RESEARCH ARTICLE OPEN ACCESS

Comparative Embryology and Transcriptomics of *Asellus infernus*, an Isopod Crustacean From Sulfidic Groundwater

Haeli J. Lomheim[1,2] | Lizet Reyes Rodas[1] | Devon Price[1] | Serban M. Sarbu[3,4,5] | Raluca I. Băncilă[3,5] | Cody Carroll[6] [iD] | Layla Freeborn[7] [iD] | Sheri Sanders[8] | Meredith E. Protas[1,9] [iD]

[1]Department of Natural Sciences and Mathematics, Dominican University of California, San Rafael, California, USA | [2]Department of Biology, Georgetown University, Washington DC, Washington DC, USA | [3]"Emil Racoviţă" Institute of Speleology, Department Cluj-Napoca, Cluj-Napoca, Romania | [4]Department of Biological Sciences, California State University, Chico, California, USA | [5]"Emil Racoviţă" Institute of Speleology, Romanian Academy, Bucharest, Romania | [6]Department of Mathematics and Statistics, Data Institute, University of San Francisco, San Francisco, California, USA | [7]Research Computing, Office of Information Technology, The Center for Research Data and Digital Scholarship, University of Colorado Boulder, Boulder, Colorado, USA | [8]Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, USA | [9]Department of Biology, University of San Francisco, San Francisco, California, USA

Correspondence: Meredith E. Protas (mprotas@usfca.edu)

ABSTRACT

Sulfidic caves are harsh and extreme environments characterized by limited oxygen, low pH, and the presence of hydrogen sulfide. Amazingly, animals can live in sulfidic caves, one such animal being *Asellus infernus*, a representative of the *Asellus aquaticus* species complex, originating from Movile Cave and from old wells that represent windows of access to a sulfidic groundwater ecosystem located in southeast Romania. Little previous work has been done on lab-reared populations of *A. infernus* as they have been historically difficult to raise in the lab. Here, we develop resources for *A. infernus*, examining questions of timing of morphological differences in cave versus surface individuals, whether the environment (lab-bred vs. wild-caught) influenced size characteristics, and the genes and pathways showing differential expression between cave and surface samples. We found that *A. infernus* did not develop pigmentation embryonically, and juveniles had increased body length and longer antenna II as compared to surface individuals. Furthermore, we found that some of these measures differed between wild-caught and lab-reared juveniles for a given population, indicating that environmental differences can also influence these

size characteristics. In addition, differential expression between cave and surface samples and allele-specific expression studies within F1 hybrids identified multiple genes, including those involved in sulfide metabolism and phototransduction. Strikingly, molecular convergence of genes involved in sulfide detoxification was observed between *A. infernus* and previous work on a fish that lives in both cave and sulfidic environments, *Poecilia mexicana*. In sum, we were able to develop embryonic and genomic tools for *A. infernus*, a model for understanding cave adaptation and adaptation to sulfidic environments.

1 | Introduction

Caves, considered extreme environments, are characterized by a lack of light, limited nutrient availability, and in some cases, extreme temperatures as compared to surface environments (Engel 2007). A specific type of cave, sulfidic cave, is further characterized by high hydrogen sulfide (H_2S), carbon dioxide (CO_2), methane (CH_4), ammonia (NH_4^+), and acidity levels. In addition, the temperature can be relatively high, such as 21°C. Copious food produced in situ by chemoautotrophic bacteria thriving in the sulfidic groundwater ecosystems can support rich and diverse invertebrate communities.

Experimental studies of cave animals are not very numerous, but even more limited are studies of sulfidic cave-dwelling animals. This is likely because sulfidic caves are rare, and they are typically hostile environments that are difficult to enter (Engel 2007; Sarbu et al. 2024). Furthermore, even if collection is successful, reproducing a given sulfidic environment in a lab setting such that the animals can be worked on is even more challenging. The majority of studies of sulfidic populations have been either morphological, taxonomic, or phylogenetic, as samples are difficult to acquire, and the species are not well suited for laboratory-based studies (Engel 2007). Despite these challenges, *Poecilia mexicana*, a cave fish that inhabits multiple environments, including surface waters, caves, and sulfidic waters, can be used to compare sulfidic cave populations versus non-sulfidic cave populations (Tobler et al. 2018). Cave populations showed less pigmentation and smaller eyes, whereas sulfidic populations showed larger heads, larger gill filaments, and differences in brain anatomy (Tobler et al. 2008; Schulz-Mirbach et al. 2016). Furthermore, both sulfidic environments and cave environments were associated with fewer but larger offspring (Riesch et al. 2010). In addition to morphological and life history studies, transcriptomic studies of cave and surface populations, both sulfidic and non-sulfidic, have also been performed for *P. mexicana* (Kelley et al. 2012; Passow, Henpita, et al. 2017; Passow, Brown et al. 2017; McGowan et al. 2019). An additional genus with sulfidic populations that have been investigated is the amphipod *Niphargus* (Flot et al. 2010; Fišer et al. 2015).

Another species that has the potential to shed insight into the biology of sulfide-adapted animals is *Asellus aquaticus*, a pan-European freshwater isopod crustacean with both surface and cave forms or ecomorphs. This represents an emerging model for studies of eco-evo as well as evo-devo biology (Protas and Jeffery 2012; Lafuente et al. 2021; Protas et al. 2023). There are multiple cave populations of the *A. aquaticus* species complex, including one known population from Movile Cave, a sulfidic cave in southeastern Romania. Much work has been done on the phylogeography of these populations (including both sulfidic and non-sulfidic cave populations) as well as comparative morphology of adult specimens (Verovnik et al. 2004; Konec et al. 2015; Bakovic et al. 2021), identifying at least four clades and sets of characteristics common to many cave populations, including eye loss, pigment loss, and increased relative length of antenna II (the longer of the two antennae in isopods). Laboratory studies involving breeding of different populations have mostly focused on populations from Slovenia, specifically the Old Subterranean Pivka and Subterranean Rak populations from Postojna-Planina Cave System (PPCS), Sweden, and Croatia, all of which are non-sulfidic caves (Protas et al. 2011; Re et al. 2018; Bakovic et al. 2021; Lukić et al. 2024). Genomic information exists from non-sulfidic cave populations including a genetic map generated from the Old Subterranean Pivka population, a draft genome generated from the Swedish Lummelunda Cave population, and an unannotated genome recently published from a surface individual from Great Britain (Protas et al. 2011; Bakovic et al. 2021; Thomas and University of Oxford and Wytham Woods Genome Acquisition Lab, Natural History Museum Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team, Wellcome Sanger Institute Scientific Operations: Sequencing Operations, Wellcome Sanger Institute Tree of Life Core Informatics team, Tree of Life Core Informatics collective and Darwin Tree of Life Consortium 2025). In addition, genomic regions responsible for eye and pigmentation phenotypes were mapped (Protas et al. 2011; Bakovic et al. 2021). Comparative embryological studies were done on the Subterranean Rak population (Mojaddidi et al. 2018), finding that cave embryos were larger and never developed pigmentation. Transcriptomic studies were performed on the Old Subterranean Pivka and Subterranean Rak populations (Stahl et al. 2015; Gross et al. 2019; Lomheim et al. 2023) and on the Molnár János Cave population (Pérez-Moreno et al. 2018), identifying many candidate genes involved in pigmentation and phototransduction. Though many tools were generated for different cave populations of *A. aquaticus*, the sulfidic cave population has been historically difficult to work with, and few tools have been generated.

A troglomorphic population of the *A. aquaticus* species complex inhabits a sulfidic groundwater ecosystem in southeastern Romania. The discovery of Movile Cave in 1986 near the town of Mangalia, close to the shore of the Black Sea, provided a subterranean window of access to this mesothermal (21°C) aquifer that occupies a surface area of 50–100 km². Of the 51 invertebrate species that inhabit this unusually rich and diverse cave ecosystem, 37 are endemic (Sarbu et al. 2019; Brad et al. 2021). One of these species, originally called *A. aquaticus infernus* (Turk-Prevorčnik et al. 1998), has been found in Movile Cave as well as in several sulfidic springs and old drinking wells in the town of Mangalia. In agreement with the results by

Summary

- Development of *Asellus infernus* as a model for understanding adaptation to cave and sulfidic environments.
- Comparative embryology analysis of *A. infernus* and a surface population of *Asellus aquaticus*.
- Generation of cave and surface transcriptomes, differential and allele-specific expression analysis.

Konec et al. (2015), it was proposed by Protas et al. (2023) that the population of the Mangalia sulfidic groundwater ecosystem was a good species and taxonomic status would be changed to: *Asellus infernus* (Turk-Prevorčnik et al. 1998). Note that we categorize *A. infernus* as a cave ecomorph because the individuals are depigmented and eyeless. However, *A. infernus* can be found in multiple subterranean environments, not all of which are caves, such as wells and upwellings from underground environments in springs or lakes. Comparative morphology studies have shown that *A. infernus* has extreme phenotypes of eye and pigment loss similar to those seen in populations from PPCS in Slovenia (Turk-Prevorčnik et al. 1998). In addition, in adults, relative antenna II length is greater in the cave population as compared to the surface population (Konec et al. 2015). Furthermore, the size of wild-caught adults of *A. infernus* was larger than the size of wild-caught adults from a nearby surface population (Herczeg et al. 2023). Despite similarities with the populations from PPCS in Slovenia in eye and pigment morphology, multiple other characters have shown differences between PPCS Slovenian populations and *A. infernus* (Konec et al. 2015). Furthermore, comparative phylogeography has shown that the Romanian populations are in a separate phylogenetic group distinct from the Slovenian cave populations (Konec et al. 2015). *A. infernus* has been historically difficult to acquire and work with, and little has been done that requires laboratory rearing or access to embryonic samples of these populations aside from a recent study generating F1 hybrids from this population to study the origin of an orange phenotype revealed within surface populations and within crosses between surface and cave populations (Rodas et al. 2023). Before the work described here, methods of working with and raising embryos outside of the female were not developed for *A. infernus*, which is one of the first steps critical in the development of a species as an evo-devo model. In addition, no sequencing information existed for this unique population.

Here, we established *A. infernus* as a model for understanding adaptations to sulfidic and cave environments, investigated the timeframe that morphological differences between cave and surface ecomorphs are established, examined the interplay between genetic and environmental inputs into morphological differences, and identified genes and pathways that might be different between this sulfidic cave population and the nearby surface population. Specifically, we asked the following questions:

- Are differences in lack of pigmentation and body size between cave and surface ecomorphs established by the end of embryonic development?

- Do wild-caught and lab-bred individuals show differences in morphological characteristics?

- Are genes and pathways involved in phototransduction, pigmentation, and sulfide detoxification differentially expressed and/or show allele-specific expression between cave and surface samples?

To address these questions, we performed comparative embryology/juvenile analysis between cave and surface individuals (both wild-caught and lab-bred) of *A. infernus* and a nearby surface population. Second, we generated transcriptomes from the cave and surface populations and performed differential expression and allele-specific expression studies to examine genes and pathways that were different in cave and surface samples.

2 | Methods

2.1 | Animals

In the following population codes, CAVE refers to the cave ecomorph and SURF to the surface ecomorph. The two-letter code after the underscore is an acronym for the collection location. The cave ecomorph, *A. infernus*, was collected from old drinking wells within the town of Mangalia in May 2021 (individuals from both wells were pooled and abbreviated CAVE_mw) and also from Karaoban Lake in August 2015 (abbreviated CAVE_kl). The wells are located near each other (0.5 km), but Karaoban Lake is located 2.7 km away from the wells. The wells represent a habitat that is fed by sulfidic water through fractures in the bedrock. Theoretically, fauna could move from location to location, but might be limited if the conduits lack dissolved O_2 or are not large enough. Therefore, it is possible that there might be some cryptic diversity within the cave ecomorph. However, the samples that were used for the gene expression analyses (detailed below) were from the two wells; they are likely to be part of the same breeding population because of their close geographical location, though molecular analysis comparing individuals from different wells has not been performed. Surface individuals of *A. aquaticus* were collected from two nearby locations, Turkish Bath (July 2018; SURF_tb) and Limanu Bridge Stream (May 2021; SURF_li), both of which do not have sulfidic waters (Table S1). These two locations are located at a distance of 2.1 km on the shore of the same pond and are therefore likely to be part of the same breeding population. Animals of each ecomorph were collected from multiple, but nearby, locations because individuals were not found at the previous location at different collection times. In the methods, we refer to the animals by the code above, so it is clear which populations were used for which assays, but in the "Results and Discussion," we use cave or *A. infernus* to include both CAVE_mw and CAVE_kl and surface to include both SURF_tb or SURF_li. Note that both *A. infernus* and *A. aquaticus* are part of the *A. aquaticus* species complex (Protas et al. 2023). For the populations we examined in this study, *A. aquaticus* refers to the Romanian surface population. However, there are subterranean populations of *A. aquaticus* from other locations, for example, in Slovenia.

We raised animals similarly to what was previously described (Protas et al. 2011), but we added algae pellets instead of decaying leaves as food, as algae pellets are easier to obtain and have less potential of introducing biological contaminants. Note that these conditions are very different than the animals' natural conditions, as we gradually changed the water that the animals were collected into synthetic freshwater (Protas et al. 2011) for both ecomorphs and did not keep the cave ecomorph in sulfidic water conditions. In addition, the food source of algae pellets was also very different than the natural food source, particularly of the cave ecomorph. Individuals from the parental populations, CAVE_mw, CAVE_kl, SURF_tb, and SURF_li, were raised separately in plastic containers (700 mL capacity) containing around 10–15 individuals. When juveniles from wild-caught individuals were found within a tank, they were separated into their own tank and allowed to grow to adulthood.

2.2 | Generation of F1 Hybrids

We set up F1 hybrid crosses by placing a CAVE_mw male and 2–3 SURF_li or SURF_tb females in a single container. Three F1 hybrid crosses were used for the transcriptomic samples described below, and five additional F1 hybrid broods were used for the juvenile comparisons. F1 hybrid crosses in the other direction (CAVE_mw females and SURF_li or SURF_tb males) were also set up, but all failed to produce offspring. CAVE_mw females very infrequently generate embryos in the laboratory, even with males from the same population, so the lack of successful crosses between CAVE_mw females and SURF_li males is likely not due to a barrier or incompatibility but rather to the low success of CAVE_mw females reproducing in the lab. Ovigerous SURF_li or SURF_tb females with hybrid offspring were separated, and F1 hybrid embryos (SURF_li_CAVE_mw_F1) were removed as described below.

