Ecology and Evolution

WILEY

Ecology and Evolution Open Access

RESEARCH ARTICLE OPEN ACCESS

Patterns of Gene Family Evolution and Selection Across *Daphnia*

Connor S. Murray[1,2] | Alan O. Bergland[1]

[1]Department of Biology, University of Virginia, Charlottesville, Virginia, USA | [2]Department of Genome Sciences, University of Virginia School of Medicine, Charlottesville, Virginia, USA

Correspondence: Connor S. Murray (csm6hg@virginia.edu)

Received: 8 November 2024 | **Revised:** 14 April 2025 | **Accepted:** 7 May 2025

Funding: A.O.B. was supported by grants from the National Institutes of Health (R35 GM119686), and by start-up funds provided by the University of Virginia. C.S.M. was supported by an Expand National Science Foundation Research Traineeship program at UVA.

Keywords: *Daphnia* | gene family evolution | positive selection | stress response

ABSTRACT

Gene family expansion underlies a host of biological innovations across the tree of life. Understanding why specific gene families expand or contract requires comparative genomic investigations clarifying further how species adapt in the wild. This study investigates the gene family change dynamics within several species of *Daphnia*, a group of freshwater microcrustaceans that are insightful model systems for evolutionary genetics' research. We employ comparative genomics approaches to understand the forces driving gene evolution and draw upon candidate gene families that change gene numbers across *Daphnia*. Our results suggest that genes related to stress responses and glycoproteins generally expand across taxa, and we investigate evolutionary hypotheses of adaptation that may underpin expansions. Through these analyses, we shed light on the interplay between gene expansions and selection within other ecologically relevant stress response gene families. While we show generalities in gene family turnover in genes related to stress response (i.e., DNA repair mechanisms), most gene family evolution is driven in a species-specific manner. Additionally, while we show general trends toward positive selection within some expanding gene families, many genes are not under selection, highlighting the complexity of diversification and evolution within *Daphnia*. Our research enhances the understanding of individual gene family evolution within *Daphnia* and provides a case study of ecologically relevant genes prone to change.

1 | Introduction

A major goal in biology is to understand how adaptive evolution changes complex phenotypes (Lewontin 1974; Mayr 1963). Gene family expansion, mediated through the novel duplication of genes, is an influential process that provides species with the opportunity for biological innovation to occur (Hahn et al. 2007; Jordan et al. 2001). Neofunctionalization is the process wherein gene copies adopt new functions (Ohno 2013), while subfunctionalization is the process in which a paralogous gene retains a part of the progenitor's role (Lynch and Force 2000). Both neofunctionalization and subfunctionalization are ways in which gene duplications lead to diversification resulting from the fixation of beneficial amino acid substitutions (Lynch 2002). Gene family expansions have long been hypothesized as the product of adaptive evolution across the tree of life, from microbes to mammals (Huang, Jiang, et al. 2023; Huang, Lu, et al. 2023; Zhang et al. 2014; Hahn et al. 2007; Jordan et al. 2001; Lugli et al. 2017; Richter et al. 2018).

When developing hypotheses for studying gene family evolution, we often view gene content change as facilitating adaptation to environmental (abiotic and biotic) shifts. For instance, animals inhabiting extreme temperature regimes typically have an increased number of heat-shock proteins (Chen et al. 2018; Zhang et al. 2012). In this way, gene family expansions have allowed taxa

to adapt and survive during temperature fluctuations (Lindquist and Craig 1988). While heat-shock proteins are known candidates of environmentally induced gene family expansions, opsins are another source of constant gene family fluctuations due to adaptation to various light environments and ultraviolet sensitivities (Novales Flamarique 2013), as well as chemosensory gene families (Peñalva-Arana et al. 2009). Additionally, innate immune response proteins and DNA repair mechanisms are also common expansion targets (Teekas et al. 2022), yet lineage-specific expansions are also expected due to specific adaptation to local environments (Lespinet et al. 2002). Overall, interpreting how gene family changes occur across related species is a worthwhile pursuit, especially for taxa prone to gene family turnover in response to environmental decay because it reveals a component of adaptation that changes many potential protein targets across an organism (Guijarro-Clarke et al. 2020). By using comparative genomics, we can test hypotheses related to the timing of gene family diversification and lay the groundwork for better understanding adaptation within the context of both individual species and groups of taxa (Mendes et al. 2021).

In this work, we assess gene family evolution and infer the strength of selection acting upon *Daphnia*, a diverse genera of small freshwater Crustaceans that live in a range of habitats from ephemeral rain puddles to lakes and even manmade estuaries (Fryer 1991). *Daphnia* adaptively radiated roughly 200 million years ago (mya) and encompasses at least 121 species to date within the Daphniidea family. Subspecies and cryptic speciation are common within *Daphnia*, and so this number of species is likely an underestimate (Forró et al. 2008). One of the most studied taxa within Daphniidea is *Daphnia pulex*, a complex of cryptic species found across both North American and European ponds (Colbourne et al. 1998; Crease et al. 2012; Murray et al. 2024; Vergilino et al. 2011). The first Crustacean genome described was *D. pulex* (Colbourne et al. 2011) and subsequent studies showed that even within species' lineages of *D. pulex* show variability in the number of genes present (Brandon et al. 2017; Lynch et al. 2017).

Daphnia species show fluctuations in the spectrum of gene gain and loss in response to environmental cues, supporting the case that gene family change is an important evolutionary mechanism in these taxa (Hamza et al. 2023; Schurko et al. 2009; Zhang et al. 2021). The focal ponds that *Daphnia* inhabit will periodically go through environmental degradation and hypoxia (Paul et al. 1998). Occasionally, this hypoxia is driven by eutrophication caused by runoff of nutrients into freshwater systems (Ebert 2022; Frisch et al. 2014). While this environmental stress largely occurs in areas devoted to agriculture, it can nonetheless affect species across entire ranges, especially in areas prone to temperature fluctuations (Smith and Schindler 2009). One way in which *Daphnia* responds to eutrophication-induced hypoxia is to increase the production of hemoglobin, which will aid oxygen binding (Fox et al. 1951). This increase in heme production results in a body-wide red hue that is potentially conserved across *Daphnia* (Zeis 2020).

Daphnia are also notable because they have a unique reproductive mode of cyclical parthenogenesis whereby females have rounds of clonal reproduction followed by a sexual event triggered through environmental cues (Rouger et al. 2016). Cyclical parthenogenesis is common across the tree of life, yet we only have a limited understanding of the gene family pathways leading to this reproductive novelty, of which meiotic pathways are thought to be especially important (Schurko et al. 2009). Because some *Daphnia* express both asexual and sexual modes, their molecular machineries to accommodate these phenotypes must be conserved in the same genome, making them an attractive model to understand reproductive-linked gene evolution (Lumer 1937; Zaffagnini and Sabelli 1972). Outside of reproductive mode variation, sperm morphology is highly polymorphic and is divergent in tail length between the *Daphnia* and *Ctenodaphnia* subgenera (Duneau et al. 2022). Sperm morphology differences motivate the question of which genes have expanded or contracted across *Daphnia* and *Ctenodaphnia*? And are there ecologically relevant gene families related to stress, immune responses, and reproduction that have evolved across *Daphnia*?

In this work, we survey gene family evolution across *Daphnia*, highlight the expanding gene families, and measure the strength of natural selection acting upon candidate genes. In this way, we test an overarching hypothesis that expanding gene families are also under positive selection. Our results show substantial gene content shifts across species and that stress response (DNA repair) and glycoprotein-associated gene families largely fluctuate across *Daphnia* genomes. We detect positive selection within some of these overrepresented gene families, indicating a link between selection and gene content expansion among ecologically relevant genes.

2 | Materials and Methods

2.1 | *Daphnia* Whole-Genome Dataset

Chromosome and scaffold-level assemblies of seven species from the genus *Daphnia* were collected from the NCBI Genome search engine (https://www.ncbi.nlm.nih.gov/datasets/genome/) accessed in January 2025 (Kitts et al. 2016). We chose North American *D. pulex* (KAP4; RefSeq: GCF_021134715.1), European *D. pulex* (D84A; GenBank: GCA_023526725.1; Barnard-Kubow et al. 2022), North American *D. pulicaria* (RefSeq: GCF_021234035.1; Wersebe et al. 2023), *D. sinensis* (GenBank: GCA_013167095.2, Jia et al. 2022), *D. carinata* (RefSeq: GCF_022539665.2), *D. galeata* (GenBank: GCA_030770115.1; Nickel et al. 2021), and *D. magna* (RefSeq: GCF_020631705.1) for analyses because they are the most complete species representatives and were annotated for protein-coding genes. These genomes were the highest quality and newest available for each unique species. We also included *Artemia franciscana* (i.e., brine shrimp; GenBank: GCF_032884065.1) and *Penaeus monodon* (i.e., black tiger shrimp; GenBank: GCF_015228065.2) as Crustacean outgroups for our tree-based analyses.

