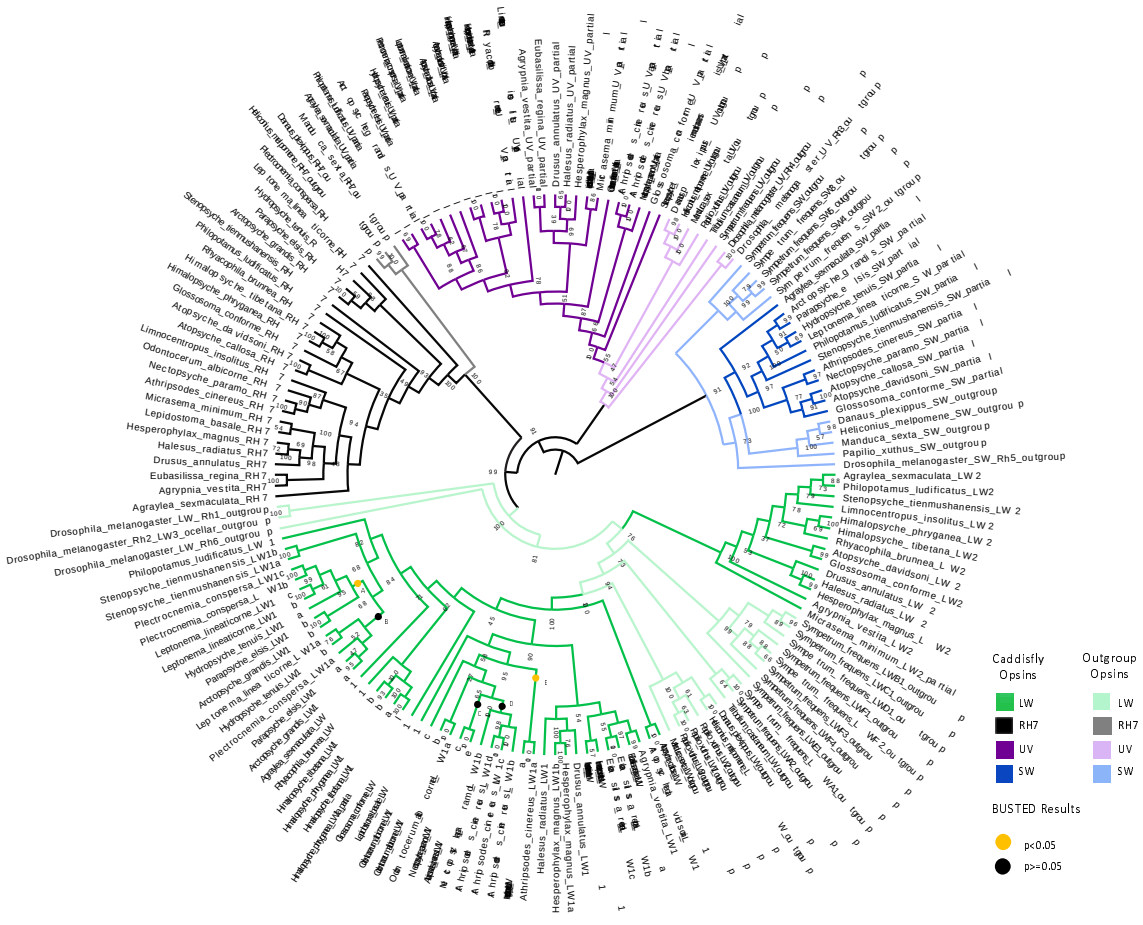
Adult Caddisfly Illustrations



Note: Adult caddisfly illustrations showing varying eye sizes and diverse wing colorations and patterns. (A) *Athripsodes cinereus* (Leptoceridae); (B) *Nectopsyche paramo* (Leptoceridae); (C) *Nectopsyche nigricapilla* (Leptoceridae); (D) *Nectopsyche ortizi* (Leptoceridae); (E) *Banyallarga vicaria* (Calamoceratidae); (F) *Phylloicus abdominalis* (Calamoceratidae). (*) Species included in this study. Illustrations by Julie Martinez and Ralph Holzenthal.