2.3 | Comparative Embryology

2.3.1 | Embryo Collection and Imaging

We anaesthetized females with embryos in a solution containing 50 mL of synthetic freshwater (Protas et al. 2011) and 20 µL of clove oil. Then, each female was washed twice in synthetic freshwater, and the embryos were removed by opening the brood pouch gently with forceps and flushing out the embryos with a glass pipette. The embryos were moved into a new Petri dish with synthetic freshwater and kept at 12°C. Embryos were monitored daily and photographed using a Leica S8 Apo Microscope. Embryos that were used for RNAseq were harvested at either 70% through embryonic development, in which the embryos have a comma shape and will shortly show eye pigmentation in the surface form, or 90% of the way through embryonic development, when eye and head pigmentation is visible and the body of the embryo is elongated (Mojaddidi et al. 2018).

We collected embryos from gravid CAVE_mw females upon arrival in the lab or females that mated immediately after arrival to the lab and then raised the embryos in vitro as described above. SURF_li females mated during transport to the lab or immediately after arrival at the lab. We then removed the embryos and raised them in vitro. We called these sets of embryos "wild-caught" as they were from wild-caught females that either mated in the wild or immediately upon arrival to the lab. Our initial comparisons of cave and surface juveniles and embryonic development used these wild-caught embryos because we did not know whether the cave animals would breed in the laboratory environment. Fortunately, we were able to raise juveniles to adulthood in the lab and generate gravid females from these lab-reared juveniles for both CAVE_mw and SURF_li populations. We labeled juveniles from these lab-reared embryos as "lab-bred." F1 hybrid embryos, generated as detailed above, could not be classified as either wild-caught or lab-bred as they were generated from wild-caught cave males and either wild-caught surface females or lab-bred surface females around a month after the animals were brought into the lab from the wild. We examined pigmentation as a presence versus absence trait in all embryos. Ideally, we would also have examined the presence versus absence of eyes (ommatidia) in embryonic development, but this requires visualization methods that often result in embryo death, and our sample sizes were too limited to successfully view enough individuals.

2.3.2 | Comparative Juvenile Analysis

For the comparative juvenile analysis, our goal was to harvest juveniles when they were first able to walk and had the appearance of miniature adults. We selected this timepoint as it is easier to flatten individuals for measurements when their appendages can be easily removed. Embryos were observed every day, and once they were able to walk, they were harvested in 100% ethanol and placed at −20°C until mounting on microscopic slides. Juveniles that were selected for mounting had at least one intact antenna II. As individuals were raised in a dish and were harvested right when they were able to walk, antennae were generally intact. To ensure that juveniles could be flattened in ethanol, all pereopods were removed under a Leica S8 Apo stereomicroscope. Some individuals also had to have their mouth parts removed to allow further flattening. After limb removal, the juvenile was washed in 50% ethanol, 100% Milli-Q water, and 50% glycerol, and then placed on a microscope slide with the dorsal side up in 100% glycerol. On the left- and right-hand sides of the juvenile, a coverslip was placed with a single drop of glycerol to create a shallow wall around the juvenile so that when a coverslip was placed over the juvenile, the third coverslip rested on top of the first two coverslips and did not crush the juvenile. Juveniles were measured using LAS Core Software (Leica, Wetzlar, Germany) on a Leica S8 Apo stereomicroscope. Antenna II length (left and right if available), head width, and body length were measured. Five groups of samples were measured CAVE_mw_wild samples, SURF_li_wild samples, CAVE_mw_lab samples, SURF_li_lab samples, and SURF_li_CAVE_mw_F1 hybrid samples (Supporting Information File 1 and Table S2). Each juvenile was photographed and measured separately by two different people using LAS Core Software. The following measures were analyzed: left and right antenna II length, body length, and

head width. The measurements from the two different people were averaged. Again, pigmentation was examined as a presence versus absence trait in all embryos.

2.3.3 | Statistical Analysis

First, we performed the Shapiro–Wilk test in R 4.1.3 (R Core Team 2022) to assess normality of measurements. Measures which did not deviate significantly from normality included antenna II lengths (left and right), body length, and head width. For these measures, ANOVA and Tukey's Honestly Significant Difference (HSD) tests for pairwise comparison were performed for each group. To address questions of allometry, we also performed principal components analysis on the measurements of body length, head width, and left and right antennae length, followed by an ANOVA and Tukey's HSD test on the resulting first and second principal component scores across groups. The Bonferroni correction for multiple testing was performed for tests of comparative juvenile analysis.

2.4 | Comparative Transcriptomics

2.4.1 | Samples

One brood of CAVE_kl embryos and one brood of SURF_tb embryos were prepared for Iso-seq to obtain full-length transcripts with the goal of combining this sequence ultimately with Illumina sequence to obtain more complete transcriptomes (Table 2). RNA extraction was performed as described (Lomheim et al. 2023), and the sequencing was performed on a PacBio machine by the Functional Genomics Lab, Vincent C. Coates Genomics Sequencing Laboratory, and UC Berkeley. For Illumina sequencing, samples were collected more recently, and individuals at the SURF_tb and CAVE_kl locations were not present, so nearby populations were used instead, CAVE_mw and SURF_li. Three CA-VE_mw broods and three SURF_li broods at two different developmental timepoints, that is, mid-stage (ca. 70%) through embryonic development and late stage (ca. 90%) through embryonic development, were harvested. Additionally, three broods of late-stage F1 hybrids (SURF_li_CAVE_mw_F1) were harvested (Table S3). Each of these samples consisted of a single brood of siblings. After harvesting, all broods were homogenized in 200 µL of TRIzol (Thermofisher, Waltham, Massachusetts, the United States). Sequencing libraries were prepared for each of these samples: three CAVE_mw mid-stage, three CAVE_mw late-stage, three SURF_li mid-stage, three SURF_li late-stage, and three late-stage SURF_li_CAVE_mw_F1. RNA was extracted using the RNeasy Plus Universal Mini Kit by the Genetic Epidemiology and Genomics Lab (GEGL), UC Berkeley. PolyA selection was performed, and libraries were constructed using the Nugen Kit low-input protocol. Sequencing was performed on a HiSeq. 4000 machine with 150 bp paired-end reads by the Functional Genomics Lab, Vincent C. Coates Genomics Sequencing Laboratory, and UC Berkeley.

2.4.2 | De novo transcriptome assembly and annotation

Two transcriptomes were made, one SURF_li transcriptome and one CAVE_mw transcriptome. First, Iso-seq sequences were run through the in-house Iso-seq pipeline of the Functional Genomics Lab, Vincent C. Coates Genomics Sequencing Laboratory, and UC Berkeley, which included the synthesis of CCS sequences, demultiplexing of sequences, primer polyA and concatamer removal, clustering, and polishing. The final transcripts were combined with the assemblies generated from Illumina reads. We used the same methods as in a previous study by Lomheim et al. (2023). Briefly, for the Illumina samples, the reads were trimmed with Trimmomatic using the following parameters: sliding window 4:24, headcrop 10, avgqual 30, and minlen 30. Then, the transcriptomes were generated using the de-novo-transcriptome-assembly pipeline of the National Center for Genome Analysis Support (NCGAS) (https://github.com/NCGAS/de-novo-transcriptome-assembly-pipeline), which incorporates multiple assemblers: SOAP version 1.03 (kmer 35, 45, 55, 65, 75, and 85; Xie et al. 2014), TransAbyss version 2.0.1 (kmer 35, 45, 55, 65, 75, and 85; Robertson et al. 2010), Trinity 2.11.0 using default parameters (Grabherr et al. 2011), and Velvet 1.2.10 (kmer 35, 45, 55, 65, 75, and 85; Zerbino and Birney 2008). All the assemblies plus the Iso-seq sequences were run through Evidential Gene's tr2aacds pipeline (Gilbert 2013). Recently, a genome was published for an individual of *A. aquaticus* from Wytham Woods, Great Britian (Thomas and University of Oxford and Wytham Woods Genome Acquisition Lab, Natural History Museum Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team, Wellcome Sanger Institute Scientific Operations: Sequencing Operations, Wellcome Sanger Institute Tree of Life Core Informatics team, Tree of Life Core Informatics collective and Darwin Tree of Life Consortium 2025). We mapped the de novo transcriptomes to the published genome in quality control (minimap2, Li 2018), but no annotation was published with this genome. As a result, once transcriptomes were generated, we identified homologous sequences to the *Tribolium castaneum* annotation from 2019 (file entitled GCF_000002335.3_Tcas5.2_protein.faa) https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/335/GCF_000002335.3_Tcas5.2/ using Blast2GO (Conesa et al. 2005), allowing us to transfer annotation information from *Tribolium* to our de novo transcriptomes. Transcriptomes were evaluated using Galaxy Version 2.10.1+galaxy2 with BUSCO version 5.3.2 (Arthropoda selected; Simão et al. 2015) and QUAST version 5.2.0 (eukaryote selected; Gurevich et al. 2013). XSEDE resources were used (Towns et al. 2014).

2.4.3 | Differential Expression of Mid- and Late-Stage Embryos

Differentially expressed genes between surface and cave samples at the mid-stage and late-stage timepoints were examined. Differential expression was assayed as previously discussed (Lomheim et al. 2023). Briefly, trimmed reads were quantified against both transcriptomes with the default parameters from Kallisto (Bray et al. 2016). The Kallisto output of estimated counts was combined into four matrices: all mid-stage embryo

samples mapped against CAVE_mw, all mid-stage embryos mapped against SURF_li, all late-stage samples mapped against CAVE_mw, and all late-stage samples mapped against SURF_li. Differential expression analysis was performed using DESeq2 (Love et al. 2014). DESeq2′s shrinkage estimation and dispersion modeling provide more stable results for analyses with fewer biological replicates by using empirical Bayes estimators. These estimators pool information across genes to gain extra information and provide more accurate estimations of dispersion in cases where there might be artificially high variance due to sample size. DESeq2 also models the likelihood of data capture using a negative binomial distribution model, which also helps combat overdispersion in smaller sample sizes and heterogeneous data such as field samples of non-model species.

Pairwise DESeq2 analysis was conducted on the four matrices described above using Trinity′s differential expression perl wrappers (run_DE_analysis.pl and analyze_diff_expr.pl, Haas 2023). This wrapper filters out transcripts with 0 CPM in any replicate, as this is ambiguously missing data, which could result from a lack of capture (technical bias) or a lack of expression (real variation). Normalization was also handled in the DESeq() function, using DESeq2′s default mean of ratios method to normalize the count data and adjust for library size effects and compositional biases. Wald tests were used to evaluate the significance of the difference between log-fold changes. Multiple test correction was done with False Discovery Rate via DESeq2′s workflow. Signal and noise in the normalized, filtered data were then visualized in a principal component analysis (PCA) of count data (Figure S1). The first two PCs explained 67.38% of the variation in the data and showed reasonably cohesive grouping of samples (in the context of expected noise due to variation in field samples) with distinct distances between the replicates (signal of different expression). Finally, we classified genes as significantly differentially expressed if they had a log2Fold change of > 2 (minimum 4× change in expression) and FDR < 0.05. We chose a moderate log fold change cutoff to guard against false positives and an FDR of 0.05 to account for our more limited statistical power. These cutoffs were further evaluated against the volcano plot to ensure we removed obviously spurious results, such as those with high fold change but no statistical support.

To unify the information from the four pairwise comparisons (mid-stage to both transcriptomes and late-stage to both transcriptomes), we needed to identify homologous loci in cave and surface transcriptomes to compare expression between the cave and surface samples. Each transcriptome had unique names as they were de novo assembled separately; there is no published annotation for this genome, and the gene name annotation transferred from *Tribolium* was not unique (i.e., multiple genes in our transcriptomes could have the same *Tribolium* ID). To establish a set of homologous sequences between the cave and surface (cave/surface pairs), a reciprocal blast technique was used. CAVE_mw and SURF_li databases were created using the web-based platform European Galaxy NCBI BLAST+ (Galaxy Version 2.10.1+galaxy2; Camacho et al. 2009; Cock et al. 2015). The CAVE_mw transcriptome was then blasted against the SURF_li database, and vice versa, and we retained the longest and highest hit for each transcript in each blast run. As these are the same species, it is expected that there will be high homology between assembled transcripts. However, due to

limitations of de novo assembly and sequencing a subset of transcripts in a sample, not all transcripts will have a complete assembly or a homologous pair retained. As such, transcripts under 400 bp and lacking a reciprocal hit were removed. A database of 15,516 cave/surface pairs resulted, with entries including the CAVE_mw transcript name and sequence, the differential expression results for that sequence (ref: CAVE_mw), the SURF_li transcript name and sequence, and the differential expression results for that sequence (ref: SURF_li).

Further filtering of possible differentially expressed genes was applied. Transcripts that had unique *Tribolium* IDs were retained as those were clearly unique sequences with clear homology to a related species. Conversely, we eliminated any transcripts that shared a *Tribolium* annotation with 9 or more other assembled transcripts in that transcriptome, rationalizing that a large number of hits to the same *Tribolium* gene were more likely to be fragments or chimeras rather than large numbers of unique paralogues. To further filter out any duplicates, we self-blasted the transcripts against their own database and removed any expressed duplicates (identity $> 90\%$, length > 150 bp, and CPM > 1). This setup allowed for the use of VLOOKUP in Excel to identify which high-confidence transcripts showed an LFC > 2 and FDR < 0.05 against both reference transcriptomes.

To perform GO enrichment, we compared this filtered set of differentially expressed transcripts against the background of the full list of genes output from DESeq2, minus any loci that shared the same Tribolium ID with 9 or more transcripts (low confidence assemblies). This was done separately for the output of mid-stage comparisons and late-stage comparisons, generating a mid-stage reference list and a late-stage reference list. Then, we used G-profiler to analyze GO enrichment at each timepoint (Raudvere et al. 2019).