2.2 | Computational Biology Pipeline

We implemented a *nextflow v23.04.1* (Di Tommaso et al. 2017) pipeline (named *nf-GeneFamilyEvolution*) to organize the "flow" of data between processes and implement reproducibility and accessibility across computational environments. Briefly, we set up the nextflow pipeline to take RefSeq/GenBank genome identifiers, download their genome and proteome from the NCBI, classify ortholog groups, run various software related to phylogenetic tree building, and perform gene family evolution analyses with minimal setup

and editing (Figure S1A). The following sections describe the individual components that are incorporated in the pipeline.

2.3 | Estimating Divergence-Time Across Species

We used benchmark universal single copy orthologous (*BUSCO*) *v5* (Manni et al. 2021) Arthropoda dataset on each genome assembly to acquire the BUSCO score and used 164 complete single-copy genes for phylogenomic analyses of divergence dating across the *Daphnia* genera (Figure S1B). To do this, we extracted each amino acid alignment for each gene using *seqkit faidx v2.2.0* (Shen et al. 2016) and aligned them using *mafft –auto v7.505* (Katoh and Standley 2013). Then, we used *clipkit-m smart-gap v2.1.1* (Steenwyk et al. 2020) to clip out regions with large gaps. After this, we concatenated all of the protein sequences together using *seqkit concat v2.2.0* and used the *mcmctree_tree_prep.py* script (https://github.com/kfuku52/mcmctree_tree_prep) to create the necessary input files for *MCMCtree v4.9e* (Dos Reis and Yang 2019), which is part of the *PAML* package of software (Yang 2007). We assembled a gene-tree using *IQtree v2.2.0.3* with the *modelfinder* option (Kalyaanamoorthy et al. 2017).

We used several time-calibration points from previous phylogenetic investigations in *Daphnia*. Specifically, we used *D. carinata—D. magna* (100.4–104.8 mya) in place of the subgenus *Daphnia—Ctenodaphnia* root comparison in Cornetti et al. (2019). *Artemia* and *D. magna* (365.1–492 mya), *Artemia* and *Penaeus* (275–541 mya), and North American *D. pulex—D. magna* (130–150 mya) were taken from *Timetree5* (Kumar et al. 2022; Mathers et al. 2013), and *D. magna—D. sinensis* (21.5–22.4 mya) was from Cornetti et al. (2019). We used the *MCMCtreeR v1.1* package in *R* to plot the 95% confidence interval of divergence estimates across taxa (Figure S1C; Puttick 2019). *MCMCtree* was run several times to ensure model convergence by showing minor deviations in the estimate of node divergence times and the mean time for each node.

2.4 | Gene Family Evolution Analyses and Ontology Enrichment

We first identified and retained the longest transcript with an open-reading frame in the sequence for each gene using the *primary_transcript.py* script within the *OrthoFinder v2.5.5* tool set (Emms and Kelly 2019). We then classified orthologous genes between the FASTA format proteomes using *OrthoFinder*. After identifying the orthologous genes between the seven species, we annotated the phylogenetic hierarchical orthologous groups (HOGs) with the most common gene name by a majority vote using the *annotate_orthogroups* function in *orthofinder-tools* (https://github.com/MrTomRod/orthofinder-tools). We used the HMMER dataset and *hmmer2go v0.18.2* functions (https://github.com/sestaton/HMMER2GO) to assign gene ontology (GO) terms to the HOGs for use in enrichment analyses. We performed quality control by hand for some single-copy ortholog gene families identified in *OrthoFinder* by BLASTing each amino acid sequence against the NCBI database using *blastp v2.13.0* (Sayers et al. 2022). This quality control step verified the GO term assignments and supported the translated function of each protein-coding gene tested ($N = 15$).

CAFE5 v1.1 (Mendes et al. 2021) was used to estimate the expansions and contractions of gene families across the *Daphnia* proteomes. Before running *CAFE5*, we excluded any HOGs that had over 80 genes present within any one species and any genes that were exclusively present in only one species, according to best practices in the *CAFE5* vignette. After this, we ran different models and found that the model with a varying gamma rate and root Poisson distribution (-k3 -p) fits our dataset and converged. After this, we extracted the HOGs found to be significantly expanding or contracting within each species and used those genes as the foreground and each species' genome as the background with *clusterProfiler v3.14.3* in *R* (Wu et al. 2021). We used *REVIGO v1.8.1* as a semantic reduction tool to minimize GO term redundancy for any enriched terms (Supek et al. 2011) and performed Bonferroni–Holm multiple testing corrections on *p*-values using *p.adjust* in *R*.

2.5 | Hypothesis Testing of Positive Selection on Codon Sequences

To test for positive selection across gene families, we used *hyphy v2.5* adaptive Branch-Site Random Effects Likelihood (*aBSREL*; Kosakovsky Pond et al. 2020) on aligned codon FASTAs generated from *OrthoFinder*. *aBSREL* tests for positive selection occurring across branches of a tree by varying the rate of selection dN/dS (ω) across both sites and branches; in this way, it models both site- and branch-level dN/dS heterogeneity (Smith et al. 2015). *aBSREL* fits a full adaptive model; a likelihood ratio test (LRT) is then performed at each branch and compares the full model to a null model where branches are not allowed to have rate classes of $dN/dS > 1$ (Kosakovsky Pond et al. 2020). As recommended by tutorials on *aBSREL*, we performed one test on each tree, comparing all leaf nodes (i.e., tips) in a pairwise manner (Spielman et al. 2019). We tested trees to understand the patterns of selection occurring on those that have potentially undergone neofunctionalization (Hou et al. 2013; Mulhair et al. 2023; Saad et al. 2018; Wang, Zhang, Yang, et al. 2023; Wang, Zhang, Wang, et al. 2023). We translated the amino acid alignments and corresponding nucleotide sequences with *pal2nal.pl v14* (Suyama et al. 2006) and excluded any premature stop codons and gaps. We next aligned the sequences with *mafft* and visually inspected alignments for quality, and removed any alignments that had evidence of artificial frameshifts. We ran each translated codon FASTA independently with 10 cores and fixed the gene tree for usage in each *aBSREL* run. We read the *aBSREL* json files into *R* for further analysis and plotting. We also used the *Datamonkey v2.0* webserver (Weaver et al. 2018) to export trees from the *aBSREL* models. We counted a gene family as being under significant positive selection where $1 < dN/dS\ (\omega) < 10$ and the multiple-testing corrected *p*-value < 0.05. We tested the odds ratio enrichment of the gene families that belong to the expanded genes identified from *CAFE5* against the genes that are non-fluctuating using a two-tailed Fisher's exact test in *R*.

2.6 | Statistics and Visualization in R

Statistical analyses were performed using *R v4.3.1* (R Core Development Team 2013). We used the following *R* packages for general analysis and visualization: *tidyverse v1.3.1* (Wickham et al. 2019), *ggplot2 v3.3.5* (Villanueva and Chen 2019),

ggtree v2.0.4 (Xu et al. 2022), *patchwork v1.0.1* (Pedersen 2022), *data.table v1.12.8* (Dowle and Srinivasan 2023), *foreach v1.4.7*, and *doMC v1.3.5* (Daniel et al. 2022).

3 | Results

3.1 | *Daphnia* Genomes and the Gene Family Dataset

From the *Daphnia* genomes, we first identified 160,498 unique genes across the whole dataset after extracting the longest open reading frame per protein coding transcript. From these, *Orthofinder* identified 1,129 single-copy orthologous gene families. Within the total gene dataset, *Orthofinder* found 13,784 hierarchical phylogenetic orthologous gene groups (HOGs). We use these 13,784 HOGs as input into *CAFE5* to estimate the evolutionary rate of gene family gain and loss and to identify any gene groups that are evolving across the phylogeny. Below, we use this gene grouping information to expand our understanding of the phylogenetic relationship between taxa.

3.2 | Phylogeny of the Represented *Daphnia* Genomes

We built a time-calibrated phylogenetic tree to understand the relatedness of each *Daphnia* genome as well as create a phylogeny to serve in our gene family analyses (Figure S1C). We find that North American and European *D. pulex* diverged 12.5 mya (95% confidence intervals; 7.2, 18.2). Recent work highlights a split at 10 mya using BUSCO gene SNPs, well in range of our estimates (Murray et al. 2024). Previous estimates of the split time for North American and European *D. pulex* place it as early as 2–3 mya based on mitochondrial genes and assorted DNA loci (Colbourne et al. 1998; Crease et al. 2012). We also show that North American *D. pulex* and *D. pulicaria* diverged 5.6 mya (2.9, 8.5), an estimate higher than previous at 0.5–2 mya (Crease et al. 2012). We also refine the split between the *Ctenodaphnia* and *Daphnia* subgenera to 35.8 mya (27.3, 45.5) (Figure S1C).