Figure 2

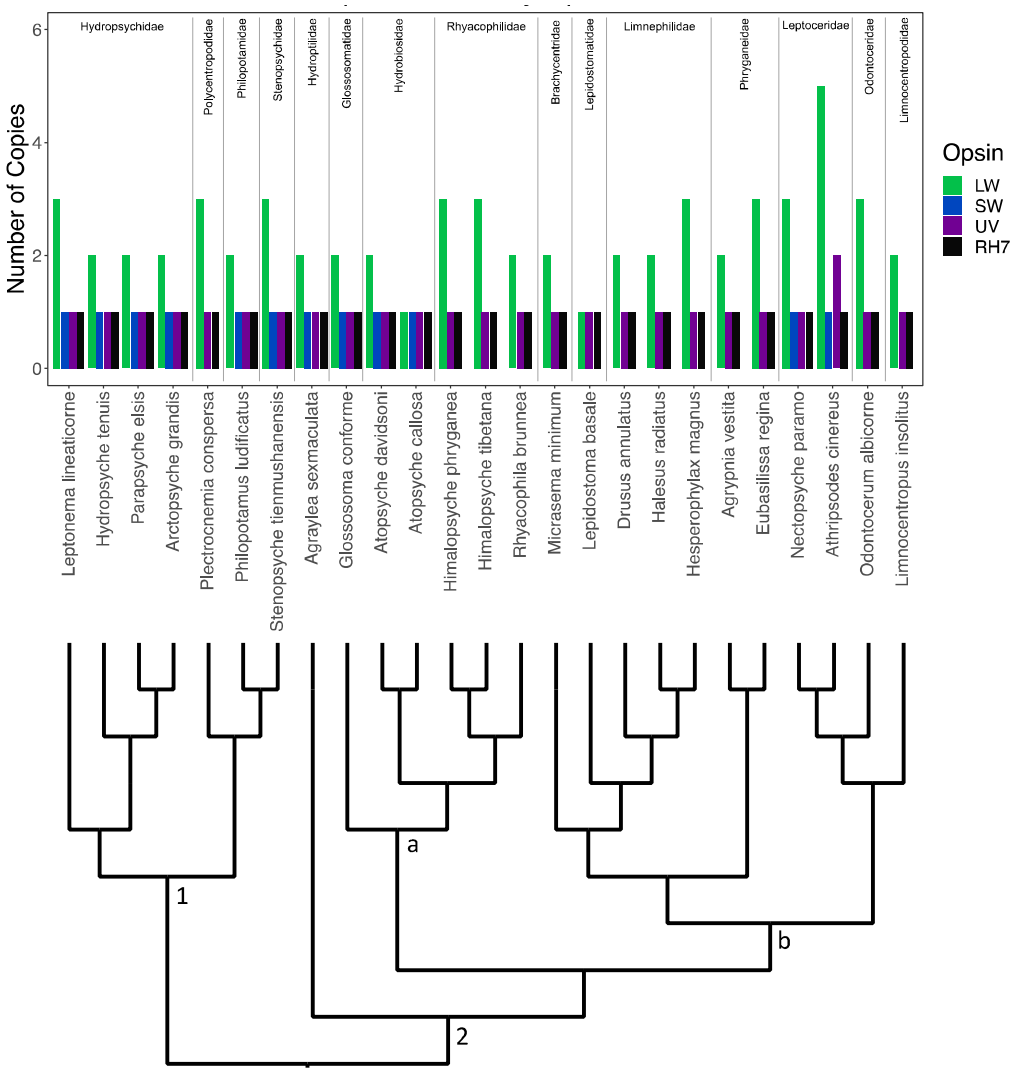
Opsin Gene Tree



Note: Maximum likelihood tree from caddisfly opsin DNA sequences found in this study along with various outgroup sequences. Nodes are labeled with bootstrap values and branches are colored by opsin type, with caddisfly opsins shown in darker colors and outgroup opsins depicted in lighter shades. Letters at the end of the node labels (e.g., LW1a, LW1b, etc.) denote multiple copies of that opsin type in a species. Clades marked with a dot (A-E) were tested for episodic diversifying selection with BUSTED (Table S3).

Figure 3

O



Note: Bar plot of the number of opsin copies in each caddisfly genome. The bars are colored by opsin type. The species are ordered by the species phylogeny indicated below the bar plot, which is based on a recent study that examined the caddisfly phylogeny in depth (Frandsen et al. 2023). Families are labeled above the bar plot. Suborders are labeled within the phylogeny as follows: (1) Annulipalpia—retreat-makers; (2) Integripalpia—cocoon- and tube case-makers; (a) (and *Agraylea sexmaculata*) basal Integripalpia—cocoon-makers; (b) tube case making Integripalpia.

Table 1

Genome Assemblies

Species	Accession number	Publication	Contig N50 (bp)	Compleasm (BUSCO)
Leptonema lineaticorne	GCA_024500535.1	Heckenhauer et al. 2023	14,931,587	C:98.30%[S:97.83%,D:0.47%], F:0.66%,I:0.00%,M:1.04%
Hydropsyche tenuis	GCA_009617725.1	Heckenhauer et al. 2019	2,190,134	C:97.93%[S:97.65%,D:0.28%], F:0.89%,I:0.00%,M:1.18%
Parapsyche elsis	GCA_022651745.1	Frandsen et al. 2019, Heckenhauer et al. 2022	5,591,679	C:96.99%[S:96.80%,D:0.19%], F:0.80%,I:0.00%,M:2.21%
Arctopsyche grandis	GCA_029955255.1	Frandsen et al. 2023	6,470,670	C:98.36%[S:95.39%,D:2.97%], F:0.71%,I:0.00%,M:0.94%
Plectrocnemia conspersa	Available on FigShare	New	32,103,979	C:97.74%[S:97.32%,D:0.42%], F:0.61%,I:0.00%,M:1.65%
Philopotamus ludificatus	GCA_022495035.1	Heckenhauer et al. 2022	35,449	C:94.63%[S:93.31%,D:1.32%], F:3.15%,I:0.05%,M:2.17%
Stenopsyche tienmushanensis	GCA_008973525.1	Luo et al. 2018	1,296,863	C:97.50%[S:95.15%,D:2.35%], F:1.04%,I:0.00%,M:1.46%
Agraylea sexmaculata	GCA_022606485.1	Heckenhauer et al. 2022	86,524	C:95.67%[S:90.96%,D:4.71%], F:1.04%,I:0.00%,M:3.30%
Glossosoma conforme	GCA_022606575.1	Heckenhauer et al. 2022	2,212,131	C:93.13%[S:92.47%,D:0.66%], F:0.66%,I:0.00%,M:6.21%
Atopsyche davidsoni	GCA_022113835.1	Ríos-Touma et al. 2021	14,095,054	C:98.68%[S:98.35%,D:0.33%], F:0.61%,I:0.00%,M:0.71%
Atopsyche callosa	Available on FigShare	New	25,586,909	C:98.54%[S:96.70%,D:1.84%], F:0.66%,I:0.00%,M:0.80%
Himalopsyche phryganea	GCA_022494535.1	Heckenhauer et al. 2019	4,634,010	C:98.21%[S:97.83%,D:0.38%], F:0.71%,I:0.00%,M:1.08%
Himalopsyche tibetana	GCA_030503985.1	Heckenhauer et al. 2022	28,889,006	C:98.91%[S:98.35%,D:0.56%], F:0.56%,I:0.00%,M:0.52%
Rhyacophila brunnea	Available on FigShare	New	1,306,779	C:97.74%[S:94.35%,D:3.39%], F:1.22%,I:0.00%,M:1.04%
Micrasema minimum	GCA_022494985.1	Heckenhauer et al. 2019	69,526	C:66.34%[S:66.20%,D:0.14%], F:7.25%,I:0.00%,M:26.41%
Lepidostoma basale	GCA_022606425.2	Heckenhauer et al. 2019	1,001,566	C:98.16%[S:97.55%,D:0.61%], F:0.85%,I:0.05%,M:0.94%
Drusus annulatus	GCA_022651775.1	Heckenhauer et al. 2019	1,032,046	C:97.92%[S:97.50%,D:0.42%], F:1.08%,I:0.00%,M:0.99%
Halesus radiatus	GCA_022606495.2	Heckenhauer et al. 2019	124,280	C:91.20%[S:89.69%,D:1.51%], F:4.38%,I:0.05%,M:4.38%
Hesperophylax magnus	GCA_026573805.1	Frandsen et al. 2023	11,205,906	C:98.68%[S:96.56%,D:2.12%], F:0.61%,I:0.00%,M:0.71%
Agrypnia vestita	GCA_016648135.1	Olsen et al. 2021	111,757	C:91.25%[S:83.10%,D:8.15%], F:4.14%,I:0.09%,M:4.52%
Eubasilissa regina	GCA_022840565.1	Kawahara et al. 2022	29,378,647	C:98.73%[S:96.28%,D:2.45%], F:0.80%,I:0.00%,M:0.47%