2.4.4 | Allele-Specific Expression Analysis

Allele-specific expression was performed as previously by Lomheim et al. (2023). First, cave/surface transcript pairs were generated, as described above, yielding 15,516 genes with cave/surface transcript pairs. The cave and surface sequences were combined (combined transcripts) using merge_pat_mat_fasta.pl script from ASE-Tigar (Nariai et al. 2016). This does not combine the actual sequences but retains two sequences for each gene, and the reads are mapped to each of the sequences.

The trimmed paired-end reads of all SURF_li, CAVE_mw, and SURF_li_CAVE_mw_F1 samples were mapped to the combined transcripts using bowtie2, -X 1000 -k 100 –very-sensitive, and ASE-TIGAR (Langmead and Salzberg 2012). Z values from ASE-Tigar are a measure of transcript abundance. The ratio of Z mapped to each surface transcript to the corresponding cave transcript was compared for each sample. Genes were selected where the surface samples mapped to the surface transcript 3.3× more than the corresponding cave transcript and where the cave samples mapped to the cave transcript 3.3× more than the corresponding surface transcript. Within this subset of genes, we selected genes where all three F1 hybrids showed a bias towards one allele, mapping to either the cave transcript

3.3× more than the surface transcript, or vice versa. Transcripts that were not annotated or where the *Tribolium* ID was present in 10 or more copies in the transcriptome were removed.

Next, to identify fixed SNPs that were different between cave and surface transcripts, all trimmed, paired-end reads of CAVE_mw, SURF_li, and SURF_li_CAVE_mw_F1 samples were mapped to combined transcripts using bowtie2 (default settings plus very fast end-to-end) and Freebayes (Garrison and Marth 2012). Mapping SURF_li reads to the CAVE_mw transcripts and reciprocally mapping CAVE_mw reads to the SURF_li transcripts identified fixed SNPs between the cave and surface transcripts. All fixed SNPs were retained that were present with at least five observations.

For all the above genes prioritized by ASE-Tigar, genes were further prioritized that also showed differential expression. For this subset of genes, we performed FreeBayes allele counting in the F1 hybrids using the fixed SNPs outlined above (Garrison and Marth 2012). To eliminate bias from the reference used, allele counting was performed both when the cave and the surface transcripts were used as a reference. Five SNPs, if available, were selected across the span of a transcript for allele counting in the F1 hybrids. Most of the time, the location of these SNPs was the same in the cave versus the surface transcript. However, if an indel was present, the location of the SNPs might be offset in the two transcripts, and alignments were performed if it was unclear where the location of the SNPs was.

For each SNP, a binomial test was performed to investigate when the allele counts deviated from the predicted fraction of 0.5. To control for multiple testing, the resulting p values were adjusted using the Benjamini–Yekutieli procedure (Benjamini and Yekutieli 2001). Any gene where at least two of the five SNPs showed adjusted p values < 0.05 for all three F1 samples to both the cave and surface transcript was identified as having allele-specific expression.

3 | Results

3.1 | Comparative Embryology

3.1.1 | *Asellus infernus* Embryos Were Depigmented Throughout Embryonic Development

We first asked when morphological differences were established in *A. infernus* as compared to the surface form, embryonically or postembryonically. *A. infernus* adults collected from wells in Mangalia (cave) were depigmented and eyeless as compared to *A. aquaticus* adults collected from the Limanu Bridge Stream population (surface) (Figure 1). We compared embryonic development of wild-caught gravid cave females and wild-caught surface females, which became gravid upon transport to the lab or immediately after arriving at the lab. Regarding pigmentation, *A. infernus* embryos never developed pigmentation, whereas surface embryos developed eye pigmentation at around 80% of the way through embryonic development (Figures 2, S2, and S3) and shortly after, developed head and body pigmentation. Additionally, F1 hybrid embryos from *A. infernus* males crossed to surface females did develop eye pigmentation, similar to the surface form (Figure 2). Furthermore, adult F1s had both head and body pigmentation and ommatidia (Figure 1D–F). Therefore, we established that pigmentation was never seen embryonically in *A. infernus,* rather than initially being present and then lost later in development.

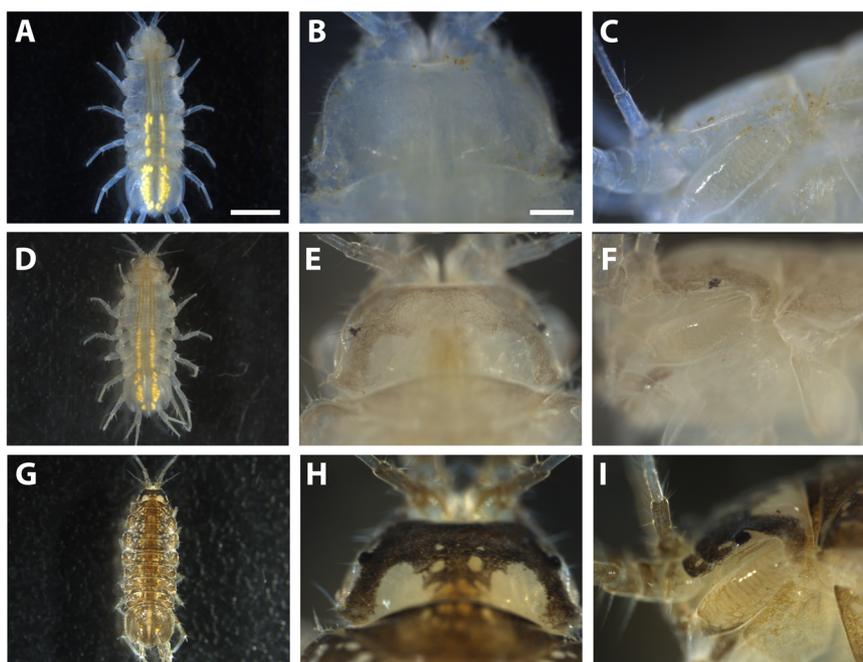**FIGURE 1** | Comparison of adult morphology between cave, F1 hybrid, and surface individuals. (A, B, C) Cave individual. (D, E, F) F1 hybrid individual. (G, H, I) Surface individual. (A, D, G) Whole body. (B, E, H) Head. (C, F, I) Profile. In A, the scale bar is 2 mm and is relevant for A, D, and G. In B, the scale bar is 200 μm and relevant for B, C, E, F, H, and I. [Color figure can be viewed at wileyonlinelibrary.com]
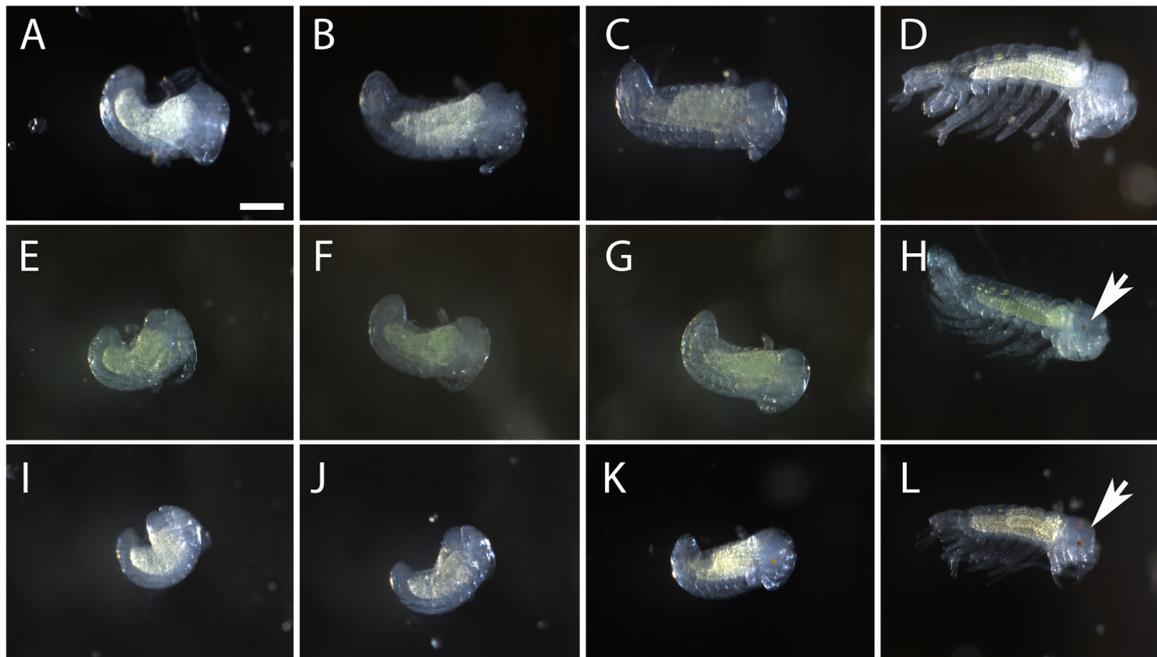
FIGURE 2 | Comparative embryology between cave, F1 hybrid, and surface embryos. (A–D) Cave embryos. (E–H) F1 hybrid embryos. (I–L) Surface embryos. (A, E, I) Embryos at approximately 70% of the way through embryonic development (mid-stage embryos). (B, F, J) Embryos at approximately 75% of the way through embryonic development. (C, G, K) Embryos at approximately 80% of the way through embryonic development. (D, H, L) Embryos at approximately 90% of the way through embryonic development. White arrow: Pigmented eye can be seen in H and L. Scale bar is 200 µM for all panels. [Color figure can be viewed at wileyonlinelibrary.com]

3.1.2 | Lab-Bred *Asellus infernus* Juveniles Had Longer Body Length and Longer Antennae II Than Lab-Bred Surface *Asellus aquaticus* Juveniles

To further address the question of the timing of the establishment of morphological characteristics, we investigated size characteristics between lab-bred *A. infernus* and surface individuals. For this comparison, we examined lab-bred *A. infernus* and surface individuals that were raised in the same conditions to eliminate, as much as possible, any environmental variables. Specifically, we compared antenna II length, body length, and head width in lab-bred surface and cave juveniles as soon as they were able to walk and would have been released from the brood pouch (Figure 3 and Table S4). For body length, antenna II length, and head width, cave samples were longer compared to surface samples ($p < 0.0005$ for all comparisons; Figure 3 and Table S4). Overall, we conclude that lab-bred cave and surface juveniles are different in body size, head width, and antenna size, and that these differences, many of which are also present in comparisons between adult cave and surface individuals, are established by the end of embryonic development.

3.1.3 | Population-Environment Interactions (Wild-Caught Vs. Lab-Bred) Influenced Several Measurements' Differences Within Both Cave and Surface Samples

To investigate whether environmental changes also influence these morphological differences, we compared wild-caught and lab-bred juveniles. For the surface wild-caught to surface lab-bred comparisons, two measures showed significant differences:

head width and body length ($p < 0.0005$ for both; Figure 3 and Table S4). For the cave wild-caught to cave lab-bred comparisons, only body length showed a significant difference ($p = 0.00073$). To ensure that differences in these measurements were not simply reflections of distinctions in overall size, we performed PCAs on the measurements of body length, head width, and left and right antennae length (Figure 4 and Table S5). The first two principal components together explained approximately 95% of the total variation, so further components were not considered. The first principal component (PC1) corresponded roughly to an average of all four body measurements and therefore can be interpreted as a general measure of overall size: large positive PC1 scores are associated simultaneously with longer body lengths, antennae, and wider heads, while more negative PC1 scores are associated with smaller measurements overall. The second principal component (PC2) corresponded to a contrast between antennae length and head and body measurements: highly positive PC2 scores are associated with longer antennae but shorter body lengths overall, and more negative PC2 scores were associated with longer bodies but shorter antennae. Figure 4 displays schematic archetypes of isopods with extreme values in each of these two directions of variation.

We performed ANOVA using juvenile's specific combination of population and environment as a predictor for each of the two principal component scores, which are uncorrelated by construction, ensuring that the differences in contrasts between antennae relative to body measurements (i.e., PC2) are not simply due to differences in size (i.e., PC1). Pairwise comparisons across subgroups were assessed using Tukey's HSD test, and p values were reported after adjusting for multiple testing (Table 1).
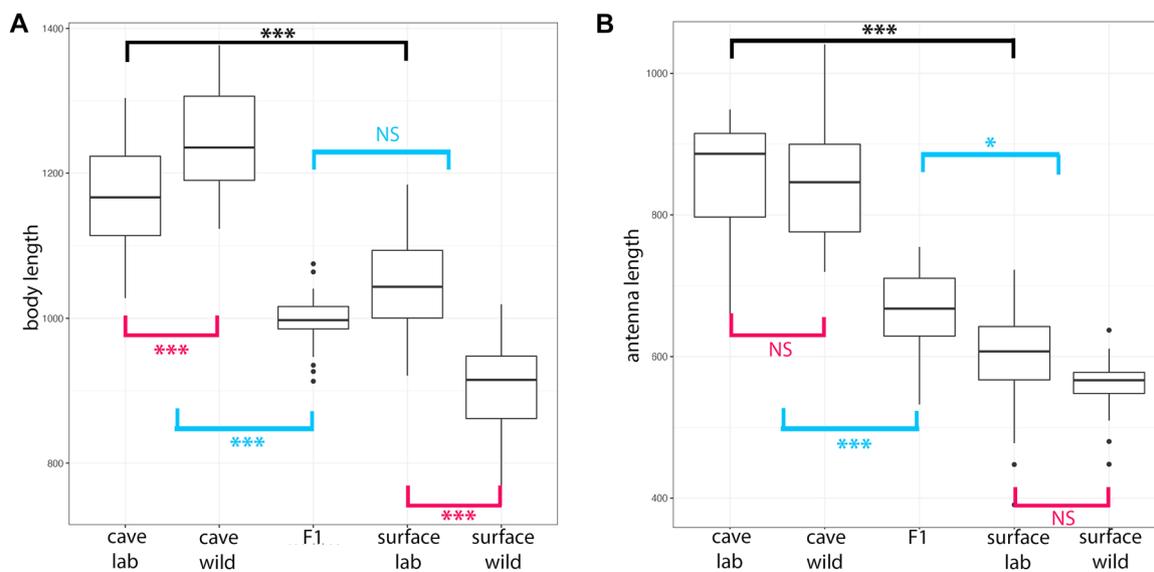
FIGURE 3 | Comparison of body and antenna length for the comparisons: cave lab versus surface lab, wild versus lab, and F1 versus cave and surface. "Lab" indicates juveniles from females that were raised from juveniles in the lab and bred in the lab. "Wild" indicates juveniles that were collected from wild-caught individuals with embryos. F1 are F1 hybrids between surface females and cave males. (A) Body length of the juvenile. (B) Size of left antenna II. Black bars and asterisks are for cave lab versus surface lab comparison. Magenta bars and asterisks are for wild versus lab comparisons. Blue bars and asterisks are for F1 versus cave and lab comparisons. Table S4 holds p values of statistical tests of all possible comparisons. ANOVA and Tukey's HSD test were performed for antenna length, body length, and head width. ***$p < 0.0005$, **$0.0005 < p < 0.005$, *$0.005 < p < 0.05$. $p > 0.05$ was labeled NS. [Color figure can be viewed at wileyonlinelibrary.com]

Significant differences in overall size (PC1) were detected for all comparisons except between cave lab-bred versus cave wild-caught. A detailed look at all comparisons together suggests that cave specimens were larger overall than surface specimens. However, size differences between lab-bred and wild-caught subgroups were not so straightforward and depended on whether the isopod was a surface or cave-dwelling specimen. For surface individuals, lab-bred subjects were generally larger than their wild-caught counterparts across the board ($p = 0.00003$), while for cave individuals, wild-caught specimens were on average larger than lab-bred individuals, though this difference was not significant ($p = 0.4604$).