3.3 | Trends of Gene Family Evolution

We used *CAFE5* to identify expanding and contracting gene families across our dataset (Figure 1; Mendes et al. 2021). The *k3p CAFE5* model minimized the negative log-likelihood value, so we are reporting its output. Our first finding is that *D. magna* has the largest expansion within its genome ($N_{\text{Genes}} = 1{,}054$; Figure 1), while European *D. pulex* has the largest contraction ($N_{\text{Genes}} = 1{,}309$; Figure 1). Also, across the dataset, there are more expansions than contractions when taking the average loss and gain across individual genomes. There are 673 individual genes being lost and 567 genes being gained. The range in gene gains (202–1,054) and loss (249–1,309) is wide for both classes. For the remainder of our work, we investigate the genes belonging to the 3,161 phylogenetic HOGs identified as significant candidates that are evolving across the species tree, as well as ecologically relevant gene families related to stress response that are tested below.

Next, we were interested in understanding how gene expansions are related to function across species (Lespinet et al. 2002; Sánchez-Gracia et al. 2009). We explicitly want to understand if

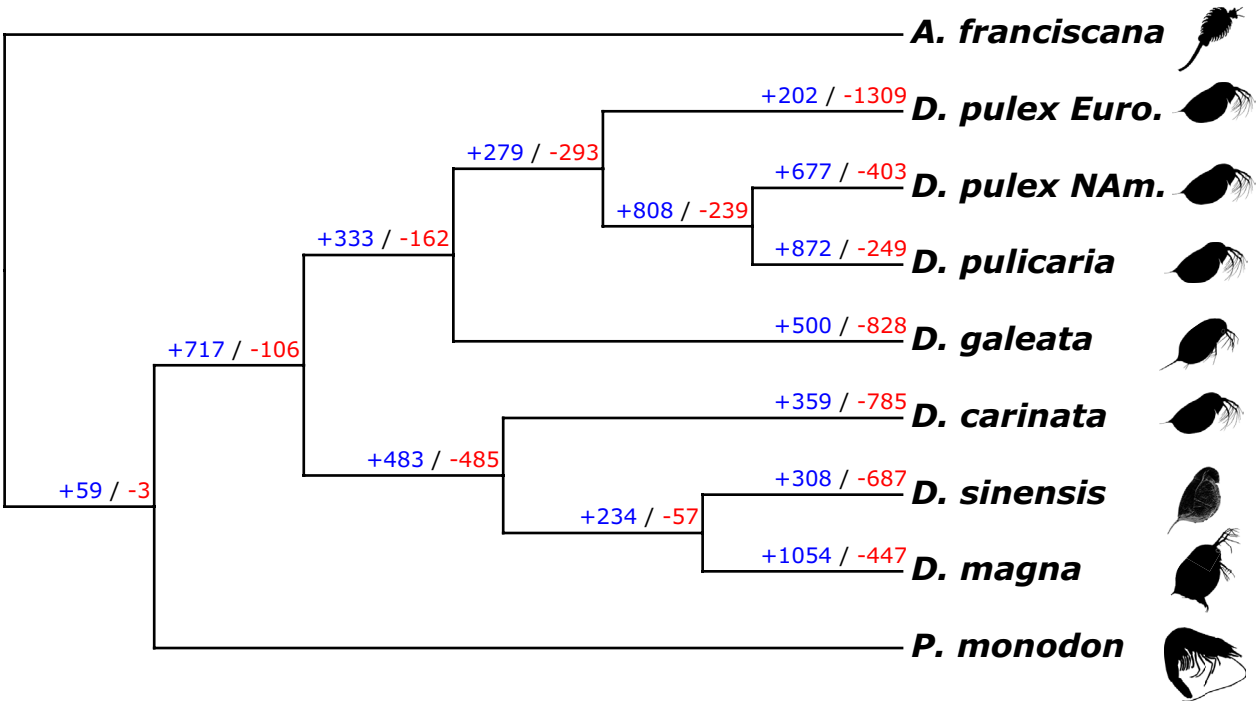

FIGURE 1 | Gene family dynamics across *Daphnia*. Results of gene family evolution analyses across the phylogenetic tree from *CAFE5*. The blue colored numbers indicate the number of genes gained and the red indicate the number of genes lost within each node and terminal leaf. Icons were taken from PhyloPic.com.

there are common functions across all species that could be related to ecologically relevant phenotypes for Daphnids. We examine whether there is a generality across all species within our data by first measuring the enrichment of GO terms belonging to the expanding genes within each species' genome. We extracted the most common expanding genes in this dataset, and one intriguing pattern is that they are largely involved with glycolysis and glycoprotein biosynthetic processes (Figure 2), potentially linked with Daphnids ability to withstand periods of anoxia, food limitation, and stress response. Additionally, most Daphnids have terms related to double-strand DNA repair enriched, except for *D. magna* and North American *D. pulex*, which may represent a general stress response to DNA damage across the genus (Figure 2). We see some general expansions for G-protein processes and phosphorylation across species as well and have a notable expansion of several terms related to immune responses in *D. magna*, which is a well-known system studied for resistance to pathogens (Ebert 2022).

We do not show any specific expansions of terms belonging to the *Ctenodaphnia* subgenera (e.g., *D. carinata*, *D. magna*, *D. sinensis*) save for lipid modification; however, there are several terms that belong specifically to the *Daphnia* subgenera, including transition metal ion transport and sensory perception of touch among others (Figure 2). Yet, there is not a specific set of terms defining the difference between subgenera (*Ctenodaphnia* vs. *Daphnia*) for both contractions and expansions (Figures S2 and S3).

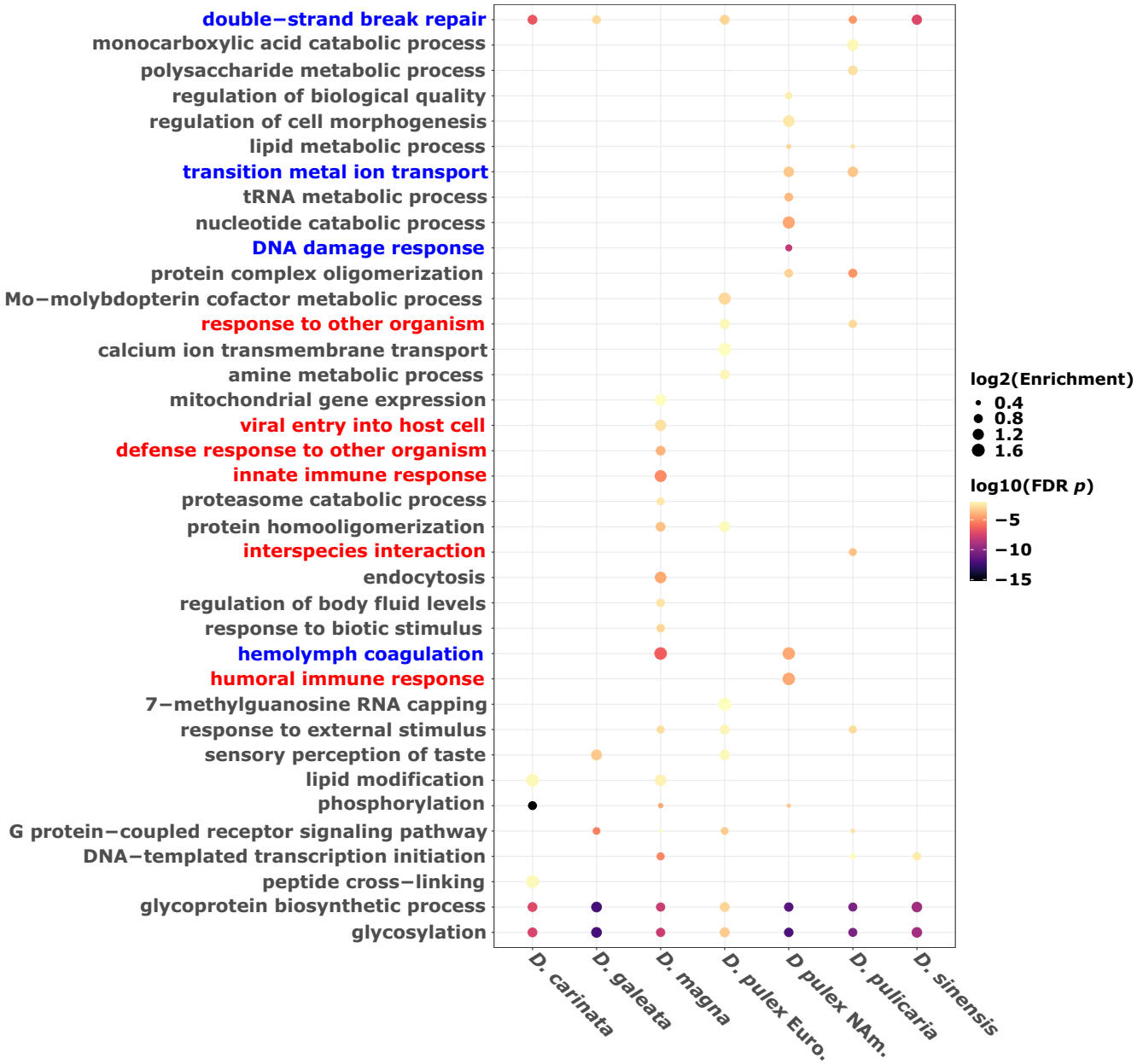

FIGURE 2 | Significantly expanding gene families and their gene ontology enrichment across species reveals an excess of stress response terms. The presence and absence data of the most enriched terms across species. The y-axis is the enriched terms within each species. The red colored GO terms indicate any terms related to immune responses and the blue terms are general stress responses. All terms have been semantically reduced and condensed long descriptors.