Nectopsyche paramo	JAWQED000000000	New	1,090,281	C:97.31%[S:95.24%,D:2.07%], F:0.80%,I:0.00%,M:1.88%
Athripsodes cinereus	GCA_947579605.1	Darwin Tree of Life Project	948,465	C:97.22%[S:96.61%,D:0.61%], F:1.08%,I:0.00%,M:1.69%
Odontocerum albicorne	GCA_949825065.1	Darwin Tree of Life Project	387,033	C:97.22%[S:96.75%,D:0.47%], F:1.18%,I:0.00%,M:1.60%
Limnocentropus insolitus	Available on FigShare	New	33,230,923	C:98.87%[S:98.26%,D:0.61%], F:0.56%,I:0.00%,M:0.56%

Note: Locations and quality scores of genome assemblies for each species. Compleasm (BUSCO) was run with the Endopterygota OrthoDB v10 gene set and the results categories are as follows: C: complete; S: single; D: duplicated; F: fragmented subclass 1 (only a portion of the gene is present in the assembly); I: fragmented subclass 2 (different sections of the gene align to different locations in the assembly); M: missing.

Acknowledgements:

We thank the BYU Office of Research Computing for the compute resources that we used in this study. We thank the BYU College of Life Sciences for providing funding through the College Undergraduate Research Award Program. We thank Ed Wilcox from the BYU DNA Sequencing Center for sequencing the new genomes presented here. J.H. acknowledges funding from the Deutsche Forschungsgemeinschaft, Project number 502865717, B.R.-T., S.U.P., R.W.H., and P.B.F. acknowledge funding from the Dirección General de Investigación, Universidad de Las Américas (Ecuador): “Montane freshwater diversity, from taxonomy to functional genomics, an approximation from Trichoptera-Part II.” (AMB.BRT.23.02).

Authors contributions:

Conceptualization: A.P. and P.B.F. Formal analysis: A.P., J.H., E.R.G., and P.B.F. Resources: J.H., B.R.-T., R.B.K., and P.B.F. Data curation: A.P., J.H., and E.R.G. Writing-original draft

prep: A.P. Writing-review and editing: all authors. Visualization: A.P., and R.W.H. Supervision: P.B.F. Funding acquisition: A.P., J.H., S.U.P., B.R.-T., and P.B.F.

References:

Arikawa K, Wakakuwa M, Qiu X, Kurasawa M, Stavenga DG. 2005. Sexual dimorphism of short-wavelength photoreceptors in the small white butterfly, Pieris rapae crucivora. *Journal of Neuroscience*. 25(25):5935-5942. https://doi.org/10.1523/JNEUROSCI.1364-05.2005.

Benson DA et al. 2013. GenBank. *Nucleic Acids Research*. 41(D1):D36–D42. https://doi.org/10.1093/nar/gks1195.

Briscoe AD et al. 2010. Positive selection of a duplicated UV-sensitive visual pigment coincides with wing pigment evolution in *Heliconius* butterflies. *Proceedings of the National Academy of Sciences*. 107(8):3628–3633. https://doi.org/10.1073/pnas.0910085107.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 32(5):1792–1797. https://doi.org/10.1093/nar/gkh340.

Feuda R, Marletaz F, Bentley MA, Holland PW. 2016. Conservation, duplication, and divergence of five opsin genes in insect evolution. *Genome Biology and Evolution.* 8(3):579-587. https://doi.org/10.1093/gbe/evw015.

Feuda R, Menon AK, Göpfert MC. 2022. Rethinking opsins. *Molecular Biology and Evolution.* 39(3):msac033. https://doi.org/10.1093/molbev/msac033.

Finkbeiner SD, Briscoe AD. 2021. True UV color vision in a female butterfly with two UV opsins. *Journal of Experimental Biology*. 224(18):jeb242802. https://doi.org/10.1242/jeb.242802.

Frandsen PB et al. 2019. Exploring the underwater silken architectures of caddisworms: comparative silkomics across two caddisfly suborders. *Philosophical Transactions of the Royal Society B: Biological Sciences.* 374(1784):20190206. https://doi.org/10.1098/rstb.2019.0206.

Frandsen PB et al. 2023. Allelic resolution of insect and spider silk genes reveals hidden genetic diversity. *Proceedings of the National Academy of Sciences.* 120(18):e2221528120. https://doi.org/10.1073/pnas.2221528120.

Frandsen PB et al. 2023. Phylogenomics recovers multiple origins of portable case-making in caddisflies (Insecta: Trichoptera), the world's most common underwater architects. *bioRxiv*. https://doi.org/10.1101/2023.12.21.572910.

French AS, Meisner S, Liu H, Weckström M, Torkkeli PH. 2015. Transcriptome analysis and RNA interference of cockroach phototransduction indicate three opsins and suggest a major role for TRPL channels. *Frontiers in Physiology*. 6:207. https://doi.org/10.3389/fphys.2015.00207.