Significant differences in contrasts between antennae and body length (PC2) were also found in most comparisons, except for those across cave wild versus surface lab-bred subgroups and cave lab-bred versus surface wild subgroups. No marginal cave versus surface trend emerged in this direction; rather, differences in PC2 scores were dependent on the interaction between surface/cave environments and lab-bred/wild breeding. The cave lab-bred and surface wild-caught groups had positive PC2 scores (i.e., longer antennae and shorter body lengths, while controlling for overall size), while cave wild-caught and surface lab-bred groups generally had negative PC2 scores (i.e., shorter antennae and longer body lengths, while controlling for overall size). See Figure 4 for visualization of group morphologies and Table 1 for specific pairwise comparisons and corresponding p values.

3.1.4 | F1 Hybrid Samples Were Intermediate and Different in Most Size Measures Between Cave and Surface Individuals

Because cave and surface forms can interbreed, we had the opportunity to compare the above size characteristics in F1

hybrids with the goal of seeing whether the F1 hybrids were similar to *A. infernus*, to surface, or intermediate between the two. Our expectation was that the F1 hybrids would be intermediate between the two measures. The F1 samples were significantly different from both cave lab-bred and wild-caught samples for the measures of head width, body length, and antenna length ($p < 0.0005$ for all measures; Figure 3 and Table S4). The F1 individuals were different for both surface lab-bred and wild-caught samples for only antenna length ($p = 0.0153$ for surface lab and $p < 0.0005$ for surface wild; Figure 3 and Table S4).

3.2 | Transcriptomic Analysis

3.2.1 | Different mRNA Levels Were Seen in Genes Involved in Phototransduction and Sulfide Detoxification in Cave and Surface Samples

Next, we investigated the question of what genes showed expression differences between *A. infernus* and surface samples. First, we generated transcriptomes for the *A. infernus* cave and *A. aquaticus* Limanu springs surface populations (Table 2). With BUSCO scores of 93%, this is similar to updated transcriptomes for the Subterranean Rak population and Rakov Škocjan of *A. aquaticus* (Lomheim et al. 2023) and transcriptomes generated from the Molnár János Cave population and its adjacent surface populations, Malom Lake and Soroksár from Hungary (Pérez-Moreno et al. 2018). Cave and surface transcriptomes had a similar number of transcripts from the previously examined populations (Table 2).

Recently, an unannotated genome for *A. aquaticus* became available (Thomas and University of Oxford and Wytham
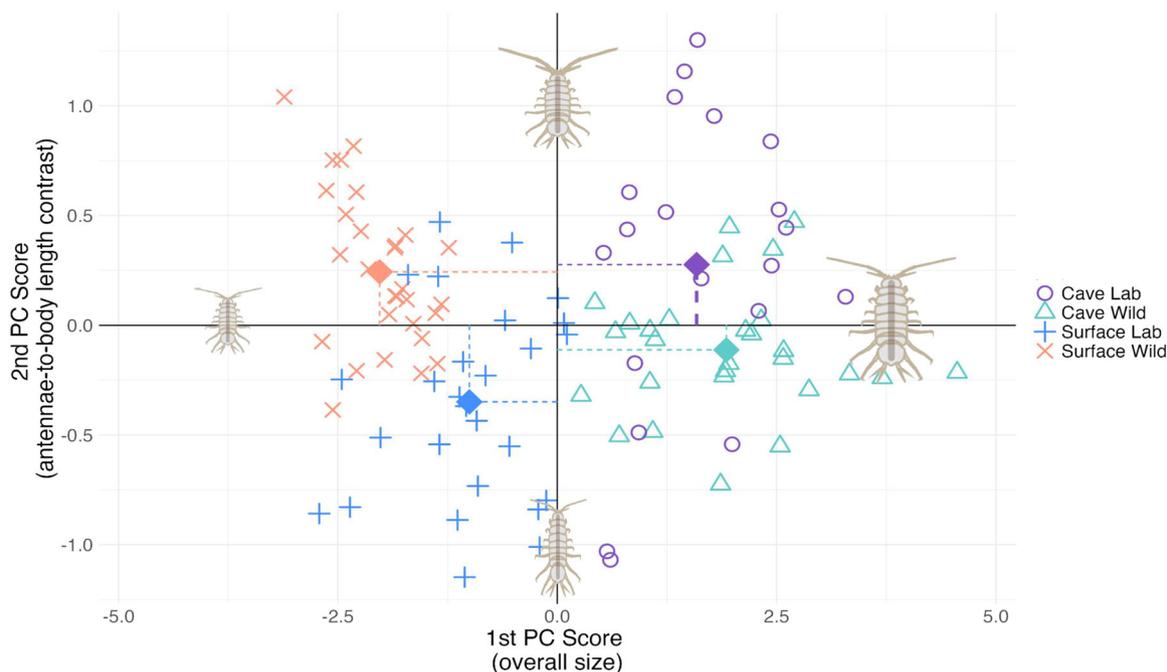
FIGURE 4 | First and second principal component scores of body measurements, which explain roughly 88% and 6.4% of total variation, respectively. Archetypal morphologies for extremes in each direction are depicted with isopod schematic diagrams, with the average for each subsample denoted by a cross. Higher first principal component scores generally correspond to larger body measurements overall and are generally held by isopods with longer body lengths, antennae, and wider heads; lower first principal component scores are associated with smaller measurements overall. Higher second principal component scores are held by isopods with longer antennae but shorter body lengths, and vice versa for isopods with lower second principal component scores. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 | Pairwise differences across environments for average first and second principal component scores evaluated with ANOVA and Tukey's HSD test, with p values adjusted for multiple comparisons.

	1st PC Score (overall size)			2nd PC Score (antennae vs. body length contrast)		
Comparison	**Diff.**	**St. Err.**	**p (adj.)**	**Diff.**	**St. Err.**	**p (adj.)**
Cave Wild–Cave Lab	0.3392	0.3389	0.4604	−0.3885	0.1887	**0.0166**
Cave Wild–Surface Wild	3.9513	0.3415	**≈0**	−0.3553	0.1707	**0.0152**
Cave Wild–Surface Lab	2.9296	0.3365	**≈0**	0.2369	0.1738	0.1952
Cave Lab–Surface Wild	3.6121	0.3123	**≈0**	0.0332	0.1873	0.9937
Cave Lab–Surface Lab	2.5904	0.3068	**≈0**	0.6254	0.1902	**≈0**
Surface Lab–Surface Wild	1.0217	0.3096	**2.76×10^{-5}**	−0.5922	0.1723	**≈0**

Note: Bold indicates significant at the 0.05 level after multiple comparison correction. Italics is used for all standard errors. Significant differences in general size (PC1 score) were detected across all comparisons except for cave wild versus cave lab subgroups. Meanwhile, differences in antennae versus body length contrasts (PC2 scores) were significant for all comparisons except cave wild versus surface lab and cave lab versus surface wild subgroups. Multiple comparison corrections for significance were performed using the Bonferroni procedure.

Woods Genome Acquisition Lab, Natural History Museum Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team, Wellcome Sanger Institute Scientific Operations: Sequencing Operations, Wellcome Sanger Institute Tree of Life Core Informatics team, Tree of Life Core Informatics collective and Darwin Tree of Life Consortium 2025), and we mapped both the cave and surface transcriptomes to the genome. For the cave transcriptome, all 88,614 transcripts had a primary alignment. Of these, 55,950 transcripts had high-quality, non-chimeric primary alignments, and another 2117 had supplementary alignments across transcripts. For the

surface transcriptome, all 80,182 transcripts had a primary alignment. 59,304 transcripts had high-quality, non-chimeric primary alignments, and another 1537 had supplementary alignments across transcripts. In all figures, tables, and files, an asterisk is placed next to any transcript that did not have a primary alignment.

We performed differential expression analysis on wild-caught embryos at two different embryonic timepoints, mid-stage and late-stage, identifying genes that were surface-biased (higher mRNA levels in surface samples as compared to cave samples) and cave-biased (higher mRNA levels in cave samples as compared to surface samples) (Supporting

TABLE 2 | Quast and BUSCO statistics for cave and surface transcriptomes.

	Surface	Cave
# contigs (≥ 0 bp)	80,182	88,614
# contigs (≥ 1000 bp)	20,989	20,380
Largest contig	46,211	45,558
Total length (≥ 0 bp)	82,452,877	88,155,484
Total length (≥ 1000 bp)	59,944,846	61,449,113
N50	2935	3088
GC (%)	37.18	36.99
# Ns per 100 kbp	1103.3	1011.18
Complete BUSCOs (%)	92.9%	93.4%
Complete and single-copy BUSCOs (%)	88.8%	89.6%
Complete and duplicated BUSCOs (%)	4.1%	3.8%
Fragmented BUSCOs (%)	1.0%	0.8%
Missing BUSCOs (%)	6.1%	5.8%

Information Files 2, 3, 4, and 5 and Figures S4 and S5). We found 166 genes that were surface-biased and 73 genes that were cave-biased at the mid-stage timepoint. At the late-stage timepoint, we found 160 genes that were surface-biased and 73 that were cave-biased. We had expected that genes involved in phototransduction would show lower mRNA levels in the cave individuals as compared to the surface samples and found that this was the case for six genes at the mid-stage timepoint and seven genes at the late-stage timepoint; some of the genes were in common between the two timepoints (Table 3). Additionally, several genes involved in melanin synthesis and the diurnal clock showed significantly different mRNA levels in cave versus surface samples.

We had also predicted that some of the genes that were differentially expressed would be in response to the sulfidic environment, such as genes involved in H_2S oxidation. As expected, we saw cave-biased differential expression in sulfide:quinone oxidoreductase and in persulfide dioxygenase ETHE1 at both the mid-stage and late-stage timepoints (Supporting Information Files 3 and 5).

GO enrichment analysis of differentially expressed genes did not highlight any categories involved with eyes, pigmentation, or sulfide metabolism. However, in the mid-stage embryos, the KEGG term of "arachidonic acid metabolism" was enriched in the genes that were surface biased (KEGG:00590, $p = 0.042$). In the late-stage embryos, the GO term "structural constituent of cuticle" was enriched in cave-biased genes (GO:0042302, $p < 0.001$). The lack of enrichment data is likely due to the comparatively small number of genes identified as showing statistically significant mRNA levels and the usage of whole embryos rather than tissue-specific RNAseq.

3.2.2 | Genes With Cave-Biased and Surface-Biased Allele-Specific Expression Were Identified Including Genes With Possible Roles in Adaptation to Sulfidic Environments

To continue examining the question of what genes showed expression differences between cave and surface samples, allele-specific expression was used. Allele-specific expression is a method of identifying genes that might have *cis*-regulatory changes. First, we used ASE-Tigar to prioritize genes with allele-specific expression (Supporting Information Files 6 and 7). In our final list of genes with allele-specific expression using FreeBayes allele counting and eliminating genes that did not also show differential expression, we identified 5 genes that were cave-biased and 11 genes that were surface-biased (Figure 5 and Supporting Information File 8).

4 | Discussion

We have generated resources for a subterranean species, *A. infernus*, that has been historically difficult to work with in the lab. We addressed the following questions: when are morphological differences established between cave and surface individuals, do both genetic and environmental inputs influence phenotype differences between cave and surface individuals, and finally, what genes and pathways are responsible for phenotype evolution between cave and surface individuals? We found that multiple morphological differences were established by the end of embryonic development: lack of pigmentation, longer body length, and longer antennae II. We also found that the origin of samples, lab-bred versus wild-caught, influenced many of the measures, both in cave and surface samples, supporting that environmental inputs, as well as genetics, influence the size phenotypes. Furthermore, we generated genomic resources for this population, including an *A. infernus* transcriptome, a surface transcriptome, and, through differential expression analysis and allele-specific expression analysis, detected differences in multiple genes involved in phototransduction and increased expression in genes involved in sulfide oxidation.

4.1 | Comparative Embryology

First, we investigated the question of when morphological differences were established between cave and surface individuals, embryonically or post-embryonically. Comparative embryological studies of *A. infernus* and surface samples showed that pigmentation never developed similarly to what had been seen in comparative embryology of the Subterranean Rak population from Slovenia (Mojaddidi et al. 2018). Therefore, the pigmentation difference was established during embryonic development in both examined populations of the *A. aquaticus* species complex. This is similar to *Astyanax mexicanus*, where cave embryos, from an albino cave population, never developed pigmentation when examined as embryos (McCauley et al. 2004). In *A. mexicanus*, though we do not know if complete loss of pigmentation is intrinsically advantageous in the cave environment, it has been seen that the gene responsible,

oca2, shows pleiotropic effects responsible for sleep loss, which is thought to be an adaptive trait (O'Gorman et al. 2021). Loss of pigmentation in *A. aquaticus* could similarly have some sort of advantageous pleiotropic effect. Support for the idea that loss of pigmentation in *A. aquaticus* is adaptive is that the same region/possibly gene is likely responsible for loss of pigmentation in multiple cave populations of the *A. aquaticus* species complex, including *A. infernus* (Re et al. 2018; Rodas et al. 2023; Fišer et al. 2024).