Outside of these examples from the enrichment data, there are many unique expanding and contracting GO terms found to be enriched in only one or two species (Figures S2 and S3). While we classify the enrichment of evolving gene families here, we test below whether these commonly enriched genes undergo selection across *Daphnia* genomes.

3.4 | General Patterns of Positive Selection on Evolving Gene Families

We tested the hypothesis that expanding gene families undergo higher rates of selection by using *hyphy v2.5 aBSREL* (Kosakovsky Pond et al. 2020) on codon sequences from the relevant gene families identified from *OrthoFinder*. We excluded gene families that identified extreme values of *dN/dS* (ω) > 10 and classified a tip, or branch, as being under positive selection if both the *dN/dS* > 1 and the multiple testing adjusted *p*-value < 0.05. We found that the expanded gene families have 6 of the 410 (1.5%) trees with positive selection compared to the non-fluctuating class (32 in 5,044; 0.63%). A Fisher's exact test shows a positive odds ratio of 2.31 (95% CI: 0.96, 5.6) (two-tailed Fisher's exact test; $z = 1.87$; $p = 0.062$), yet non-significant.

3.5 | Evolution of Stress Response Gene Families and Natural Selection in *Daphnia*

We investigated whether gene family expansions associated with stress response GO terms identified in Figure 2 are subject to positive selection using the *aBSREL* method. Specifically, we examined gene families involved in iron-ion binding, heme production, and hemolymph processes due to their relevance in hypoxia adaptation and stress responses in *Daphnia* (Zeis et al. 2009). Initially, we focused on the *hemoglobin-1* gene family, which showed significant expansions within the transition metal-ion binding and hemolymph coagulation GO terms (Figure 2) and is directly involved in stress response and heme production among *Daphnia*.

Positive selection was assessed using LRTs with an adjusted threshold of $p < 0.05$ after multiple testing corrections. Trees were rooted using available outgroups when present. Using *aBSREL*, we identified positive selection specifically in *D. sinensis* ($p = 0.006$; Figure 3A). Despite substantial expansions in *D. galeata* ($N = 5$) and *D. pulicaria* ($N = 7$), no selection was detected in these expanded lineages (Figure 3A). We further explored the hemolymph response by examining the *2-oxoglutarate and iron-dependent oxygenase JMJD4-like* protein family, noting expansions ($N = 3$) in *D. magna*, although without any detectable selection. This gene family remains single-copy across all tested species and outgroups (including crustaceans *Artemia* and *Penaeus*), except *D. magna*, suggesting potential non-selective functional diversification (Figure S4).

Signals of selection were observed in the *hsc70-interacting* protein family within *D. galeata* ($p = 0.046$; Figure 3B), with expansions noted in both *D. magna* and the crustacean outgroup *Penaeus*. This indicates a broad evolutionary significance likely related to chaperone and heat shock protein function, which are widely conserved stress response components. Another related family, *heat shock protein beta-1-like isoform X1*, had expansions ($N = 2$) within *D. magna*, yet lacked detectable selection. Additionally, we explored genes related to physical defense responses in *Daphnia*, such as the *chitin-binding type-2 domain-containing* protein family, which had expansions ($N = 1$) in European *D. pulex*, but showed no selection signals.

Within the DNA damage and repair pathways, the *centrosomal protein of 164kDa-like* family exhibited selection signals across multiple branches (five of nine terminal branches analyzed). Specifically, *D. carinata* demonstrated strong positive selection ($p = 1.82 \times 10^{-6}$; Figure 3C) with gene expansions ($N = 2$). Interestingly, selection signals (adjusted $p < 0.05$) were widespread throughout the genus except in *D. sinensis*, and although North American *D. pulex* also exhibited expansions, no selection signals were detected. This family is likely critical for DNA repair and chromosomal organization and has broadly expanded among *Daphnia* while absent in examined outgroups (Figure 3C).

Immune-related pathways highlighted strong selection within the *clotting factor G beta subunit-like* protein family specific to North American *D. pulex* ($p = 6.93 \times 10^{-12}$) and associated with expansions ($N = 2$; Figure 2). Similarly, *putative defense protein 3*, potentially important in pathogen defense, expanded within *D. magna* ($N = 1$) and showed positive selection specifically in *D. galeata* ($p = 0.0032$; Figure 3D).

Reproductive gene families also showed lineage-specific dynamics, with a *testes-specific* protein family expanding and undergoing strong positive selection exclusively in *D. sinensis* ($p = 4.92 \times 10^{-5}$). Notably, related orthologs were restricted to *D. sinensis* and *D. magna*, suggesting specialized reproductive roles.

Analysis of glycosylation and glycoprotein synthesis terms, significantly expanded across the genus (Figure 2), revealed that the *glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase* family underwent positive selection in *D. carinata* ($p = 9.3 \times 10^{-5}$) without expansions, while *D. pulicaria* exhibited both selection ($p = 2.13 \times 10^{-5}$) and expansion ($N = 1$), illustrating species-specific adaptive divergence (Figure S5).

Beyond specific expansions, general patterns of selection included structural and cytoskeletal GO terms related to heme production, where 6 of 133 genes (4.5%) showed selection signals. In defense and immune-related GO terms, 2 of 66 genes (3%) had detectable selection. Among reproductive genes, 1 of 60 genes (1.7%) showed selection at terminal branches. Additionally, many GO terms remained under-classified or uncharacterized through simple term matching and majority voting; among these under-classified families, 39 of 1,676 genes (2.33%) exhibited positive selection. Collectively, our findings suggest that gene family expansions combined with lineage-specific episodes of positive selection are likely crucial adaptive responses across stress response, immune defense, reproductive processes, DNA damage repair, and glycosylation pathways within the *Daphnia* genus.