Frentiu FD, Bernard GD, Sison-Mangus MP, Brower AVZ, Briscoe AD. 2007. Gene duplication is an evolutionary mechanism for expanding spectral diversity in the long-wavelength photopigments of butterflies. *Molecular Biology and Evolution*. 24(9):2016–2028. https://doi.org/10.1093/molbev/msm132.

Futahashi R et al. 2015. Extraordinary diversity of visual opsin genes in dragonflies. *Proceedings of the National Academy of Sciences.* 112(11):E1247-E1256. https://doi.org/10.1073/pnas.1424670112.

Guignard Q, Allison JD, Slippers B. 2022. The evolution of insect visual opsin genes with specific consideration of the influence of ocelli and life history traits. *BMC Ecology and Evolution*. 22(2). https://doi.org/10.1186/s12862-022-01960-8.

Heckenhauer J et al. 2019. Annotated draft genomes of two caddisfly species *Plectrocnemia conspersa* CURTIS and *Hydropsyche tenuis* NAVAS (Insecta: Trichoptera). *Genome Biology and Evolution.* 11(12):3445-3451. https://doi.org/10.1093/gbe/evz264.

Heckenhauer J et al. 2022. Genome size evolution in the diverse insect order Trichoptera. *GigaScience*. 11:giac011. https://doi.org/10.1093/gigascience/giac011.

Heckenhauer J et al. 2023. Characterization of the primary structure of the major silk gene, *h-fibroin*, across caddisfly (Trichoptera) suborders. *iScience*. 26(8):107253. https://doi.org/10.1016/j.isci.2023.107253.

Holzenthal R, Blahnik R, Prather A, Kjer K. 2007. Order Trichoptera kirby, 1813 (Insecta), caddisflies. *Zootaxa*. 1668:639-698. https://doi.org/10.11646/zootaxa.1668.1.29.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*. 14:587–589. https://doi.org/10.1038/nmeth.4285.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*. 30(4):772-780. https://doi.org/10.1093/molbev/mst010.

Kawahara AY et al. 2022. Long-read HiFi sequencing correctly assembles repetitive *heavy fibroin* silk genes in new moth and caddisfly genomes. *Gigabyte*. gigabyte64. https://doi.org/10.46471/gigabyte.64.

Kosakovsky Pond SL et al. 2020. Hyphy 2.5—a customizable platform for evolutionary hypothesis testing using phylogenies. *Molecular Biology and Evolution*. 37(1):295–299. https://doi.org/10.1093/molbev/msz197.

Liénard MA et al. 2021. The evolution of red color vision is linked to coordinated rhodopsin tuning in lycaenid butterflies. *Proceedings of the National Academy of Sciences*. 118(6):e2008986118. https://doi.org/10.1073/pnas.2008986118.

Lord NP et al. 2016. A cure for the blues: Opsin duplication and subfunctionalization for short-wavelength sensitivity in jewel beetles (Coleoptera: Buprestidae). *BMC Evolutionary Biology*. 16:107. https://doi.org/10.1186/s12862-016-0674-4.

Luo S, Tang M, Frandsen PB, Stewart RJ, Zhou X. 2018. The genome of an underwater architect, the caddisfly *Stenopsyche tienmushanensis* Hwang (Insecta: Trichoptera). *GigaScience*. 7(12):giy143. https://doi.org/10.1093/gigascience/giy143

Macias-Muñoz A, Olguin AGR, Briscoe AD. 2019. Evolution of phototransduction genes in Lepidoptera. *Genome Biology and Evolution*. 11(8):2107–2124 https://doi.org/10.1093/gbe/evz150.

Minh BQ et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*. 37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015.

Morse JC, Frandsen PB, Graf W, Thomas JA. 2019. Diversity and ecosystem services of Trichoptera. *Insects*. 10(5):125. https://doi.org/10.3390/insects10050125.

Mulhair PO, Crowley L, Boyes DH, Lewis OT, Holland PWH. 2023. Opsin gene duplication in Lepidoptera: retrotransposition, sex linkage, and gene expression. *Molecular Biology and Evolution*. 40(11):msad241. https://doi.org/10.1093/molbev/msad241.

Murrell B et al. 2015. Gene-wide identification of episodic selection. *Molecular Biology and Evolution*. 32(5):1365–1371. https://doi.org/10.1093/molbev/msv035.

Ni JD, Baik LS, Holmes TC, Montell C. 2017. A rhodopsin in the brain functions in circadian photoentrainment in *Drosophila*. *Nature*. 545:340–344. https://doi.org/10.1038/nature22325.

Olsen LK et al. 2021. Draft genome assemblies and annotations of *Agrypnia vestita* Walker, and *Hesperophylax magnus* Banks reveal substantial repetitive element expansion in tube case-making caddisflies (Insecta: Trichoptera). *Genome Biology and Evolution.* 13(3):evab013. https://doi.org/10.1093/gbe/evab013.

Rambaut, A. (2018). FigTree (Version 1.4.4) [Computer software]. http://tree.bio.ed.ac.uk/software/figtree.

Ríos-Touma B et al. 2021. De novo genome assembly and annotation of an Andean caddisfly, *Atopsyche davidsoni* Sykora, 1991, a model for genome research of high-elevation adaptations. *Genome Biology and Evolution.* 14(1):evab286. https://doi.org/10.1093/gbe/evab286.

Sharkey CR et al. 2017. Overcoming the loss of blue sensitivity through opsin duplication in the largest animal group, beetles. *Scientific Reports.* 7:8. https://doi.org/10.1038/s41598-017-00061-7

Sharkey CR, Blanco J, Lord NP, Wardill TJ. 2023. Jewel beetle opsin duplication and divergence is the mechanism for diverse spectral sensitivities. *Molecular Biology and Evolution*. 40(2):msad023. https://doi.org/10.1093/molbev/msad023.

Shichida Y, Matsuyama T. 2009. Evolution of opsins and phototransduction. *Philosophical Transactions of the Royal Society B: Biological Sciences.* 364(1531):2881–2895. https://doi.org/10.1098/rstb.2009.0051.