In addition, we investigated whether size characteristics were established between cave and surface individuals embryonically or post-embryonically. The specific traits of body length, head width, and antenna II length were all significantly different between lab-bred *A. infernus* and surface samples. Previous work had shown that the Subterranean Rak population also had increased body length and increased antenna length (Mojaddidi et al. 2018). As these cave populations are ecologically and geographically distinct, as reviewed by Protas et al. (2023), it is striking that similar size differences are present in adults but also in juveniles, suggesting that the size differences might be adaptive in the cave environment. In fact, one of the commonly described features of cave animals is smaller numbers of larger embryos (Culver 1982), and *A. infernus* has previously been described as having fewer embryos per brood than the adjacent surface population (Turk-Prevorčnik et al. 1998). It could be that larger adults have larger embryos, but previous studies in the surface form of *A. aquaticus* did not see a correlation between adult size and embryo size (Ridley and Thompson 1979). In addition, sulfidic environments have also been associated with larger embryo size in different species. For example, in *P. mexicana*, both cave environments and sulfidic environments housed individuals with larger embryos (Riesch et al. 2010). One reason for this could be resource scarcity (Riesch et al. 2014). Or, perhaps minimizing the surface area to volume ratio with larger individuals minimizes the region

TABLE 3 | Light-interacting genes that exhibit different mRNA levels in cave versus surface samples in mid- and late-stage embryos.

Mid-stage embryos	Cave or surface biased	Light-interacting gene category
NP_001155991.1rhodopsin 1/6-like+	Surface	Phototransduction
XP_008192672.1 serine/threonine-protein phosphatase rdgC	Surface	Phototransduction
XP_015835336.1 high affinity cGMP-specific 3′,5′-cyclic phosphodiesterase 9A	Surface	Phototransduction
XP_015838325.1 G-protein coupled receptor 161	Surface	Phototransduction
XP_015834977.1 protein kinase C-binding protein NELL1 isoform X1	Surface	Phototransduction
XP_008197970.1 regulator of G-protein signaling 7 isoform X2	Surface	Phototransduction
NP_001139379.1dopamine N acetyltransferase isoform 2+	Surface	Melanin synthesis
XP_015839422.1 laccase 1 isoform X1	Surface	Melanin synthesis
XP_972873.2 dopamine N-acetyltransferase°	Cave, surface	Melanin synthesis
XP_008199627.1 carotenoid isomerooxygenase	Cave	Carotenoid pathway
XP_008201448.1 putative ferric-chelate reductase 1	Cave	Heme pathway
XP_967100.1 retinol dehydrogenase 11	Cave	Retinoid pathway
Late-stage embryos		
XP_015834291.1 c-opsin isoform X1	Surface	Phototransduction
NP_001155991.1rhodopsin 1/6-like+	Surface	Phototransduction
XP_015835000.1 transient receptor potential protein isoform X1	Surface	Phototransduction
XP_015837026.1 transient receptor potential channel pyrexia+	Surface	Phototransduction
XP_015835336.1 high affinity cGMP-specific 3′,5′-cyclic phosphodiesterase 9A	Surface	Phototransduction
XP_015834977.1 protein kinase C-binding protein NELL1 isoform X1	Surface	Phototransduction
XP_008197970.1 regulator of G-protein signaling 7 isoform X2	Surface	Phototransduction
XP_015835617.1 period isoform X1	Surface	Diurnal clock
XP_015835556.1 PDF receptor isoform X4	Surface	Diurnal clock
XP_972873.2 dopamine N-acetyltransferase	Cave	Melanin synthesis
XP_967100.1 retinol dehydrogenase 11	Cave	Retinoid pathway

Note: +gene-name also showed differential expression between the Subterranean Rak population versus Rakov Škocjan (Lomheim et al. 2023). °Two paralogues of this gene were present; one was cave-biased and the other was surface-biased.
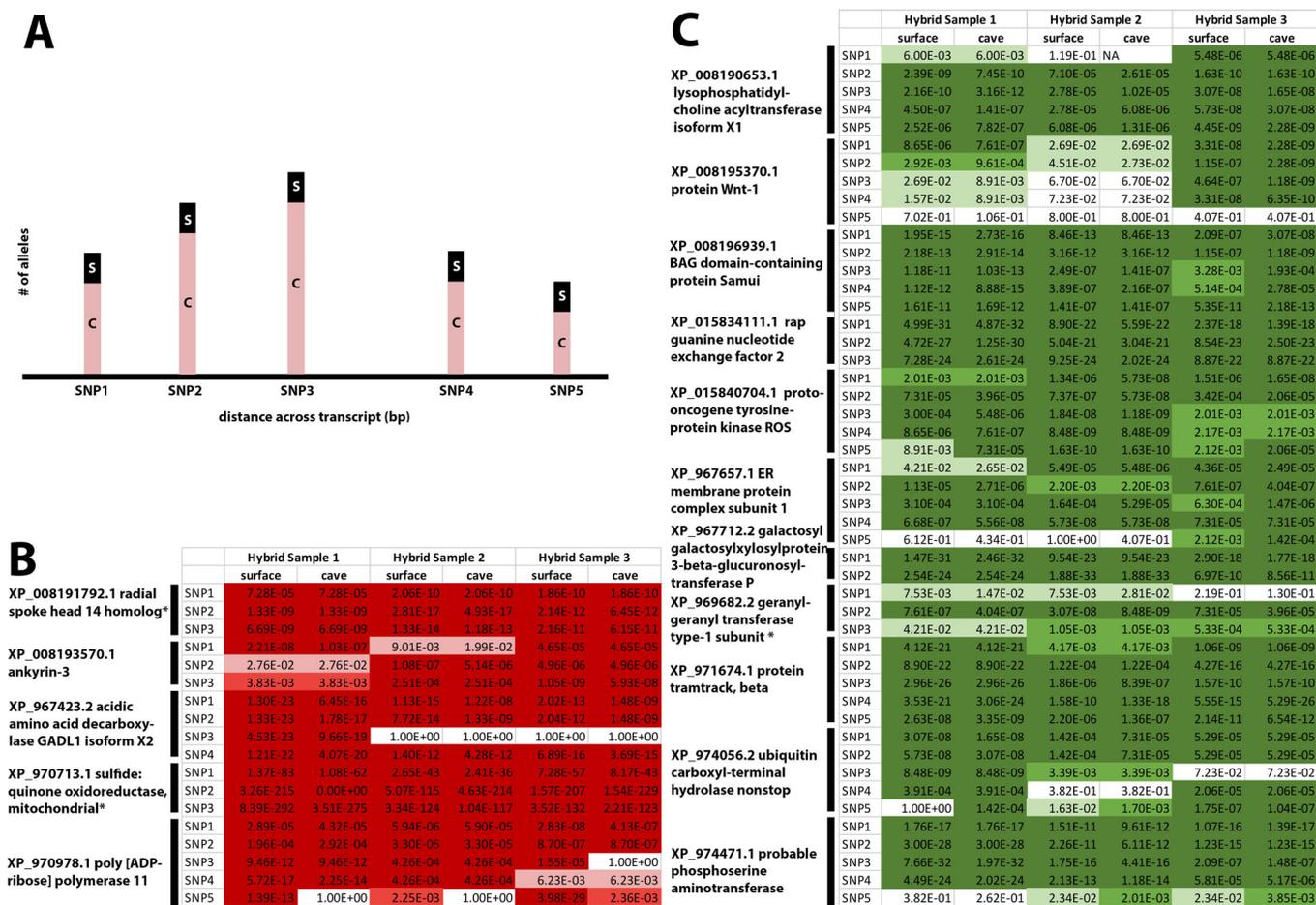
A

# of alleles (y-axis) vs. *distance across transcript (bp)* (x-axis), SNP1–SNP5. S = surface allele, C = cave allele.

B

		Hybrid Sample 1		Hybrid Sample 2		Hybrid Sample 3	
		surface	cave	surface	cave	surface	cave
XP_008191792.1 radial spoke head 14 homolog*	SNP1	7.28E-05	7.28E-05	2.06E-10	2.06E-10	1.86E-10	1.86E-10
	SNP2	1.33E-09	1.33E-09	2.81E-17	4.93E-17	2.14E-12	6.45E-12
	SNP3	6.69E-09	6.69E-09	1.33E-14	1.18E-13	2.16E-11	6.15E-11
XP_008193570.1 ankyrin-3	SNP1	2.21E-08	1.03E-07	9.01E-03	1.99E-02	4.65E-05	4.65E-05
	SNP2	2.76E-02	2.76E-02	1.08E-07	5.14E-06	4.96E-06	4.96E-06
	SNP3	3.83E-03	3.83E-03	2.51E-04	2.51E-04	1.05E-09	5.93E-08
XP_967423.2 acidic amino acid decarboxylase GADL1 isoform X2	SNP1	1.30E-23	6.45E-16	1.13E-15	1.22E-08	2.02E-13	1.48E-09
	SNP2	1.33E-23	1.78E-17	7.72E-14	1.33E-09	2.04E-12	1.48E-09
	SNP3	4.53E-23	9.66E-19	1.00E+00	1.00E+00	1.00E+00	1.00E+00
	SNP4	1.21E-22	4.07E-20	1.40E-12	4.28E-12	6.89E-16	3.69E-15
XP_970713.1 sulfide: quinone oxidoreductase, mitochondrial*	SNP1	1.37E-83	1.08E-62	2.65E-43	2.41E-36	7.28E-57	8.17E-43
	SNP2	3.26E-215	0.00E+00	5.07E-115	4.63E-214	1.57E-207	1.54E-229
	SNP3	8.39E-292	3.51E-275	3.34E-124	1.04E-117	3.52E-132	2.21E-123
XP_970978.1 poly [ADP-ribose] polymerase 11	SNP1	2.89E-05	4.32E-05	5.94E-06	5.90E-05	2.83E-08	4.13E-07
	SNP2	1.96E-04	2.92E-04	3.30E-05	3.30E-05	8.70E-07	8.70E-07
	SNP3	9.46E-12	9.46E-12	4.26E-04	4.26E-04	1.55E-05	1.00E+00
	SNP4	5.72E-17	2.25E-14	4.26E-04	4.26E-04	6.23E-03	6.23E-03
	SNP5	1.39E-13	1.00E+00	2.25E-03	1.00E+00	3.98E-29	2.36E-03

C

		Hybrid Sample 1		Hybrid Sample 2		Hybrid Sample 3	
		surface	cave	surface	cave	surface	cave
XP_008190653.1 lysophosphatidylcholine acyltransferase isoform X1	SNP1	6.00E-03	6.00E-03	1.19E-01	NA	5.48E-06	5.48E-06
	SNP2	2.39E-09	7.45E-10	7.10E-05	2.61E-05	1.63E-10	1.63E-10
	SNP3	2.16E-10	3.16E-12	2.78E-05	1.02E-05	3.07E-08	1.65E-08
	SNP4	4.50E-07	1.41E-07	2.78E-05	6.08E-06	5.73E-08	3.07E-08
	SNP5	2.52E-06	7.82E-07	6.08E-06	1.31E-06	4.45E-09	2.28E-09
XP_008195370.1 protein Wnt-1	SNP1	8.65E-06	7.61E-07	2.69E-02	2.69E-02	3.31E-08	2.28E-09
	SNP2	2.92E-03	9.61E-04	4.51E-02	2.73E-02	1.15E-07	2.28E-09
	SNP3	2.69E-02	8.91E-03	6.70E-02	6.70E-02	4.64E-07	1.18E-09
	SNP4	1.57E-02	8.91E-03	7.23E-02	7.23E-02	3.31E-08	6.35E-10
	SNP5	7.02E-01	1.06E-01	8.00E-01	8.00E-01	4.07E-01	4.07E-01
XP_008196939.1 BAG domain-containing protein Samui	SNP1	1.95E-15	2.73E-16	8.46E-13	8.46E-13	2.09E-07	3.07E-08
	SNP2	2.18E-13	2.91E-14	3.16E-12	1.15E-07	1.15E-07	1.18E-09
	SNP3	1.18E-11	1.03E-13	2.49E-07	1.41E-07	3.28E-03	1.93E-04
	SNP4	1.12E-12	8.88E-15	3.89E-07	2.16E-07	5.14E-04	2.78E-05
	SNP5	1.61E-11	1.69E-12	1.41E-07	1.41E-07	5.35E-11	2.18E-13
XP_015834111.1 rap guanine nucleotide exchange factor 2	SNP1	4.99E-31	4.87E-32	8.90E-22	5.59E-22	2.37E-18	1.39E-18
	SNP2	4.72E-27	1.25E-30	5.04E-21	3.04E-21	8.54E-23	2.50E-23
	SNP3	7.28E-24	2.61E-24	9.25E-24	2.02E-24	8.87E-22	8.87E-22
XP_015840704.1 proto-oncogene tyrosine-protein kinase ROS	SNP1	2.01E-03	2.01E-03	1.34E-06	5.73E-08	1.51E-06	1.65E-08
	SNP2	7.31E-05	3.96E-05	7.37E-07	5.73E-08	3.42E-04	2.06E-05
	SNP3	3.00E-04	5.48E-06	1.84E-08	1.18E-09	2.01E-03	2.01E-03
	SNP4	8.65E-06	7.61E-07	8.48E-09	8.48E-09	2.17E-03	2.17E-03
	SNP5	8.91E-03	7.31E-05	1.63E-10	1.63E-10	2.12E-03	2.06E-05
XP_967657.1 ER membrane protein complex subunit 1	SNP1	4.21E-02	2.65E-02	5.49E-05	5.48E-06	4.36E-05	2.49E-05
	SNP2	1.13E-05	2.71E-06	2.20E-03	2.20E-03	7.61E-07	4.04E-07
	SNP3	3.10E-04	3.10E-04	1.64E-04	5.29E-05	6.30E-04	1.47E-06
	SNP4	6.68E-07	5.56E-08	5.73E-08	5.73E-08	7.31E-05	7.31E-05
	SNP5	6.12E-01	4.34E-01	1.00E+00	4.07E-01	2.12E-03	1.42E-04
XP_967712.2 galactosyl galactosylxylosylprotein 3-beta-glucuronosyltransferase P	SNP1	1.47E-31	2.46E-32	9.54E-23	9.54E-23	2.90E-18	1.77E-18
	SNP2	2.54E-24	2.54E-24	1.88E-33	1.88E-33	6.97E-10	8.56E-11
XP_969682.2 geranylgeranyl transferase type-1 subunit *	SNP1	7.53E-03	1.47E-02	7.53E-03	2.81E-02	2.19E-01	1.30E-01
	SNP2	7.61E-07	4.04E-07	3.07E-08	8.48E-09	7.31E-05	3.96E-05
	SNP3	4.21E-02	4.21E-02	1.05E-03	1.05E-03	5.33E-04	5.33E-04
XP_971674.1 protein tramtrack, beta	SNP1	4.12E-21	4.12E-21	4.17E-03	4.17E-03	1.06E-09	1.06E-09
	SNP2	8.90E-22	8.90E-22	1.22E-04	1.22E-05	4.27E-16	4.27E-16
	SNP3	2.96E-26	2.96E-26	1.86E-06	8.39E-07	1.57E-10	1.57E-10
	SNP4	3.53E-21	3.06E-24	1.58E-10	1.33E-18	5.55E-15	5.29E-26
	SNP5	2.63E-08	3.35E-09	2.20E-06	1.36E-07	2.14E-11	6.54E-12
XP_974056.2 ubiquitin carboxyl-terminal hydrolase nonstop	SNP1	3.07E-08	1.65E-08	1.42E-04	7.31E-05	5.29E-05	5.29E-05
	SNP2	5.73E-08	3.07E-08	1.42E-04	7.31E-05	5.29E-05	5.29E-05
	SNP3	8.48E-09	8.48E-09	3.39E-03	3.39E-03	7.23E-02	7.23E-02
	SNP4	3.91E-04	3.91E-04	3.82E-01	3.82E-01	2.06E-05	2.06E-05
	SNP5	1.00E+00	1.42E-04	1.63E-02	1.70E-03	1.75E-07	1.04E-07
XP_974471.1 probable phosphoserine aminotransferase	SNP1	1.76E-17	1.76E-17	1.51E-11	9.61E-12	1.07E-16	1.39E-17
	SNP2	3.00E-28	3.00E-28	2.26E-11	6.11E-12	1.23E-15	1.23E-15
	SNP3	7.66E-32	1.97E-32	1.75E-16	4.41E-16	2.09E-07	1.48E-07
	SNP4	4.49E-24	2.02E-24	2.13E-13	1.18E-14	5.81E-05	5.17E-06
	SNP5	3.82E-01	2.62E-01	2.34E-02	2.01E-03	2.34E-02	3.85E-03