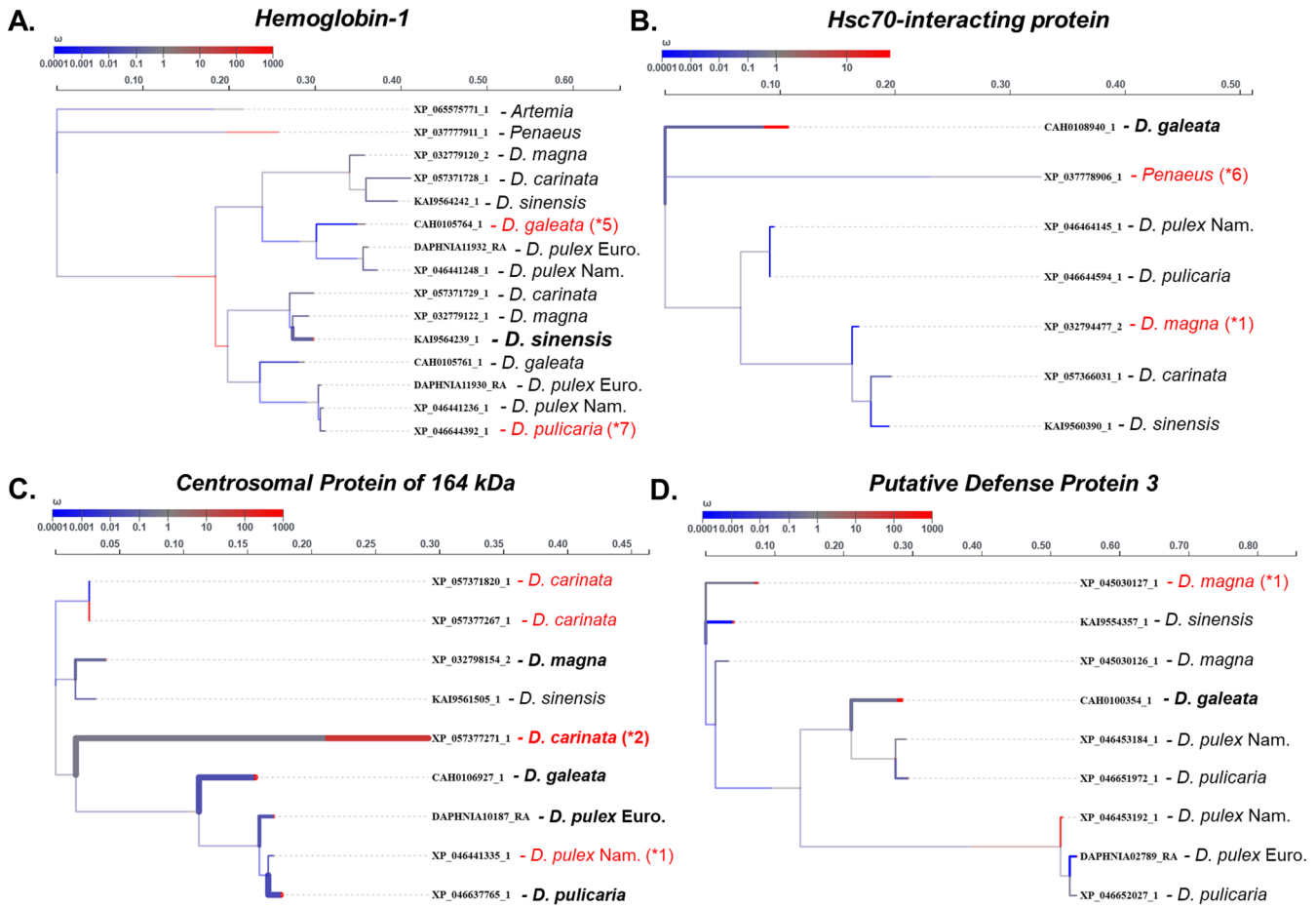

FIGURE 3 | General stress response gene families undergoing positive selection and expansions. (A–D) The thickness of branches indicates the level of significance, and any tips that are significant are bolded. If a tip label is colored red, that indicates that a significant expansion occurred with the number of expanded genes in parentheses. The gene names are included as tip labels with the species name. The color of branches indicates the estimated *dN/dS*.

4 | Discussion

In this work, we describe which gene families have expanded or contracted across *Daphnia*. We find overrepresented GO terms related to stress response. We show evidence for elevated positive selection across the gene families identified as being expanded across species, roughly affecting 1.5% of the expanding genes. Overall, we overview the general patterns of gene family expansion and contraction and identify candidate families under both positive selection that ultimately help our understanding of the evolutionary dynamics occurring within *Daphnia*.

4.1 | Evolutionary Dynamics of Gene Family Evolution in Daphnia

Daphnia are an interesting group of taxa to study from the perspective of gene family evolution. For instance, *Daphnia* have undergone adaptive radiations, have varying modes of asexual and sexual reproduction, and phenotypic plasticity occurring within and between populations related to predator defense and male production (Chin and Cristescu 2021; Hebert and Wilson 1994; Xu et al. 2013, 2015). Yet broadly speaking, *Daphnia* are algae grazers that make up an integral part of the biomass within freshwater systems, and as such, they may go through similar selection regimes related to predation (Schwartz 1984) and/or seasonal adaptation (Bergland et al. 2014; Winder et al. 2004). Our hypothesis was that *Daphnia* would have similar expansion relationships within their genomes related to these selective pressures in the wild. We show that there is a general trend toward expansions in metabolic (glycosylation and glycoprotein synthesis) and a number of stress responses in *Daphnia*, most notably DNA repair (Figure 2).

These results align in part with previous comparative analyses of *Daphnia* gene families. A comparative gene family analysis conducted by Zhang et al. (2021) using three *Daphnia* and several pancrustacean genomes revealed significant expansions that align with several enriched terms we identified. Particularly noteworthy were expansions related to methylation in *D. pulicaria* and North American *D. pulex*, as well as structural morphogenesis terms in *D. carinata* and North American *D. pulex*. Notably, Zhang et al. (2021) demonstrated that in the presence of fish kairomones, *D. mitsukuri* downregulates terms associated with heme production and iron binding while upregulating those linked to chemosensory and visual perception. Therefore, it is plausible that similar gene expression patterns could manifest in other *Daphnia* species in the presence of fish predators, a large selection pressure in ponds and lakes. Ye et al. (2017)

conducted a gene evolution analysis of two *Daphnia* species, primarily uncovering terms related to chitin binding and oxidative stress processes. Additionally, heme production genes were shown to be highly variable in North American *D. pulex* and *D. magna*, in relation to their ability to tackle the issues of hypoxia in small ponds (Ye et al. 2017). This is interesting because *Daphnia* are known to adapt to hypoxic environments through several hemoglobin proteins (Fox et al. 1951; Kobayashi et al. 1994).

Additionally, glycosylation terms are likely composed of genes related to stress and defense response to other organisms. Previous experimental results of exposing *D. magna* to kairomones from *Triops cancriformis* (i.e., eastern tadpole shrimp) showed upregulation of chitin production and subsequent cuticle changes related to glycosylation (Otte et al. 2014). The *Daphnia* cuticle is composed of lipids and waxes, chitin, and glycosylated/unglycosylated proteins (Minelli et al. 2016). In an experimental exposure of a Chinese-derived *D. pulex* to *Microcystis*, microbes known to produce toxic algae blooms, showed that *D. pulex* upregulates genes related to morphological change and glycoprotein synthesis (Huang, Jiang, et al. 2023; Huang, Lu, et al. 2023). *D. galeata* have similar responses to *Microcystis* and upregulate many overlapping terms with our expanding gene families like glycosylation and terms related to cuticle development (Kim et al. 2024). We show expansions in a glycoprotein synthesis family (Figure S5) and evidence for selection occurring in *D. carinata* and *D. pulicaria*. It could be likely that similar proteins are actively evolving and will likely reveal mechanisms that could be indicative of specific environmental or biotic responses.

We show overlap between our results and Ye et al. (2017), where we primarily observe expansions in glycosylation terms, a distinction possibly stemming from our utilization of a more extensive set of ortholog groups and inclusion of additional *Daphnia* species. This underscores the value of comparative genomics in elucidating key biological processes within *Daphnia*. This brings to light that comparative genomics in *Daphnia* has led to several interesting findings. The first *Daphnia* genome, *D. pulex arenata*, a subspecies of circumarctic *D. pulex*, had discovered over 30,000 unique genes (Colbourne et al. 2011). At the time, this number was over twice the amount in *Drosophila melanogaster* and humans. Upon reinvestigation, many are thought to be erroneous gene models due to fragmented draft genomes (Denton et al. 2014); however, some of these erroneous genes could nonetheless be describing evolutionarily significant events or splicing variants (Ye et al. 2017). We used the error prediction feature in *CAFE5* (Han et al. 2013), which calculates a predicted influence of genome assembly error on the estimates of our gene expansion and contraction and found it to be 8%. This estimate of error is in range with other projects (Neale et al. 2017) and is similar to *Drosophila* genomes (Da Lage et al. 2019). Also, the gene family gain and loss rate (λ) across the phylogeny is $\lambda = 0.0012$, a similar estimate to projects in *Drosophila* (Da Lage et al. 2019; Hahn et al. 2007). Therefore, the genomes tested here tend to have similar evolutionary rates of gene gain and loss across Arthropods and Crustaceans. We chose to include high-quality genomes available to minimize assembly bias, and we hope to include more in the future when created, especially those with chromosome-level scaffolds.

5 | Conclusion

Our study elucidates the gene family evolution of several members of *Daphnia*, and we provide supporting evidence that stress response genes are undergoing gene number evolution. We additionally show that some of these genes prone to expansions are also under positive selection, leading us to understand the gene diversification within *Daphnia*. Our study has important implications for continuing the work to elucidate the mechanisms that drive divergence across species, and we highlight the need to further validate how specific stress response genes are functional within species and populations of *Daphnia* (Genereux et al. 2020). Ultimately, although, we began the knowledge-building process necessary to link gene evolution with function across an interesting group of taxa prone to rapid adaptation.

Author Contributions

Connor S. Murray: conceptualization (equal), data curation (lead), formal analysis (lead), investigation (lead), methodology (lead), project administration (equal), resources (equal), software (equal), supervision (equal), visualization (lead), writing – original draft (equal), writing – review and editing (equal). **Alan O. Bergland:** conceptualization (equal), formal analysis (equal), funding acquisition (equal), investigation (equal), methodology (equal), project administration (equal), supervision (equal), writing – review and editing (equal).

Acknowledgments

The authors wish to acknowledge members of the Bergland lab for their discussion and feedback related to the manuscript's development. The authors acknowledge Research Computing at the University of Virginia for providing computational resources and technical support that have contributed to the results reported within this publication. URL: https://rc.virginia.edu. We also thank the two anonymous reviewers who's suggestions greatly improved our work.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All scripts and data used in every analysis are deposited in our GitHub repository along with the nextflow pipeline: https://github.com/connor122721/nf-GeneFamilyEvolution. All relevant data are deposited in data dryad at DOI: 10.5061/dryad.gqnk98t02.