Shirley M, Ma Z, Pedersen B, Wheelan S. 2015. Efficient "pythonic" access to FASTA files using pyfaidx. *PeerJ PrePrints.* 3:e970v1.

Sison-Mangus MP, Bernard GD, Lampel J, Briscoe AD. 2006. Beauty in the eye of the beholder: the two blue opsins of lycaenid butterflies and the opsin gene-driven evolution of sexually dimorphic eyes. *Journal of Experimental Biology*. 209(16):3079–3090. https://doi.org/10.1242/jeb.02360.

Sison-Mangus MP, Briscoe AD, Zaccardi G, Knu□ttel H, Kelber A. 2008 The lycaenid butterfly Polyommatus icarus uses a duplicated blue opsin to see green. *Journal of Experimental Biology*. 211(3):361–369. https://doi.org/10.1242/jeb.012617.

Sondhi Y, Ellis EA, Bybee SM, Theobald JC, Kawahara AY. 2021. Light environment drives evolution of color vision genes in butterflies and moths. *Communications Biology*. 4: 177. https://doi.org/10.1038/s42003-021-01688-z.

Spaethe J, Briscoe AD. 2004. Early duplication and functional diversification of the opsin gene family in insects. *Molecular Biology and Evolut*ion. 21(8):1583–1594. https://doi.org/10.1093/molbev/msh162.

Stanke M et al. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Research*. 34:W435–W439. https://doi.org/10.1093/nar/gkl200.

Suvorov A et al. 2017. Opsins have evolved under the permanent heterozygote model: Insights from phylotranscriptomics of Odonata. *Molecular Ecology*. 26(5):1306-1322. https://doi.org/10.1111/mec.13884.

Suyama M, Torrents D, Bork P. 2009. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*. 34:W609–W612. https://doi.org/10.1093/nar/gkl315.

van der Kooi CJ, Stavenga DG, Arikawa K, Belusic G, Kelber A. 2021. Evolution of insect color vision: from spectral sensitivity to visual ecology. *Annual Review of Entomology*. 66: 435-461. https://doi.org/10.1146/annurev-ento-061720-071644.

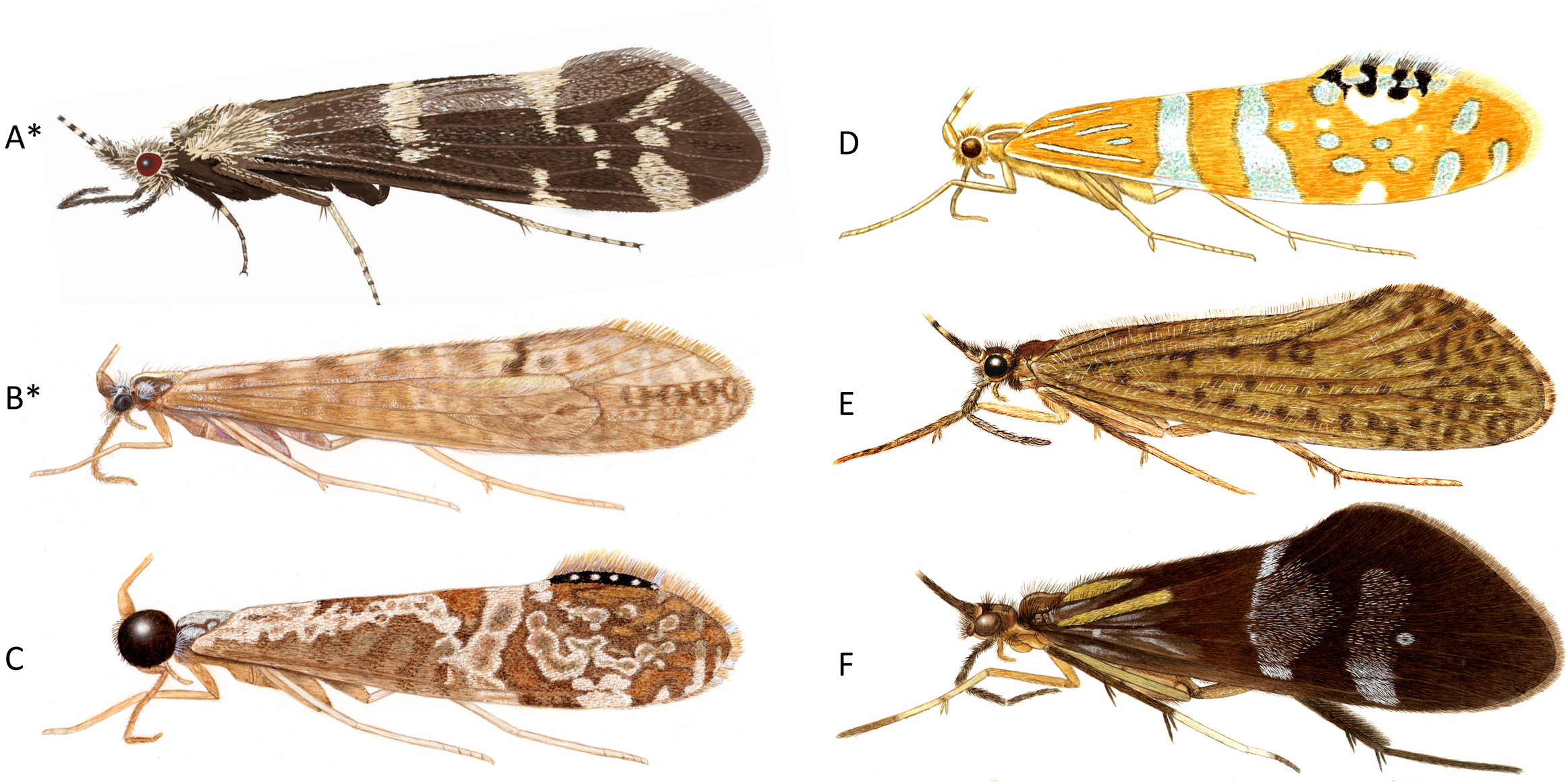
The Darwin Tree of Life Project Consortium. 2022. Sequence locally, think globally: The Darwin Tree of Life Project. *Proceedings of the National Academy of Sciences.* 119(4): e2115642118. https://doi.org/10.1073/pnas.2115642118.