FIGURE 5 | Genes showing allele-specific expression. Three F1 hybrid samples were examined both in reference to the surface transcriptome and the cave transcriptome. A. Allele counting was performed using FreeBayes variant counting for five SNPs along the location of each gene, both the cave and surface versions of the transcript (shown as rows labeled as SNP1-5 in parts B and C). C represents the cave allele, and S represents the surface allele. Here, a schematic of a hypothetical gene with cave-biased expression is shown as an example. B. Genes that show cave-biased allele-specific expression are shown in red. C. Genes that show surface-biased allele-specific expression are in green. We performed binomial tests to show whether the ratio of surface to cave alleles deviates significantly from the expected 1:1 ratio. Multiple comparison correction was performed using the Benjamini–Yekutieli procedure. White cells denote non-significance with $p > 0.05$, light green/red cells denote significance with $0.005 < p < 0.05$, plain green/red cells denote significance with $0.0005 < p < 0.005$, and dark green/red cells denote significance with $p < 0.0005$. Genes shown in B and C are those for which at least two SNPs showed a $p < 0.05$ for all hybrid samples to both the surface and cave transcript. The code before the gene name is the *Tribolium castaneum* gene ID. An asterisk indicates a gene that did not have a high-quality, non-chimeric primary alignment for either the surface or cave transcript. [Color figure can be viewed at wileyonlinelibrary.com]

exposed to the toxin (pertinent to the sulfidic environment). Another possibility is that the size difference is due to genome size (nucleotypic effect) (Gregory 2001). Subterranean *Proasellus* isopods have been shown to have a larger genome size than non-subterranean *Proasellus* isopods (Lefébure et al. 2017). Interestingly, though we have seen this size difference in juveniles comparing two cave/surface pairs, a recent study comparing adult individuals from six cave populations and nine surface populations for adult individuals of *A. aquaticus* did not show a general trend of larger cave adults as compared to surface adults (Herczeg et al. 2023). However, *A. infernus* did have the largest adult size of all the populations examined and showed a larger size than the Turkish bath surface population (two of the populations examined here; Herczeg et al. 2023). Future work could examine juveniles of other cave populations to see if there is a trend in cave versus surface juvenile size, though absent as an overall trend in adults.

With the cave and surface samples, we were able to compare wild and lab-reared juveniles of both populations. Interestingly, many of the measures differed between wild and lab-reared individuals within a population. For example, body length was statistically different between wild and lab surface populations as well as between wild and lab cave populations. Body length was higher in wild cave samples as compared to lab cave samples. Conversely, body length was lower in wild surface samples as compared to lab surface populations. In terms of allometry, lab cave and wild surface samples had longer antennae and shorter body lengths, controlling for overall size, while wild cave and lab surface samples generally had shorter antennae and longer body lengths, controlling for overall size. Therefore, there was no overall cave versus surface trend in allometry. A possible reason for the opposite trends seen for wild-caught cave versus wild-caught surface samples is that, as mentioned above, the wild-caught cave samples were mostly taken from females that were already gravid when individuals

were collected but the majority of the wild-caught surface samples were taken from females that became gravid on transport to the lab which was likely a stressful and different environment. As for the lab-bred versus wild-caught environmental conditions, the supply of food is likely to be greater and of higher nutritional value in the lab for both cave and surface populations. In the wild, surface populations are susceptible to seasonal fluctuations in food availability. Additionally, it is unclear whether the type of food used in the laboratory populations, algae pellets, recapitulates the type of nutrition the animals would be getting in either the cave or surface environment. Furthermore, other conditions, such as water composition and light regime, could also influence juvenile size by influencing gene expression changes. Finally, our lab-bred individuals were from a single-generation bred in the lab; it is possible that an increased number of generations in the lab might also affect some size measures if some sort of epigenetic modifications were affecting gene expression changes. Rearing *A. mexicanus* surface fish in the dark also resulted in morphological changes (Bilandžija et al. 2020). Additionally, studies of *A. mexicanus* showed many gene expression differences in population comparisons of both cave and surface populations that were wild-caught and lab-reared (Krishnan et al. 2020). Furthermore, altering conditions that cave and surface *A. mexicanus* were housed in (light–dark cycle vs. dark rearing) resulted in many gene expression changes (Sears et al. 2020). Therefore, the difference in size characters between wild and lab-raised *A. aquaticus* juveniles is not surprising.

F1 hybrids were investigated to see if they were more similar in size to surface or *A. infernus* samples or were intermediate in size between the two. In an ideal world, we would have generated F1s from cave and surface individuals that were raised from embryos in the lab, and then, by our previous definition, these F1s would count as lab-bred, and we could compare them exclusively to lab-bred surface and cave samples. However, F1s of these populations are not straightforward to achieve, and the F1s we generated were from cave males that were recently brought from the wild and surface females, of which some were recently brought from the wild or others that were lab-bred. These crosses were performed in the lab, mostly after they had been living in the lab for at least a month. As a result, these F1s cannot be classified by our previous definitions as wild-caught or lab-bred but instead are somewhere in between. Therefore, we compared the F1 measurements to both wild-caught and lab-bred samples to determine if the F1 measurement was significantly different for each measure. F1 hybrid juveniles were significantly different than and intermediate between cave and surface samples for antenna length. Furthermore, F1 hybrid juveniles were significantly different than cave samples for body length and head width, though not significantly different than surface samples. In *A. mexicanus*, F1 hybrids were different than both cave and surface parents for melanophore number, width of lower jaw, sclera size, and eye diameter (Wilkens 1988; Ma et al. 2018). Furthermore, some of these measures in *A. mexicanus* showed maternal effect and were different for F1 hybrids generated from cave females × surface males as compared to F1 hybrids generated from cave males × surface females (Ma et al. 2018). We, unfortunately, were not yet able to investigate maternal effect in *A. aquaticus* as the females for all F1 hybrids were surface, and the males were cave due to the

difficulty in setting up the reciprocal crosses. This is likely not an incompatibility issue; cave females do not breed robustly in the lab, so it is more difficult to get them to breed (even if with males from the same population).

4.2 | Transcriptomic Analysis

We had predicted that we would see lower mRNA levels of genes involved in phototransduction in the mid- and late-stage cave embryos (surface-biased expression). Supporting our prediction, six genes within the phototransduction group were surface-biased in the mid-stage timepoint, and seven genes were surface-biased at the late-stage timepoint (some of these genes are overlapping). This resembles what was seen in the Subterranean Rak population where three genes involved in phototransduction showed significantly different mRNA levels in cave and surface samples at both timepoints (Lomheim et al. 2023); two genes in particular were in common between the analyses here and those involving the Subterranean Rak population—rhodopsin 1/6-like and transient receptor potential channel pyrexia. Additionally, differential expression in cave versus surface spiders was also seen for opsins (Gainett et al. 2020). What is somewhat surprising is that no genes involved in ommochrome synthesis were differentially expressed in the cave versus surface comparison, as ommochromes are thought to be the pigments present in *A. aquaticus* (Needham and Brunet 1957). In comparisons of the Subterranean Rak population and Rakov Škocjan surface (Lomheim et al. 2023), a single gene in the ommochrome pathway showed significantly different mRNA levels, *scarlet* (Lomheim et al. 2023). It is possible that genes involved in the ommochrome synthesis pathway are lowly expressed at the timepoints examined, and that is why we did not detect significant differential expression in this set of genes.

We found that genes involved in melanin synthesis showed different mRNA levels between cave and surface samples: one paralogue of *dopamine N-acetyltransferase* was cave-biased, another paralogue of *dopamine N-acetyltransferase* was surface-biased, and *laccase 1* was surface-biased. Interestingly, in comparisons of Subterranean Rak versus Rakov Škocjan surface, two paralogues of *dopamine N-acetyltransferase* were surface-biased (Lomheim et al. 2023). However, in the current study, the cave and surface sequences of *dopamine N-acetyltransferase* do appear to be the orthologue of the *dopamine N-acetyltransferase* sequences from Subterranean Rak versus Rakov Škocjan surface that showed differential expression in the opposite direction indicating that for this gene, there is likely a difference in bias in the two cave populations. Differences in melanin synthesis are not surprising as they have been seen in other cave animals and might also relate to other functions of melanin pathway genes, such as melatonin synthesis (Bilandžija et al. 2012; Bilandžija et al. 2013; Bilandžija et al. 2018). One of the differentially expressed genes discussed above, dopamine N-acetyltransferase, blasts to *aanat*, which has been previously described as targeted for inactivating mutations in cave and deep-sea organisms (Huang et al. 2022). Furthermore, *aanat2* has been disrupted in *A. mexicanus*, by CRISPR, which resulted in decreased nighttime sleep (Mack et al. 2021). In

addition, two genes in the diurnal clock category showed decreased mRNA levels in cave samples as compared to surface samples, *period* and *PDF receptor*. Whether these expression differences indicate a behavioral difference between cave and surface populations, such as sleep, remains to be seen. Behavioral studies of another cave population of *A. aquaticus*, from Molnár János Cave, have shown a greater likelihood of feeding and higher movement when compared to a nearby surface population (Herczeg et al. 2020; Berisha et al. 2022). Additionally, a previous study involving *A. infernus* reported that cave individuals fed less but moved more as compared to surface individuals (Mösslacher and Creuzé des Châtelliers 1996). Further behavioral studies between *A. infernus* and surface individuals will help to understand some of these gene expression differences.

Differential expression analysis allowed us to investigate genes and pathways that might be involved in cave-specific phenotypes, such as eye and pigmentation loss, but also genes and pathways that were involved in sulfidic versus non-sulfidic environments. Differential expression experiments had previously been performed on sulfidic and non-sulfidic populations of the fish, *P. mexicana*. Lists of genes were identified that were overexpressed and underexpressed in multiple sulfidic populations versus multiple non-sulfidic populations (Tobler et al. 2016). Several genes were in common in the list of genes overexpressed in sulfidic populations of *P. mexicana* that also showed higher mRNA levels in *A. infernus* cave samples as compared to surface samples including *sulfide:quinone oxidoreductase (sqrdl)*, *persulfide dioxygenase ETHE1 (ethe1)*, *acidic amino acid decarboxylase GADL1 (gadl1)*, and *carbonic anhydrase 3*. In addition, one gene was biased toward surface samples in *A. aquaticus*, which was more lowly expressed in sulfidic populations in *P. mexicana*, *adenosylhomocysteinase 2*. Additional studies in *P. mexicana* showed evidence for genetic changes and selection in two of these genes, *sqrdl* and *ethe1*, in multiple sulfidic populations of *P. mexicana* (Brown et al. 2018; Brown et al. 2019; Greenway et al. 2020; Ryan et al. 2023). Both genes encode key enzymes in mitochondrial sulfide oxidation. Interestingly, one of the above genes also showed cave-biased allele-specific expression in our study, *sqrdl*. Additionally, we also saw cave-biased allele-specific expression in *gadl1*; when blasted, this gene showed the top hit in crustaceans as cysteine sulfinic acid decarboxylase (*csad*). *Csad* is involved in the biosynthesis of taurine, a sulfur-containing amino acid which was shown to protect against oxidative stress (Baliou et al. 2021). Interestingly, *csad* facilitates the production of hypotaurine, a precursor to taurine, and hypotaurine and thiotaurine are found at high concentrations in hydrothermal vent-dwelling organisms and are thought to be protective in environments of high H_2S (Yancey et al. 2009). The parallels between the genes identified as likely to play a role in adaptation to the sulfidic environment between other organisms discussed above and *A. infernus* support the idea that multiple sulfidic organisms (from invertebrates to vertebrates) may have evolved similar solutions to adapt to sulfidic environments. However, it is important to keep in mind that in our current study in *A. aquaticus*, the comparison is between sulfidic cave and non-sulfidic surface with two different parameters that are not separable: sulfidic versus non-sulfidic and cave versus surface. However, because the *A. aquaticus* species complex contains both sulfidic cave and non-sulfidic cave populations, comparison of sulfidic versus non-sulfidic cave populations should be possible in the future. We do not currently know of any sulfidic surface *A. aquaticus* populations, but if discovered, the additional comparison of sulfidic surface to non-sulfidic surface could be made.