References

Barnard-Kubow, K. B., D. Becker, C. S. Murray, et al. 2022. "Genetic Variation in Reproductive Investment Across an Ephemerality Gradient in *Daphnia Pulex*." *Molecular Biology and Evolution* 39, no. 6: msac121. https://doi.org/10.1093/molbev/msac121.

Bergland, A. O., E. L. Behrman, K. R. O'Brien, P. S. Schmidt, and D. A. Petrov. 2014. "Genomic Evidence of Rapid and Stable Adaptive Oscillations Over Seasonal Time Scales in *Drosophila*." *PLoS Genetics* 10, no. 11: e1004775. https://doi.org/10.1371/journal.pgen.1004775.

Brandon, C. S., M. J. Greenwold, and J. L. Dudycha. 2017. "Ancient and Recent Duplications Support Functional Diversity of *Daphnia* Opsins." *Journal of Molecular Evolution* 84, no. 1: 12–28. https://doi.org/10.1007/s00239-016-9777-1.

Chen, B., M. E. Feder, and L. Kang. 2018. "Evolution of Heat-Shock Protein Expression Underlying Adaptive Responses to Environmental

Stress." *Molecular Ecology* 27, no. 15: 3040–3054. https://doi.org/10.1111/mec.14769.

Chin, T. A., and M. E. Cristescu. 2021. "Speciation in *Daphnia*." *Molecular Ecology* 30: 15824. https://doi.org/10.1111/mec.15824.

Colbourne, J. K., T. J. Crease, L. J. Weider, P. D. N. Hebert, F. Duferesne, and A. Hobaek. 1998. "Phylogenetics and Evolution of a Circumarctic Species Complex (Cladocera: *Daphnia pulex*)." *Biological Journal of the Linnean Society* 65, no. 3: 347–365. https://doi.org/10.1111/j.1095-8312.1998.tb01146.x.

Colbourne, J. K., M. E. Pfrender, D. Gilbert, et al. 2011. "The Ecoresponsive Genome of *Daphnia Pulex*." *Science* 331, no. 6017: 555–561. https://doi.org/10.1126/science.1197761.

Cornetti, L., P. D. Fields, K. Van Damme, and D. Ebert. 2019. "A Fossil-Calibrated Phylogenomic Analysis of *Daphnia* and the Daphniidae." *Molecular Phylogenetics and Evolution* 137: 250–262. https://doi.org/10.1016/j.ympev.2019.05.018.

Crease, T. J., A. R. Omilian, K. S. Costanzo, and D. J. Taylor. 2012. "Transcontinental Phylogeography of the *Daphnia pulex* Species Complex." *PLoS One* 7, no. 10: e46620. https://doi.org/10.1371/journal.pone.0046620.

Da Lage, J.-L., G. W. C. Thomas, M. Bonneau, and V. Courtier-Orgogozo. 2019. "Evolution of Salivary Glue Genes in *Drosophila* Species." *BMC Evolutionary Biology* 19, no. 1: 36. https://doi.org/10.1186/s12862-019-1364-9.

Daniel, F., Revolution Analytics, and S. Weston. 2022. "doMC: Foreach Parallel Adaptor for 'parallel'."

Denton, J. F., J. Lugo-Martinez, A. E. Tucker, D. R. Schrider, W. C. Warren, and M. W. Hahn. 2014. "Extensive Error in the Number of Genes Inferred From Draft Genome Assemblies." *PLoS Computational Biology* 10, no. 12: e1003998. https://doi.org/10.1371/journal.pcbi.1003998.

Di Tommaso, P., M. Chatzou, E. W. Floden, P. P. Barja, E. Palumbo, and C. Notredame. 2017. "Nextflow Enables Reproducible Computational Workflows." *Nature Biotechnology* 35, no. 4: 316–319. https://doi.org/10.1038/nbt.3820.

Dos Reis, M., and Z. Yang. 2019. "Bayesian Molecular Clock Dating Using Genome-Scale Datasets." In *Evolutionary Genomics*, edited by M. Anisimova, vol. 1910, 309–330. Springer New York.

Dowle, M., and A. Srinivasan. 2023. "data.table: Extension of 'data.frame'." https://r-datatable.com, https://rdatatable.gitlab.io/data.table, https://github.com/rdatatable/data.table.

Duneau, D., M. Möst, and D. Ebert. 2022. "Evolution of Sperm Morphology in a Crustacean Genus With Fertilization Inside an Open Brood Pouch." *Peer Community Journal* 2: e63. https://doi.org/10.24072/pcjournal.182.

Ebert, D. 2022. "*Daphnia* as a Versatile Model System in Ecology and Evolution." *EvoDevo* 13, no. 1: 16. https://doi.org/10.1186/s13227-022-00199-0.

Emms, D. M., and S. Kelly. 2019. "OrthoFinder: Phylogenetic Orthology Inference for Comparative Genomics." *Genome Biology* 20, no. 1: 238. https://doi.org/10.1186/s13059-019-1832-y.

Forró, L., N. M. Korovchinsky, A. A. Kotov, and A. Petrusek. 2008. "Global Diversity of Cladocerans (Cladocera; Crustacea) in Freshwater." *Hydrobiologia* 595, no. 1: 177–184. https://doi.org/10.1007/s10750-007-9013-5.

Fox, H. M., B. M. Gilchrist, and E. A. Phear. 1951. "Functions of Haemoglobin in *Daphnia*." *Proceedings of the Royal Society of London, Series B: Biological Sciences* 138, no. 893: 514–528. https://doi.org/10.1098/rspb.1951.0038.

Frisch, D., P. K. Morton, P. R. Chowdhury, et al. 2014. "A Millennial-Scale Chronicle of Evolutionary Responses to Cultural Eutrophication in *Daphnia*." *Ecology Letters* 17, no. 3: 360–368. https://doi.org/10.1111/ele.12237.

Fryer, G. 1991. "Functional Morphology and the Adaptive Radiation of the Daphniidae (Branchiopoda: Anomopoda)." *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 331, no. 1259: 1–99. https://doi.org/10.1098/rstb.1991.0001.

Genereux, D. P., A. Serres, J. Armstrong, et al. 2020. "A Comparative Genomics Multitool for Scientific Discovery and Conservation." *Nature* 587, no. 7833: 240–245. https://doi.org/10.1038/s41586-020-2876-6.

Guijarro-Clarke, C., P. W. H. Holland, and J. Paps. 2020. "Widespread Patterns of Gene Loss in the Evolution of the Animal Kingdom." *Nature Ecology & Evolution* 4, no. 4: 519–523. https://doi.org/10.1038/s41559-020-1129-2.

Hahn, M. W., M. V. Han, and S.-G. Han. 2007. "Gene Family Evolution Across 12 *Drosophila* Genomes." *PLoS Genetics* 3, no. 11: e197. https://doi.org/10.1371/journal.pgen.0030197.

Hamza, W., K. M. Hazzouri, N. Sudalaimuthuasari, et al. 2023. "Genome Assembly of a Relict Arabian Species of *Daphnia* O. F. Müller (Crustacea: Cladocera) Adapted to the Desert Life." *International Journal of Molecular Sciences* 24, no. 1: 889. https://doi.org/10.3390/ijms24010889.

Han, M. V., G. W. C. Thomas, J. Lugo-Martinez, and M. W. Hahn. 2013. "Estimating Gene Gain and Loss Rates in the Presence of Error in Genome Assembly and Annotation Using CAFE 3." *Molecular Biology and Evolution* 30, no. 8: 1987–1997. https://doi.org/10.1093/molbev/mst100.

Hebert, P. D. N., and C. C. Wilson. 1994. "Provincialism in Plankton: Endemism and Allopatric Speciation in Australian *Daphnia*." *Evolution* 48, no. 4: 1333–1349. https://doi.org/10.1111/j.1558-5646.1994.tb05317.x.

Hou, Y., R. Sierra, D. Bassen, et al. 2013. "Molecular Evidence for β-Tubulin Neofunctionalization in Retaria (Foraminifera and Radiolarians)." *Molecular Biology and Evolution* 30, no. 11: 2487–2493. https://doi.org/10.1093/molbev/mst150.

Huang, Y., N. Lu, T. Yang, J. Yang, L. Gu, and Z. Yang. 2023. "Transcriptome Analysis Reveals the Molecular Basis of Anti-Predation Defence in *Daphnia Pulex* Simultaneously Responding to *Microcystis Aeruginosa*." *Freshwater Biology* 68, no. 8: 1372–1385. https://doi.org/10.1111/fwb.14110.

Huang, Z., C. Jiang, J. Gu, et al. 2023. "Duplications of Human Longevity-Associated Genes Across Placental Mammals." *Genome Biology and Evolution* 15, no. 10: evad186. https://doi.org/10.1093/gbe/evad186.