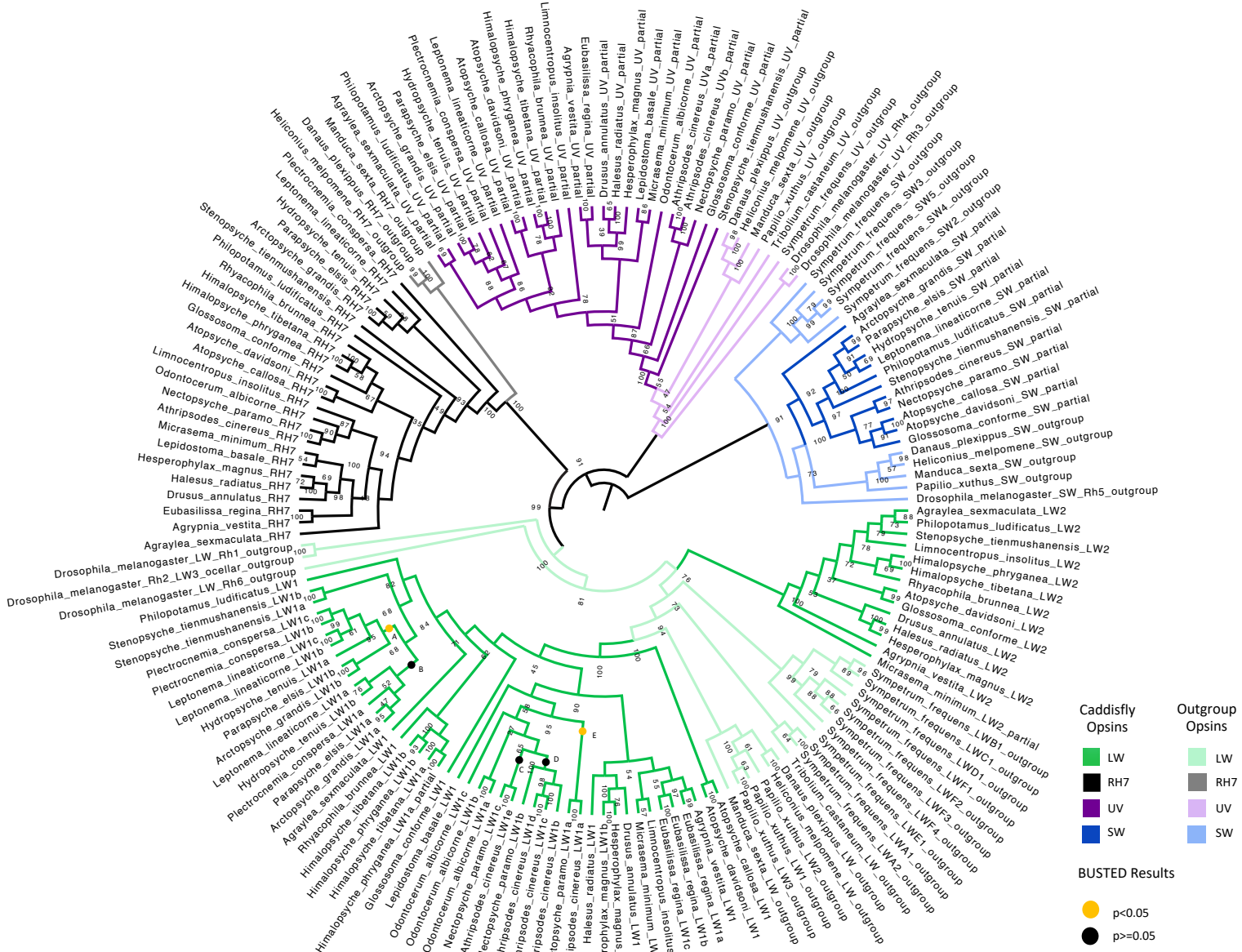
Thomas JA, Frandsen PB, Prendini E, Zhou X, Holzenthal RW. 2020. A multigene phylogeny and timeline for Trichoptera (Insecta). *Systematic Entomology*. 45(3):670-686. https://doi.org/10.1111/syen.12422.