As previously mentioned, *gadl1*, blasting to *csad*, showed cave-biased allele-specific expression in F1 hybrids. This could also be a possible candidate for eye phenotype as *csad* is in the cysteine metabolism pathway, similar to a gene contributing to eye degeneration in *A. mexicanus* is *cystathionine ß-synthase a* (*cbsa*) (Ma et al. 2020). *Csad* is also a possible pigment candidate gene as CRISPR gene editing in the planthopper, *Nilaparvata lugens*, showed increased pigmentation (Chen et al. 2021).

Multiple other genes showed surface-biased allele-specific expression, including *ER membrane protein complex subunit 1*, *rap guanine nucleotide exchange factor 2*, and *tramtrack*. Humans with defects in *ER membrane protein complex subunit 1* have visual impairment (Geetha et al. 2018), and therefore, this could be a candidate gene for the ommatidia loss phenotype. *Rap guanine nucleotide exchange factor 2* is also known as *pdz-gef*, which is required for photoreceptor differentiation in *D. melanogaster* (Baril et al. 2014). Interestingly, allele-specific expression analyses in *A. mexicanus* showed allele-specific expression in *rap guanine exchange factor 2* as well (Leclercq et al. 2024), though cave-biased rather than surface-biased. *Tramtrack* blasted to *bric-a-brac*, which has been associated with pigmentation variation in *Drosophilidae* (reviewed by Massey and Wittkopp (2016)). Furthermore, *bric-a-brac* is known to play a role in *D. melanogaster* limb segmentation (Chu et al. 2002) and therefore could also be involved in appendage differences between the cave and surface forms.

There was no overlap in the lists of genes that showed both allele-specific expression and differential expression with log2 fold of 2 and $p < 0.05$ between the Romanian samples described here (*A. infernus* and surface) and the previously described populations of Subterranean Rak and Rakov Škocjan (Lomheim et al. 2023). However, if the log2 fold were lowered to 1.8 for cave versus surface analysis, there would be one gene in common between both cave/surface pair comparisons that showed cave-biased allele-specific expression and differential expression: EFR3 homolog cmp44E (*efr3*). The *Drosophila melanogaster* orthologue of *efr3* has been documented to play a role in phototransduction, hypoxia, and olfaction (Kain et al. 2009; Balakrishnan et al. 2018; Lu et al. 2022). It is striking that this gene shows allele-specific expression in both populations, which are geographically and ecologically disparate. Examination of other populations in the *A. aquaticus* species complex would help to inform what overexpression of this gene in the cave ecomorph might accomplish and whether this expression difference could result in adaptations in multiple cave environments. The lack of overlap in genes that show allele-specific expression, aside from *efr3*, between the two cave/surface pair comparisons could be due to the restrictive nature of how we defined allele-specific expression, or it could be because these cave populations have come to different solutions in the evolution to their respective subterranean environments. We do know, however, that *A. infernus* likely has the same genetic

mechanism for an orange phenotype masked by depigmentation (Rodas et al. 2023). Therefore, it is likely that there are both genetic commonalities and differences in the genetic basis of subterranean phenotypes in *A. infernus* and the Subterranean Rak population.

One limitation of our allele-specific expression studies, as previously mentioned, was that we were only able to obtain F1 hybrids with cave males and surface females and not the reciprocal cross of cave females and surface males. To address mis-identifying genes with parent-of-origin effect as genes with allele-specific expression, all our genes with allele-specific expression also showed differential expression, which one would not expect to see with genes of parent-of-origin effect. Future work will prioritize investigating other laboratory living conditions for *A. infernus* to improve the survival/reproduction of individuals and hopefully generate F1 crosses with cave females × surface males.

Another potential limitation of our studies is that we pooled together individuals from two different wells for the differential expression, allele-specific expression, and comparative embryology studies. It is possible that these wells have connections between them. However, it is also possible that some degree of cryptic diversity might be present. Future studies can examine the genetic structure between the different wells and other populations of *A. infernus*.

In sum, we have established *A. infernus* as a model for both understanding cave-specific phenotypes and adaptation to sulfidic environments. We now have the tools for future comparative work, both between sulfidic and non-sulfidic cave populations of the *A. aquaticus* species complex, but also comparative work between sulfidic cave species—*A. infernus* and *P. mexicana*. Our transcriptomic work has identified candidate genes of interest that could be involved in adaptation to the cave environment and/or adaptation to sulfidic environments. Of particular interest are genes that show molecular convergence between *A. infernus* and *P. mexicana*. The next step will be to test these genes functionally to investigate what advantages they could be providing.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in PRJNA957305 at https://www.ncbi.nlm.nih.gov, reference number PRJNA957305. All sequences discussed in this report are present in the National Center for Biotechnology Information, Sequencing Reads Archive (BioProject ID: PRJNA957305).

References

Bakovic, V., M. L. Martin Cerezo, A. Höglund, et al. 2021. "The Genomics of Phenotypically Differentiated *Asellus aquaticus* Cave, Surface Stream and Lake Ecotypes." *Molecular Ecology* 30: 3530–3547.

Balakrishnan, S. S., U. Basu, D. Shinde, R. Thakur, M. Jaiswal, and P. Raghu. 2018. "Regulation of PI4P Levels by PI4KIIIα During G-Protein-Coupled PLC Signaling in Drosophila Photoreceptors." *Journal of Cell Science* 131: jcs217257.

Baliou, S., M. Adamaki, P. Ioannou, et al. 2021. "Protective Role of Taurine Against Oxidative Stress." *Molecular Medicine Reports* 24: 1–19.

Baril, C., M. Lefrançois, M. Sahmi, H. Knævelsrud, and M. Therrien. 2014. "Apical Accumulation of the Sevenless Receptor Tyrosine Kinase During Drosophila Eye Development Is Promoted by the Small GTPase Rap1." *Genetics* 197: 1237–1250.

Benjamini, Y., and D. Yekutieli. 2001. "The Control of the False Discovery Rate in Multiple Testing Under Dependency." *Annals of Statistics* 4: 1165–1188.

Berisha, H., G. Horváth, Ž. Fišer, G. Balázs, C. Fišer, and G. Herczeg. 2022. "Sex-Dependent Increase of Movement Activity in the Freshwater Isopod *Asellus aquaticus* Following Adaptation to a Predator-Free Cave Habitat." *Current Zoology* 69: 418–425.

Bilandžija, H., L. Abraham, L. Ma, K. J. Renner, and W. R. Jeffery. 2018. "Behavioural Changes Controlled by Catecholaminergic Systems Explain Recurrent Loss of Pigmentation in Cavefish." *Proceedings of the Royal Society B: Biological Sciences* 285: 20180243.

Bilandžija, H., B. Hollifield, M. Steck, et al. 2020. "Phenotypic Plasticity as a Mechanism of Cave Colonization and Adaptation." *eLife* 9: e51830.

Bilandžija, H., L. Ma, A. Parkhurst, and W. R. Jeffery. 2013. "A Potential Benefit of Albinism in Astyanax Cavefish: Downregulation of the Oca2 Gene Increases Tyrosine and Catecholamine Levels as an Alternative to Melanin Synthesis." *PLoS One* 8: e80823.

Bilandžija, H., H. Ćetković, and W. R. Jeffery. 2012. "Evolution of Albinism in Cave Planthoppers by a Convergent Defect in the First Step of Melanin Biosynthesis." *Evolution & Development* 14: 196–203.

Brad, T., S. Iepure, and S. Sarbu. 2021. "The Chemoautotrophically Based Movile Cave Groundwater Ecosystem, a Hotspot of Subterranean Biodiversity." *Diversity* 13: 128.

Bray, N. L., H. Pimentel, P. Melsted, and L. Pachter. 2016. "Near-Optimal Probabilistic RNA-seq Quantification." *Nature Biotechnology* 34: 525–527.

Brown, A. P., L. Arias-Rodriguez, M. Yee, M. Tobler, and J. L. Kelley. 2018. "Concordant Changes in Gene Expression and Nucleotides Underlie Independent Adaptation to Hydrogen-Sulfide-Rich Environments." *Genome Biology and Evolution* 10: 2867–2881.

Brown, A. P., K. L. McGowan, E. J. Schwarzkopf, et al. 2019. "Local Ancestry Analysis Reveals Genomic Convergence in Extremophile Fishes." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 374: 20180240.

Camacho, C., G. Coulouris, V. Avagyan, et al. 2009. "BLAST: Architecture and Applications." *BMC Bioinformatics* 10: 421.

Chen, J. X., W. X. Li, J. Lyu, Y. T. Hu, G. Huang, and W. Q. Zhang. 2021. "CRISPR/Cas9-Mediated Knockout of the NlCSAD Gene Results in Darker Cuticle Pigmentation and a Reduction in Female Fecundity in *Nilaparvata lugens* (Hemiptera: Delphacidae)." *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology* 256: 110921.

Chu, J., P. D. S. Dong, and G. Panganiban. 2002. "Limb Type-Specific Regulation of Bric a Brac Contributes to Morphological Diversity." *Development* 129: 695–704.

Cock, P. J. A., J. M. Chilton, B. Grüning, J. E. Johnson, and N. Soranzo. 2015. "NCBI BLAST Integrated Into Galaxy." *GigaScience* 4: s13742–s13747.

Conesa, A., S. Götz, J. M. García-Gómez, J. Terol, M. Talón, and M. Robles. 2005. "Blast2GO: A Universal Tool for Annotation, Visualization and Analysis in Functional Genomics Research." *Bioinformatics* 21: 3674–3676.

Culver, D. C. 1982. *Cave Life: Evolution and Ecology*. Harvard University Press.

Engel, A. S. 2007. "Observations on the Biodiversity of Sulfidic Karst Habitats." *Journal of Cave and Karst Studies* 69: 187–206.

Fišer, C., R. Luštrik, S. Sarbu, J. F. Flot, and P. Trontelj. 2015. "Morphological Evolution of Coexisting Amphipod Species Pairs From Sulfidic Caves Suggests Competitive Interactions and Character Displacement, but No Environmental Filtering and Convergence." *PLoS One* 10: e0123535.

Fišer, Ž., H. Whitehorn, T. Furness, P. Trontelj, and M. Protas. 2024. "Genetic Bias in Repeated Evolution of Pigment Loss in Cave Populations of the *Asellus aquaticus* Species Complex." *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 342: 425–436.

Flot, J. F., G. Wörheide, and S. Dattagupta. 2010. "Unsuspected Diversity of Niphargus Amphipods in the Chemoautotrophic Cave Ecosystem of Frasassi, Central Italy." *BMC Evolutionary Biology* 10: 171.

Gainett, G., J. A. Ballesteros, C. R. Kanzler, et al. 2020. "Systemic Paralogy and Function of Retinal Determination Network Homologs in Arachnids." *BMC Genomics* 21: 811.

Garrison, E., and G. Marth, 2012. Haplotype-Based Variant Detection From Short-Read Sequencing. arXiv, 1207.

Geetha, T. S., L. Lingappa, A. R. Jain, et al. 2018. "A Novel Splice Variant in EMC1 Is Associated With Cerebellar Atrophy, Visual Impairment, Psychomotor Retardation With Epilepsy." *Molecular Genetics & Genomic Medicine* 6: 282–287.

Gilbert, D. 2013. "EvidentialGene: tr2aacds, mRNA transcript assembly software." http://arthropods.eugenes.org/EvidentialGene.

Grabherr, M. G., B. J. Haas, M. Yassour, et al. 2011. "Full-Length Transcriptome Assembly From RNA-Seq Data Without a Reference Genome." *Nature Biotechnology* 29: 644–652.

Greenway, R., N. Barts, C. Henpita, et al. 2020. "Convergent Evolution of Conserved Mitochondrial Pathways Underlies Repeated Adaptation to Extreme Environments." *Proceedings of the National Academy of Sciences* 117: 16424–16430.

Gregory, T. R. 2001. "Coincidence, Coevolution, or Causation? DNA Content, Cell Size, and the C-Value Enigma." *Biological Reviews of the Cambridge Philosophical Society* 76: 65–101.

Gross, J. B., D. A. Sun, B. M. Carlson, S. Brodo-Abo, and M. E. Protas. 2019. "Developmental Transcriptomic Analysis of the Cave-Dwelling Crustacean, *Asellus aquaticus*." *Genes* 11: 42.

Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler. 2013. "QUAST: Quality Assessment Tool for Genome Assemblies." *Bioinformatics* 29: 1072–1075.

Haas, B. 2023. Trinity Differential Expression. https://github.com/trinityrnaseq/trinityrnaseq/wiki/Trinity-Differential-Expression.

Herczeg, G., G. Balázs, A. Biró, Ž. Fišer, S. Kralj-Fišer, and C. Fišer. 2023. "Island and Rensch's Rules Do Not Apply to Cave Vs. Surface Populations of *Asellus aquaticus*." *Frontiers in Ecology and Evolution* 11: 1155261.

Herczeg, G., V. P. Hafenscher, G. Balázs, Ž. Fišer, S. Kralj-Fišer, and G. Horváth. 2020. "Is Foraging Innovation Lost Following Colonization of a Less Variable Environment? A Case Study in Surface- Vs. Cave-Dwelling *Asellus aquaticus*." *Ecology and Evolution* 10: 5323–5331.

Huang, Y., J. Li, C. Bian, R. Li, X. You, and Q. Shi. 2022. "Evolutionary Genomics Reveals Multiple Functions of Arylalkylamine N-Acetyltransferase in Fish." *Frontiers in Genetics* 13: 820442.

Kain, P., S. Chandrashekaran, V. Rodrigues, and G. Hasan. 2009. "Drosophila Mutants in Phospholipid Signaling Have Reduced Olfactory Responses as Adults and Larvae." *Journal of Neurogenetics* 23: 303–312.