Jia, J., C. Dong, M. Han, et al. 2022. "Multi-Omics Perspective on Studying Reproductive Biology in *Daphnia Sinensis*." *Genomics* 114, no. 2: 110309. https://doi.org/10.1016/j.ygeno.2022.110309.

Jordan, I. K., K. S. Makarova, J. L. Spouge, Y. I. Wolf, and E. V. Koonin. 2001. "Lineage-Specific Gene Expansions in Bacterial and Archaeal Genomes." *Genome Research* 11, no. 4: 555–565. https://doi.org/10.1101/gr.166001.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. "ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates." *Nature Methods* 14, no. 6: 587–589. https://doi.org/10.1038/nmeth.4285.

Katoh, K., and D. M. Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30, no. 4: 772–780. https://doi.org/10.1093/molbev/mst010.

Kim, E. J., D. Jeon, Y. J. Park, H. Woo, and S. I. Eyun. 2024. "Dietary Exposure of the Water Flea *Daphnia galeata* to *Microcystin-LR*." *Animal Cells and Systems* 28, no. 1: 25–36. https://doi.org/10.1080/19768354.2024.2302529.

Kitts, P. A., D. M. Church, F. Thibaud-Nissen, et al. 2016. "Assembly: A Resource for Assembled Genomes at NCBI." *Nucleic Acids Research* 44, no. D1: D73–D80. https://doi.org/10.1093/nar/gkv1226.

Kobayashi, M., K. Ishigaki, M. Kobayashi, Y. Igarashi, and K. Imai. 1994. "Oxygen Transport Efficiency of Multiple-Component Hemoglobin in *Daphnia Magna*." *Canadian Journal of Zoology* 72, no. 12: 2169–2171. https://doi.org/10.1139/z94-289.

Kosakovsky Pond, S. L., A. F. Y. Poon, R. Velazquez, et al. 2020. "HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies." *Molecular Biology and Evolution* 37, no. 1: 295–299. https://doi.org/10.1093/molbev/msz197.

Kumar, S., M. Suleski, J. M. Craig, et al. 2022. "TimeTree 5: An Expanded Resource for Species Divergence Times." *Molecular Biology and Evolution* 39, no. 8: msac174. https://doi.org/10.1093/molbev/msac174.

Lespinet, O., Y. I. Wolf, E. V. Koonin, and L. Aravind. 2002. "The Role of Lineage-Specific Gene Family Expansion in the Evolution of Eukaryotes." *Genome Research* 12, no. 7: 1048–1059. https://doi.org/10.1101/gr.174302.

Lewontin, R. C. 1974. *The Genetic Basis of Evolutionary Change*. Vol. 560. Columbia University Press.

Lindquist, S., and E. A. Craig. 1988. "The Heat-Shock Proteins." *Annual Review of Genetics* 22, no. 1: 631–677. https://doi.org/10.1146/annurev.ge.22.120188.003215.

Lugli, G. A., C. Milani, F. Turroni, et al. 2017. "Comparative Genomic and Phylogenomic Analyses of the Bifidobacteriaceae Family." *BMC Genomics* 18, no. 1: 568. https://doi.org/10.1186/s12864-017-3955-4.

Lumer, H. 1937. "Growth and Maturation in the Parthenogenetic Eggs of *Daphnia Magna* Strauss." *Cytologia* 8, no. 1: 1–14. https://doi.org/10.1508/cytologia.8.1.

Lynch, M. 2002. "Gene Duplication and Evolution." *Science* 297, no. 5583: 945–947. https://doi.org/10.1126/science.1075472.

Lynch, M., and A. Force. 2000. "The Probability of Duplicate Gene Preservation by Subfunctionalization." *Genetics* 154, no. 1: 459–473. https://doi.org/10.1093/genetics/154.1.459.

Lynch, M., R. Gutenkunst, M. Ackerman, et al. 2017. "Population Genomics of *Daphnia pulex*." *Genetics* 206, no. 1: 315–332. https://doi.org/10.1534/genetics.116.190611.

Manni, M., M. R. Berkeley, M. Seppey, F. A. Simão, and E. M. Zdobnov. 2021. "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* 38, no. 10: 4647–4654. https://doi.org/10.1093/molbev/msab199.

Mathers, T. C., R. L. Hammond, R. A. Jenner, B. Hänfling, and A. Gómez. 2013. "Multiple Global Radiations in Tadpole Shrimps Challenge the Concept of 'Living Fossils'." *PeerJ* 1: e62. https://doi.org/10.7717/peerj.62.

Mayr, E. 1963. *Animal Species and Evolution*. Harvard University Press.

Mendes, F. K., D. Vanderpool, B. Fulton, and M. W. Hahn. 2021. "CAFE 5 Models Variation in Evolutionary Rates Among Gene Families." *Bioinformatics* 36, no. 22–23: 5516–5518. https://doi.org/10.1093/bioinformatics/btaa1022.

Minelli, A., G. Boxshall, and G. Fusco. 2016. *Arthropod Biology and Evolution*. Springer-Verlag Berlin An.

Mulhair, P. O., L. Crowley, D. H. Boyes, O. T. Lewis, and P. W. H. Holland. 2023. "Opsin Gene Duplication in Lepidoptera: Retrotransposition, Sex Linkage, and Gene Expression." *Molecular Biology and Evolution* 40, no. 11: msad241. https://doi.org/10.1093/molbev/msad241.

Murray, C. S., M. Karram, D. J. Bass, et al. 2024. "Trans-Specific Polymorphisms Between Cryptic *Daphnia* Species Affect Fitness and Behavior." *Molecular Ecology* 34, no. 3: e17632. https://doi.org/10.1111/mec.17632.

Neale, D. B., P. E. McGuire, N. C. Wheeler, et al. 2017. "The Douglas-Fir Genome Sequence Reveals Specialization of the Photosynthetic Apparatus in Pinaceae." *G3: Genes, Genomes, Genetics* 7, no. 9: 3157–3167. https://doi.org/10.1534/g3.117.300078.

Nickel, J., T. Schell, T. Holtzem, et al. 2021. "Hybridization Dynamics and Extensive Introgression in the *Daphnia Longispina* Species Complex: New Insights From a High-Quality *Daphnia Galeata* Reference Genome." *Genome Biology and Evolution* 13, no. 12: evab267. https://doi.org/10.1093/gbe/evab267.

Novales Flamarique, I. 2013. "Opsin Switch Reveals Function of the Ultraviolet Cone in Fish Foraging." *Proceedings of the Royal Society B: Biological Sciences* 280, no. 1752: 20122490. https://doi.org/10.1098/rspb.2012.2490.

Ohno, S. 2013. *Evolution by Gene Duplication*. Springer Science & Buisness Media.

Otte, K. A., T. Fröhlich, G. J. Arnold, and C. Laforsch. 2014. "Proteomic Analysis of *Daphnia Magna* Hints at Molecular Pathways Involved in Defensive Plastic Responses." *BMC Genomics* 15: 1–17. https://doi.org/10.1186/1471-2164-15-306.

Paul, R. J., M. Colmorgen, R. Pirow, Y.-H. Chen, and M.-C. Tsai. 1998. "Systemic and Metabolic Responses in *Daphnia magna* to Anoxia." *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 120, no. 3: 519–530. https://doi.org/10.1016/S1095-6433(98)10062-4.

Pedersen, T. L. 2022. "patchwork: The Composer of Plots." https://patchwork.data-imaginist.com, https://github.com/thomasp85/patchwork.

Peñalva-Arana, D. C., M. Lynch, and H. M. Robertson. 2009. "The Chemoreceptor Genes of the Waterflea *Daphnia pulex*: Many Grs but no Ors." *BMC Evolutionary Biology* 9, no. 1: 79. https://doi.org/10.1186/1471-2148-9-79.

Puttick, M. N. 2019. "MCMCtreeR: Functions to Prepare MCMCtree Analyses and Visualize Posterior Ages on Trees." *Bioinformatics* 35, no. 24: 5321–5322. https://doi.org/10.1093/bioinformatics/btz554.

R Core Development Team. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Richter, D. J., P. Fozouni, M. B. Eisen, and N. King. 2018. "Gene Family Innovation, Conservation and Loss on the Animal Stem Lineage." *eLife* 7: e34226. https://doi.org/10.7554/eLife.34226.

Rouger, R., K. Reichel, F. Malrieu, J. P. Masson, and S. Stoeckel. 2016. "Effects of Complex Life Cycles on Genetic Diversity: Cyclical Parthenogenesis." *Heredity* 117, no. 5: 336–347. https://doi.org/10.1038/hdy.2016.52.

Saad, R., A. B. Cohanim, M. Kosloff, and E. Privman. 2018. "Neofunctionalization in Ligand Binding Sites of Ant Olfactory Receptors." *Genome Biology and Evolution* 10, no. 9: 2490–2500. https://doi.org/10.1093/gbe/evy131.