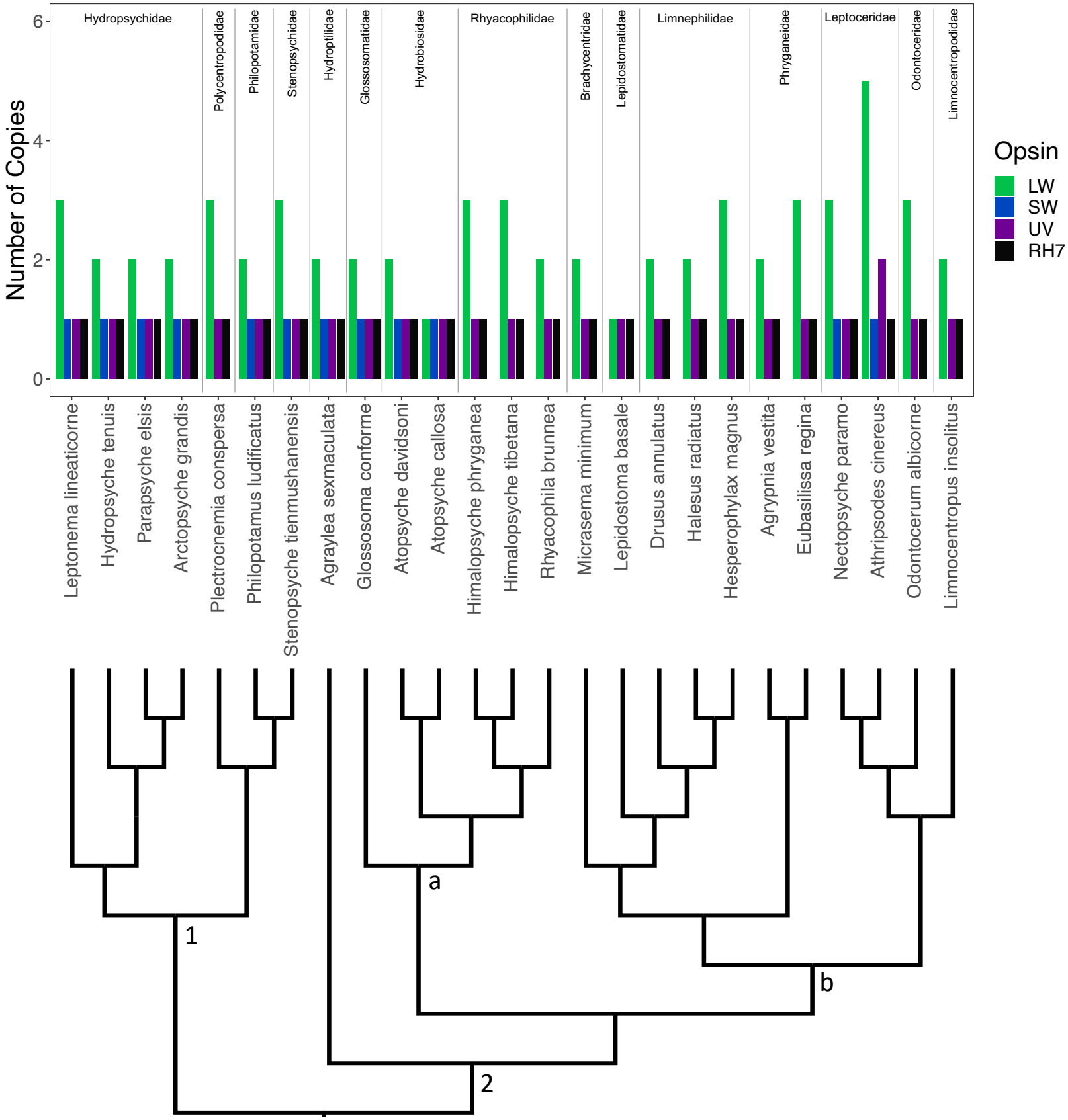
Wickham H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org.

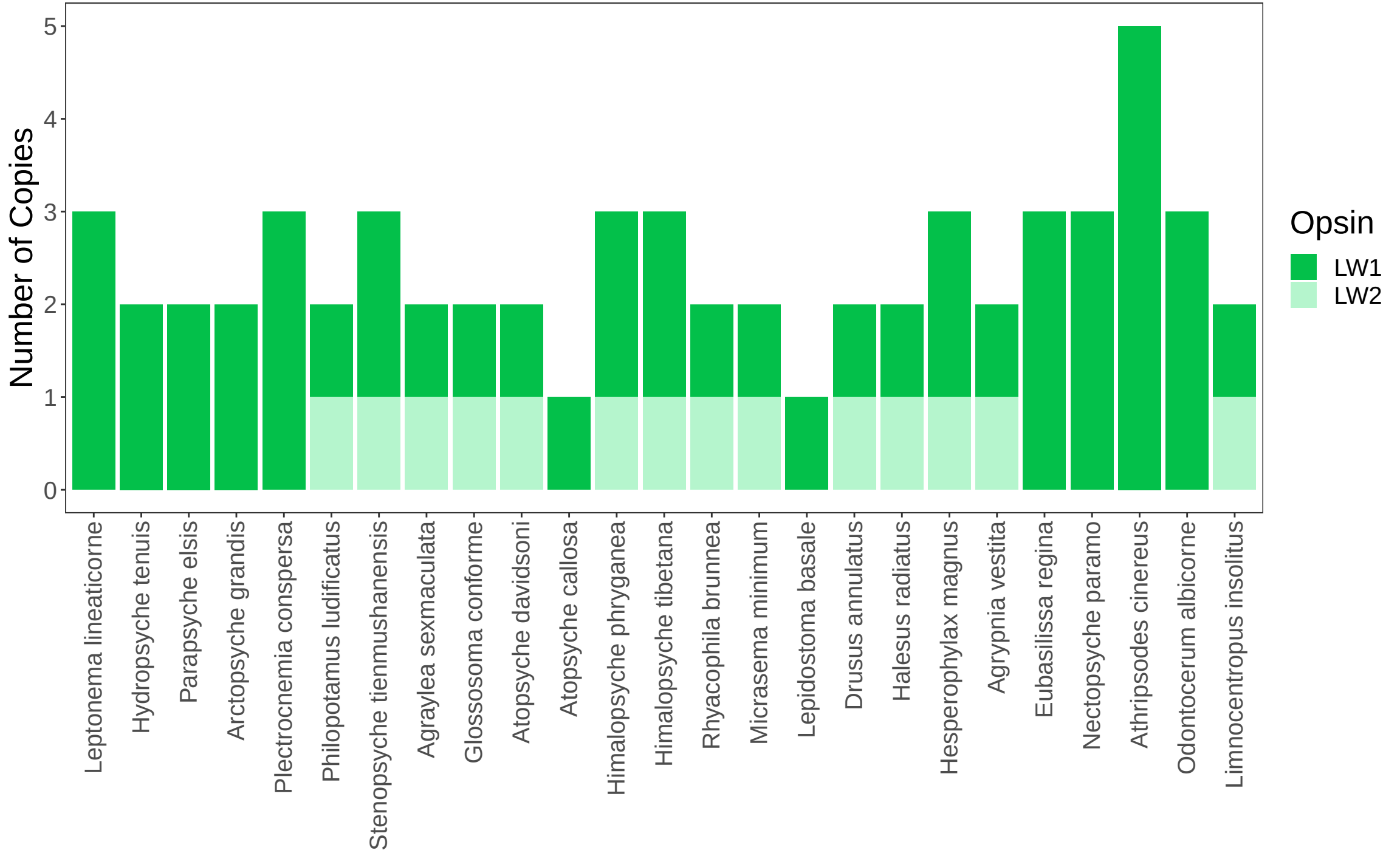
Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. 2016. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*. 8(1):28-36. https://doi.org/10.1111/2041-210X.12628.