Kelley, J. L., C. N. Passow, M. Plath, L. A. Rodriguez, M. C. Yee, and M. Tobler. 2012. "Genomic Resources for a Model in Adaptation and Speciation Research: Characterization of the *Poecilia mexicana* Transcriptome." *BMC Genomics* 13: 652.

Konec, M., S. Prevorčnik, S. M. Sarbu, R. Verovnik, and P. Trontelj. 2015. "Parallels Between Two Geographically and Ecologically Disparate Cave Invasions by the Same Species, *Asellus aquaticus* (Isopoda, Crustacea)." *Journal of Evolutionary Biology* 28: 864–875.

Krishnan, J., J. L. Persons, R. Peuß, et al. 2020. "Comparative Transcriptome Analysis of Wild and Lab Populations of *Astyanax mexicanus* Uncovers Differential Effects of Environment and Morphotype on Gene Expression." *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 334: 530–539.

Lafuente, E., M. D. Lürig, M. Rövekamp, et al. 2021. "Building on 150 Years of Knowledge: The Freshwater Isopod *Asellus aquaticus* as an Integrative Eco-Evolutionary Model System." *Frontiers in Ecology and Evolution* 9: 748212.

Langmead, B., and S. L. Salzberg. 2012. "Fast Gapped-Read Alignment With Bowtie 2." *Nature Methods* 9: 357–359.

Leclercq, J., J. Torres-Paz, M. Policarpo, F. Agnès, and S. Rétaux. 2024. "Evolution of the Regulation of Developmental Gene Expression in Blind Mexican Cavefish." *Development* 151: dev202610.

Lefébure, T., C. Morvan, F. Malard, et al. 2017. "Less Effective Selection Leads to Larger Genomes." *Genome Research* 27: 1016–1028.

Li, H. 2018. "Minimap2: Pairwise Alignment for Nucleotide Sequences." *Bioinformatics* 34: 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

Lomheim, H. J., L. Reyes Rodas, L. Mulla, et al. 2023. "Transcriptomic Analysis of Cave, Surface, and Hybrid Samples of the Isopod *Asellus aquaticus* and Identification of Chromosomal Location of Candidate Genes for Cave Phenotype Evolution." *EvoDevo* 14: 9.

Love, M. I., W. Huber, and S. Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-seq Data With DESeq2." *Genome Biology* 15: 550.

Lu, J., W. Dong, G. R. Hammond, and Y. Hong. 2022. "Hypoxia Controls Plasma Membrane Targeting of Polarity Proteins by Dynamic Turnover of PI4P and PI (4, 5) P2." *Frontiers in Ecology and Evolution* 11: e79582.

Lukić, M., L. Jovović, J. Bedek, et al. 2024. "A Practical Guide for the Husbandry of Cave and Surface Invertebrates as the First Step in Establishing New Model Organisms." *PLoS One* 19, no. 4: e0300962.

Ma, L., A. V. Gore, D. Castranova, et al. 2020. "A Hypomorphic Cystathionine ß-Synthase Gene Contributes to Cavefish Eye Loss by Disrupting Optic Vasculature." *Nature Communications* 11: 2772.

Ma, L., A. G. Strickler, A. Parkhurst, M. Yoshizawa, J. Shi, and W. R. Jeffery. 2018. "Maternal Genetic Effects in Astyanax Cavefish Development." *Developmental Biology* 441: 209–220.

Mack, K. L., J. B. Jaggard, J. L. Persons, et al. 2021. "Repeated Evolution of Circadian Clock Dysregulation in Cavefish Populations." *PLoS Genetics* 17: e1009642.

Massey, J. H., and P. J. Wittkopp. 2016. "The Genetic Basis of Pigmentation Differences Within and Between Drosophila Species." *Current Topics in Developmental Biology* 119: 27–61.

McCauley, D. W., E. Hixon, and W. R. Jeffery. 2004. "Evolution of Pigment Cell Regression in the Cavefish Astyanax: A Late Step in Melanogenesis." *Evolution & Development* 6: 209–218.

McGowan, K. L., C. N. Passow, L. Arias-Rodriguez, M. Tobler, and J. L. Kelley. 2019. "Expression Analyses of Cave Mollies (*Poecilia mexicana*) Reveal Key Genes Involved in the Early Evolution of Eye Regression." *Biology Letters* 15: 20190554.

Mojaddidi, H., F. E. Fernandez, P. A. Erickson, and M. E. Protas. 2018. "Embryonic Origin and Genetic Basis of Cave Associated Phenotypes in the Isopod Crustacean *Asellus aquaticus*." *Scientific Reports* 8: 16589.

Mösslacher, F., and M. Creuzé des Châtelliers. 1996. "Physiological and Behavioural Adaptations of an Epigean and a Hypogean Dwelling Population of *Asellus aquaticus* (L.) (Crustacea, Isopoda)." *Archiv für Hydrobiologie* 138: 187–198.

Nariai, N., K. Kojima, T. Mimori, Y. Kawai, and M. Nagasaki. 2016. "A Bayesian Approach for Estimating Allele-Specific Expression From RNA-Seq Data With Diploid Genomes." *BMC Genomics* 17: 2.

Needham, A. E., and P. Brunet. 1957. "The Integumental Pigment of Asellus." *Comparative Biochemistry and Physiology* 13: 207–209.

O'Gorman, M., S. Thakur, G. Imrie, et al. 2021. "Pleiotropic Function of the oca2 Gene Underlies the Evolution of Sleep Loss and Albinism in Cavefish." *Current Biology* 31: 3694–3701.e4.

Passow, C. N., A. P. Brown, L. Arias-Rodriguez, et al. 2017. "Complexities of Gene Expression Patterns in Natural Populations of an Extremophile Fish (*Poecilia mexicana*, Poeciliidae)." *Molecular Ecology* 26: 4211–4225.

Passow, C. N., C. Henpita, J. H. Shaw, et al. 2017. "The Roles of Plasticity and Evolutionary Change in Shaping Gene Expression Variation in Natural Populations of Extremophile Fish." *Molecular Ecology* 26: 6384–6399.

Pérez-Moreno, J. L., G. Balázs, and H. D. Bracken-Grissom. 2018. "Transcriptomic Insights Into the Loss of Vision in Molnár János Cave's Crustaceans." *Integrative and Comparative Biology* 58: 452–464.

Protas, M., and W. R. Jeffery. 2012. "Evolution and Development in Cave Animals: From Fish to Crustaceans." *WIREs Developmental Biology* 1: 823–845.

Protas, M., P. Trontelj, S. Prevorčnik, and Ž. Fišer. 2023. "The *Asellus aquaticus* Species Complex: An Invertebrate Model in Subterranean Evolution." In *Groundwater Ecology and Evolution*, 329–350. Elsevier.

Protas, M. E., P. Trontelj, and N. H. Patel. 2011. "Genetic Basis of Eye and Pigment Loss in the Cave Crustacean, *Asellus aquaticus*." *Proceedings of the National Academy of Sciences* 108: 5702–5707.

R Core Team. 2022. "*R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing." https://www.Rproject.org/.

Raudvere, U., L. Kolberg, I. Kuzmin, et al. 2019. "G: Profiler: A Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 Update)." *Nucleic Acids Research* 47: W191–W198.

Re, C., Ž. Fišer, J. Perez, A. Tacdol, P. Trontelj, and M. E. Protas. 2018. "Common Genetic Basis of Eye and Pigment Loss in Two Distinct Cave Populations of the Isopod Crustacean *Asellus aquaticus*." *Integrative and Comparative Biology* 58: 421–430.

Ridley, M., and D. J. Thompson. 1979. "Size and Mating in *Asellus aquaticus* (Crustacea: Isopoda)." *Zeitschrift für Tierpsychologie* 51: 380–397.

Riesch, R., M. Plath, and I. Schlupp. 2010. "Toxic Hydrogen Sulfide and Dark Caves: Life-History Adaptations in a Livebearing Fish (*Poecilia mexicana*, Poeciliidae)." *Ecology* 91: 1494–1505.

Riesch, R., M. Plath, I. Schlupp, M. Tobler, and R. Brian Langerhans. 2014. "Colonisation of Toxic Environments Drives Predictable Life-History Evolution in Livebearing Fishes (Poeciliidae)." *Ecology Letters* 17: 65–71.

Robertson, G., J. Schein, R. Chiu, et al. 2010. "De Novo Assembly and Analysis of RNA-Seq Data." *Nature Methods* 7: 909–912.

Rodas, L. R., S. M. Sarbu, R. Bancila, D. Price, Ž. Fišer, and M. Protas. 2023. "Standing Genetic Variation as a Potential Mechanism of Novel Cave Phenotype Evolution in the Freshwater Isopod, *Asellus aquaticus*." *Evolution & Development* 25: 137–152.

Ryan, K., R. Greenway, J. Landers, L. Arias-Rodriguez, M. Tobler, and J. L. Kelley. 2023. "Selection on Standing Genetic Variation Mediates Convergent Evolution in Extremophile Fish." *Molecular Ecology* 32: 5042–5054.

Sarbu, S. M., T. Brad, R. I. Băncilă, and A. Ştefan. 2024. "Exploring Biodiversity and Food Webs in Sulfur Cave in the Vromoner Canyon on the Greek–Albanian Border." *Diversity* 16: 477.

Sarbu, S. M., C. Lascu, and T. Brad. 2019. "Dobrogea: Movile Cave." In *Cave and Karst Systems of Romania*, 429–436. Springer.

Schulz-Mirbach, T., C. Eifert, R. Riesch, et al. 2016. "Toxic Hydrogen Sulphide Shapes Brain Anatomy: A Comparative Study of Sulphide-Adapted Ecotypes in the *Poecilia mexicana* Complex." *Journal of Zoology* 300: 163–176.

Sears, C. R., T. E. Boggs, and J. B. Gross. 2020. "Dark-Rearing Uncovers Novel Gene Expression Patterns in an Obligate Cave-Dwelling Fish." *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 334: 518–529.

Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. "BUSCO: Assessing Genome Assembly and Annotation Completeness With Single-Copy Orthologs." *Bioinformatics* 31: 3210–3212.

Stahl, B. A., J. B. Gross, D. I. Speiser, et al. 2015. "A Transcriptomic Analysis of Cave, Surface, and Hybrid Isopod Crustaceans of the Species *Asellus aquaticus*." *PLoS One* 10: e0140484.

Thomas, J., University of Oxford and Wytham Woods Genome Acquisition Lab, Natural History Museum Genome Acquisition Lab, Darwin

Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team, Wellcome Sanger Institute Scientific Operations: Sequencing Operations, Wellcome Sanger Institute Tree of Life Core Informatics team, Tree of Life Core Informatics collective and Darwin Tree of Life Consortium. 2025. "The Genome Sequence of the Pond Louse, *Asellus aquaticus* (Linnaeus, 1758)." *Wellcome Open Research* 10: 291.

Tobler, M., T. J. DeWitt, I. Schlupp, et al. 2008. "Toxic Hydrogen Sulfide and Dark Caves: Phenotypic and Genetic Divergence Across Two Abiotic Environmental Gradients in *Poecilia mexicana*." *Evolution* 62: 2643–2659.

Tobler, M., J. L. Kelley, M. Plath, and R. Riesch. 2018. "Extreme Environments and the Origins of Biodiversity: Adaptation and Speciation in Sulphide Spring Fishes." *Molecular Ecology* 27: 843–859.

Tobler, M., C. N. Passow, R. Greenway, J. L. Kelley, and J. H. Shaw. 2016. "The Evolutionary Ecology of Animals Inhabiting Hydrogen Sulfide-Rich Environments." *Annual Review of Ecology, Evolution, and Systematics* 47: 239–262.

Towns, J., T. Cockerill, M. Dahan, et al. 2014. "XSEDE: Accelerating Scientific Discovery." *Computing in Science & Engineering* 16: 62–74.

Turk-Prevorčnik, S., A. Blejec, and S. Turk-Prevorcnik. 1998. "*Asellus aquaticus* Infernus, New Subspecies (Isopoda: Asellota: Asellidae), From Romanian Hypogean Waters." *Journal of Crustacean Biology* 18: 763–773.

Verovnik, R., B. Sket, and P. Trontelj. 2004. "Phylogeography of Subterranean and Surface Populations of Water Lice *Asellus aquaticus* (Crustacea: Isopoda)." *Molecular Ecology* 13: 1519–1532.

Wilkens, H. 1988. "Evolution and Genetics of Epigean and Cave *Astyanax fasciatus* (Characidae, Pisces) Support for the Neutral Mutation Theory." In *Evolutionary Biology*, 271–367. Springer.

Xie, Y., G. Wu, J. Tang, et al. 2014. "SOAPdenovo-Trans: De Novo Transcriptome Assembly With Short RNA-Seq Reads." *Bioinformatics* 30: 1660–1666.

Yancey, P. H., J. Ishikawa, B. Meyer, P. R. Girguis, and R. W. Lee. 2009. "Thiotaurine and Hypotaurine Contents in Hydrothermal-Vent Polychaetes Without Thiotrophic Endosymbionts: Correlation With Sulfide Exposure." *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* 311: 439–447.

Zerbino, D. R., and E. Birney. 2008. "Velvet: Algorithms for De Novo Short Read Assembly Using De Bruijn Graphs." *Genome Research* 18: 821–829.

Supporting Information

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

SupFiles: Infernus_6_7_25. Supplementary Figure 1: Principal Count Analysis (PCA) of count data of RNA-seq data from Asellus infernus (cave) and A. aquaticus (surface) embryos A PCA was performed on the RNA-seq count data from Asellus infernus (cave) and A. aquaticus (surface) embryos sampled at ~70% (mid-stage) and ~90% (late-stage) of embryonic development. **Supplementary Figure 2:** Embryonic development of cave samples of A. infernus. **Supplementary Figure 3:** Embryonic development of surface samples. **Supplementary Figure 4:** Gene expression matrix of genes mapped to the cave transcriptome. **Supplementary Figure 5:** Gene expression matrix of genes mapped to the surface transcriptome. **Supplementary Table 1:** Populations collected of A. aquaticus and A. infernus. **Supplementary Table 2:** Samples used for hatchling analysis. **Supplementary Table 3:** Samples used for Illumina sequencing. **Supplementary Table 4:** P-values for statistical tests comparing cave, surface, and F1 samples. **Supplementary Table 5:** Factor loadings and fraction of variance explained (FVE) for first two principal components.