Sánchez-Gracia, A., F. G. Vieira, and J. Rozas. 2009. "Molecular Evolution of the Major Chemosensory Gene Families in Insects." *Heredity* 103, no. 3: 208–216. https://doi.org/10.1038/hdy.2009.55.

Sayers, E. W., E. E. Bolton, J. R. Brister, et al. 2022. "Database Resources of the National Center for Biotechnology Information." *Nucleic Acids Research* 50, no. D1: D20–D26. https://doi.org/10.1093/nar/gkab1112.

Schurko, A. M., J. M. Logsdon, and B. D. Eads. 2009. "Meiosis Genes in *Daphnia pulex* and the Role of Parthenogenesis in Genome Evolution." *BMC Evolutionary Biology* 9, no. 1: 78. https://doi.org/10.1186/1471-2148-9-78.

Schwartz, S. S. 1984. "Life History Strategies in *Daphnia*: A Review and Predictions." *Oikos* 42, no. 1: 114. https://doi.org/10.2307/3544616.

Shen, W., S. Le, Y. Li, and F. Hu. 2016. "SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation." *PLoS One* 11, no. 10: e0163962. https://doi.org/10.1371/journal.pone.0163962.

Smith, M. D., J. O. Wertheim, S. Weaver, B. Murrell, K. Scheffler, and S. L. Kosakovsky Pond. 2015. "Less Is More: An Adaptive Branch-Site Random Effects Model for Efficient Detection of Episodic Diversifying Selection." *Molecular Biology and Evolution* 32, no. 5: 1342–1353. https://doi.org/10.1093/molbev/msv022.

Smith, V. H., and D. W. Schindler. 2009. "Eutrophication Science: Where Do We Go From Here?" *Trends in Ecology & Evolution* 24, no. 4: 201–207. https://doi.org/10.1016/j.tree.2008.11.009.

Spielman, S. J., S. Weaver, S. D. Shank, B. R. Magalis, M. Li, and S. L. Kosakovsky Pond. 2019. "Evolution of Viral Genomes: Interplay Between Selection, Recombination, and Other Forces." In *Evolutionary Genomics*, edited by M. Anisimova, vol. 1910, 427–468. Springer New York.

Steenwyk, J. L., T. J. Buida, Y. Li, X.-X. Shen, and A. Rokas. 2020. "ClipKIT: A Multiple Sequence Alignment Trimming Software for Accurate Phylogenomic Inference." *PLoS Biology* 18, no. 12: e3001007. https://doi.org/10.1371/journal.pbio.3001007.

Supek, F., M. Bošnjak, N. Škunca, and T. Šmuc. 2011. "REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms." *PLoS One* 6, no. 7: e21800. https://doi.org/10.1371/journal.pone.0021800.

Suyama, M., D. Torrents, and P. Bork. 2006. "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.

Teekas, L., S. Sharma, and N. Vijay. 2022. "Lineage-Specific Protein Repeat Expansions and Contractions Reveal Malleable Regions of Immune Genes." *Genes & Immunity* 23, no. 7: 218–234. https://doi.org/10.1038/s41435-022-00186-4.

Vergilino, R., S. Markova, M. Ventura, M. Manca, and F. Dufresne. 2011. "Reticulate Evolution of the *Daphnia pulex* Complex as Revealed by Nuclear Markers: Reticulate Evolution in the *Daphnia pulex* Complex." *Molecular Ecology* 20, no. 6: 1191–1207. https://doi.org/10.1111/j.1365-294X.2011.05004.x.

Villanueva, R. A. M., and Z. J. Chen. 2019. "ggplot2: Elegant Graphics for Data Analysis." *Measurement: Interdisciplinary Research and Perspectives* 17, no. 2nd ed.: 160–167. https://doi.org/10.1080/15366367.2019.1565254.

Wang, S., Y. Zhang, W. Yang, et al. 2023. "Duplicate Genes as Sources for Rapid Adaptive Evolution of Sperm Under Environmental Pollution in Tree Sparrow." *Molecular Ecology* 32, no. 7: 1673–1684. https://doi.org/10.1111/mec.16833.

Wang, W., X.-S. Zhang, Z.-N. Wang, and D.-X. Zhang. 2023. "Evolution and Phylogenetic Diversity of the Aquaporin Gene Family in Arachnids." *International Journal of Biological Macromolecules* 240: 124480. https://doi.org/10.1016/j.ijbiomac.2023.124480.

Weaver, S., S. D. Shank, S. J. Spielman, M. Li, S. V. Muse, and S. L. Kosakovsky Pond. 2018. "Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes." *Molecular Biology and Evolution* 35, no. 3: 773–777. https://doi.org/10.1093/molbev/msx335.

Wersebe, M. J., R. E. Sherman, P. D. Jeyasingh, and L. J. Weider. 2023. "The Roles of Recombination and Selection in Shaping Genomic Divergence in an Incipient Ecological Species Complex." *Molecular Ecology* 32, no. 6: 1478–1496. https://doi.org/10.1111/mec.16383.

Wickham, H., M. Averick, J. Bryan, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4, no. 43: 1686. https://doi.org/10.21105/joss.01686.

Winder, M., P. Spaak, and W. M. Mooij. 2004. "Trade-Offs in *Daphnia* Habitat Selection." *Ecology* 85, no. 7: 2027–2036. https://doi.org/10.1890/03-3108.

Wu, T., E. Hu, S. Xu, et al. 2021. "clusterProfiler 4.0: A Universal Enrichment Tool for Interpreting Omics Data." *Innovation* 2, no. 3: 100141. https://doi.org/10.1016/j.xinn.2021.100141.

Xu, S., D. J. Innes, M. Lynch, and M. E. Cristescu. 2013. "The Role of Hybridization in the Origin and Spread of Asexuality in *Daphnia*." *Molecular Ecology* 22, no. 17: 4549–4561. https://doi.org/10.1111/mec.12407.

Xu, S., L. Li, X. Luo, et al. 2022. "Ggtree: A Serialized Data Object for Visualization of a Phylogenetic Tree and Annotation Data." *iMeta* 1, no. 4: e56. https://doi.org/10.1002/imt2.56.

Xu, S., K. Spitze, M. S. Ackerman, et al. 2015. "Hybridization and the Origin of Contagious Asexuality in *Daphnia pulex*." *Molecular Biology and Evolution* 20: msv190. https://doi.org/10.1093/molbev/msv190.

Yang, Z. 2007. "PAML 4: Phylogenetic Analysis by Maximum Likelihood." *Molecular Biology and Evolution* 24, no. 8: 1586–1591. https://doi.org/10.1093/molbev/msm088.

Ye, Z., S. Xu, K. Spitze, et al. 2017. "A New Reference Genome Assembly for the Microcrustacean *Daphnia pulex*." *G3: Genes, Genomes, Genetics* 7, no. 5: 1405–1416. https://doi.org/10.1534/g3.116.038638.

Zaffagnini, F., and B. Sabelli. 1972. "Karyologic Observations on the Maturation of the Summer and Winter Eggs of *Daphnia pulex* and *Daphnia middendorffiana*." *Chromosoma* 36, no. 2: 5213. https://doi.org/10.1007/BF00285213.

Zeis, B. 2020. "Hemoglobin in Arthropods—*Daphnia* as a Model." In *Vertebrate and Invertebrate Respiratory Proteins, Lipoproteins and Other Body Fluid Proteins*, edited by U. Hoeger and J. R. Harris, vol. 94, 163–194. Springer International Publishing. https://doi.org/10.1007/978-3-030-41769-7_6.

Zeis, B., T. Lamkemeyer, R. J. Paul, et al. 2009. "Acclimatory Responses of the *Daphnia pulex* Proteome to Environmental Changes. I. Chronic Exposure to Hypoxia Affects the Oxygen Transport System and Carbohydrate Metabolism." *BMC Physiology* 9, no. 1: 7. https://doi.org/10.1186/1472-6793-9-7.

Zhang, G., X. Fang, X. Guo, et al. 2012. "The Oyster Genome Reveals Stress Adaptation and Complexity of Shell Formation." *Nature* 490, no. 7418: 49–54. https://doi.org/10.1038/nature11413.

Zhang, G., C. Li, Q. Li, et al. 2014. "Comparative Genomics Reveals Insights Into Avian Genome Evolution and Adaptation." *Science* 346, no. 6215: 1311–1320. https://doi.org/10.1126/science.1251385.

Zhang, X., D. Blair, J. Wolinska, et al. 2021. "Genomic Regions Associated With Adaptation to Predation in *Daphnia* Often Include Members of Expanded Gene Families." *Proceedings of the Royal Society B: Biological Sciences* 288, no. 1955: 20210803. https://doi.org/10.1098/rspb.2021.0803.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.