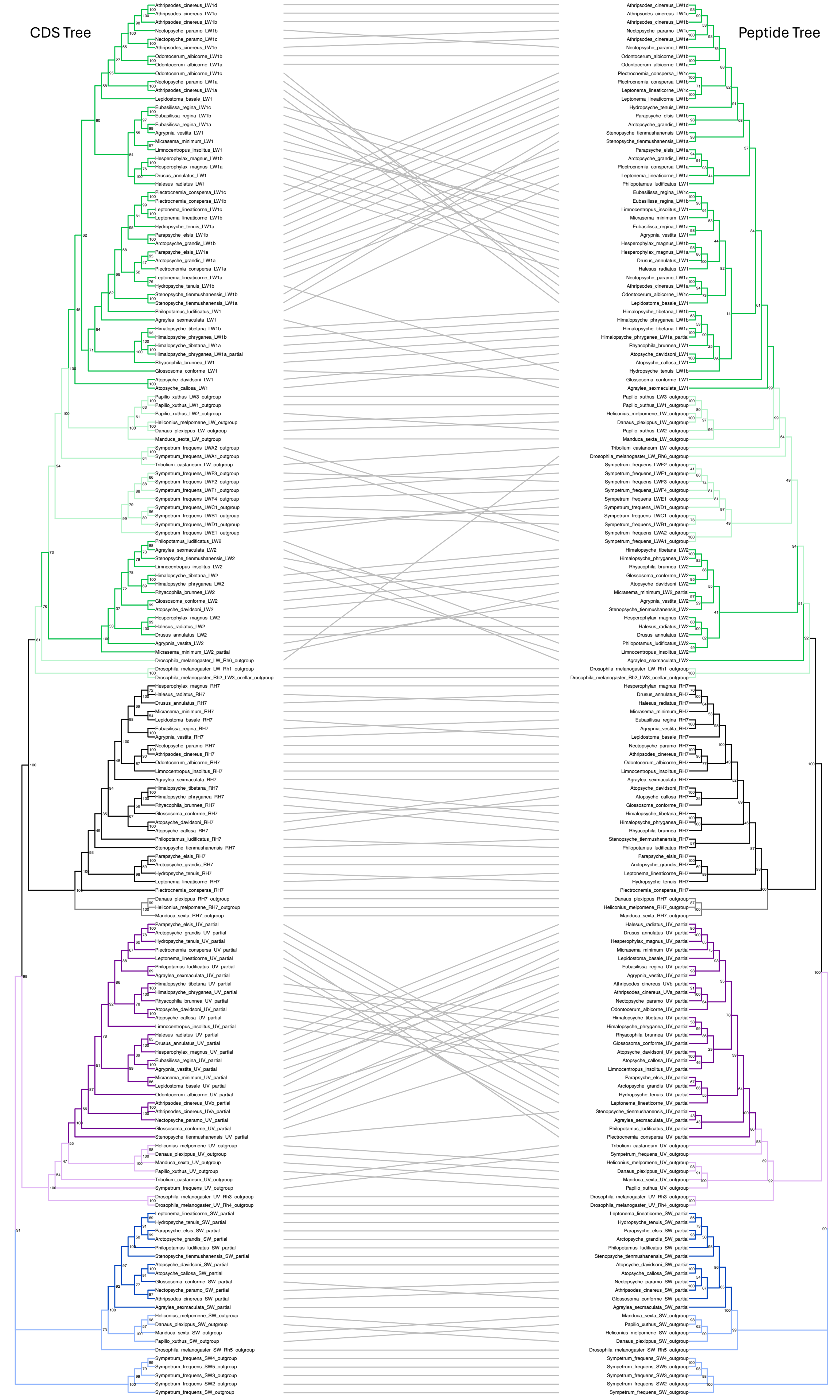
Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19(6):153. https://doi.org/10.1186/s12859-018-2129-y.


A*
B*
C
D
E
F


Caddisfly Opsins
LW
RH7
UV
SW
Outgroup Opsins
LW
RH7
UV
SW
BUSTED Results
p<0.05
p>=0.05


Number of Copies
0
2
4
6
Hydropsychidae
Polycentropodidae
Philopotamidae
Stenopsychidae
Hydroptilidae
Glossosomatidae
Hydrobiosidae
Rhyacophilidae
Brachycentridae
Lepidostomatidae
Limnephilidae
Phryganeidae
Leptoceridae
Odontoceridae
Limnocentropodidae
Opsin
LW
SW
UV
RH7
Leptonema lineaticorne
Hydropsyche tenuis
Parapsyche elsis
Arctopsyche grandis
Plectrocnemia conspersa
Philopotamus ludificatus
Stenopsyche tienmushanensis
Agraylea sexmaculata
Glossosoma conforme
Atopsyche davidsoni
Atopsyche callosa
Himalopsyche phryganea
Himalopsyche tibetana
Rhyacophila brunnea
Micrasema minimum
Lepidostoma basale
Drusus annulatus
Halesus radiatus
Hesperophylax magnus
Agrypnia vestita
Eubasilissa regina
Nectopsyche paramo
Athripsodes cinereus
Odontocerum albicorne
Limnocentropus insolitus
1
2
a
b


Number of Copies
0
1
2
3
4
5
Leptonema lineaticorne
Hydropsyche tenuis
Parapsyche elsis
Arctopsyche grandis
Plectrocnemia conspersa
Philopotamus ludificatus
Stenopsyche tienmushanensis
Agraylea sexmaculata
Glossosoma conforme
Atopsyche davidsoni
Atopsyche callosa
Himalopsyche phryganea
Himalopsyche tibetana
Rhyacophila brunnea
Micrasema minimum
Lepidostoma basale
Drusus annulatus
Halesus radiatus
Hesperophylax magnus
Agrypnia vestita
Eubasilissa regina
Nectopsyche paramo
Athripsodes cinereus
Odontocerum albicorne
Limnocentropus insolitus
Opsin
LW1
LW2


CDS Tree
Peptide Tree