

# Lawrence Berkeley National Laboratory

## LBL Publications

### Title

The Toxicogenome of *Hyalella azteca*: A Model for Sediment Ecotoxicology and Evolutionary Toxicology

### Permalink

<https://escholarship.org/uc/item/97m7340m>

### Journal

Environmental Science and Technology, 52(10)

### ISSN

0013-936X

### Authors

Poynton, Helen C  
Hasenbein, Simone  
Benoit, Joshua B  
[et al.](#)

### Publication Date

2018-05-15

### DOI

10.1021/acs.est.8b00837

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



# EPA Public Access

Author manuscript

*Environ Sci Technol.* Author manuscript; available in PMC 2019 May 15.

About author manuscripts

Submit a manuscript

Published in final edited form as:

*Environ Sci Technol.* 2018 May 15; 52(10): 6009–6022. doi:10.1021/acs.est.8b00837.

## The Toxicogenome of *Hyalella azteca*: A Model for Sediment Ecotoxicology and Evolutionary Toxicology.

HC Poynton<sup>1</sup>, S Hasenbein<sup>2</sup>, JB Benoit<sup>3</sup>, MS Sepulveda<sup>4</sup>, MF Poelchau<sup>5</sup>, DST Hughes<sup>6</sup>, SC Murali<sup>6</sup>, S Chen<sup>4,7</sup>, KM Glastad<sup>8</sup>, MAD Goodisman<sup>9</sup>, JH Werren<sup>10</sup>, JH Vineis<sup>11</sup>, JL Bowen<sup>11</sup>, M Friedrich<sup>12</sup>, J Jones<sup>12</sup>, HM Robertson<sup>13</sup>, R Feyereisen<sup>14</sup>, A Mechler-Hickson<sup>15</sup>, N Mathers<sup>15</sup>, CE Lee<sup>15</sup>, JK Colbourne<sup>16</sup>, A Biales<sup>17</sup>, JS Johnston<sup>18</sup>, GA Wellborn<sup>19</sup>, AJ Rosendale<sup>3</sup>, AG Cridge<sup>20</sup>, MC Munoz-Torres<sup>21</sup>, PA Bain<sup>22</sup>, AR Manny<sup>23</sup>, KM Major<sup>1</sup>, FN Lambert<sup>24</sup>, CD Vulpe<sup>24</sup>, P Tuck<sup>1</sup>, BJ Blalock<sup>1</sup>, YY Lin<sup>25</sup>, ME Smith<sup>26</sup>, H Ochoa-Acuña<sup>4</sup>, MM Chen<sup>25</sup>, CP Childers<sup>5</sup>, J Qu<sup>6</sup>, S Dugan<sup>6</sup>, SL Lee<sup>6</sup>, H Chao<sup>6</sup>, H Dinh<sup>6</sup>, Y Han<sup>6</sup>, H Doddapaneni<sup>6</sup>, KC Worley<sup>6,27</sup>, DM Muzny<sup>6</sup>, RA Gibbs<sup>6</sup>, and S Richards<sup>6</sup>

<sup>1</sup>School for the Environment, University of Massachusetts Boston, Boston, Massachusetts 02125 United States. <sup>2</sup>Aquatic Systems Biology Unit, Technical University of Munich, D-85354 Freising, Germany. <sup>3</sup>Department of Biological Sciences, University of Cincinnati, Cincinnati, Ohio 45221 United States. <sup>4</sup>Forestry and Natural Resources, Purdue University, West Lafayette, Indiana 47907 United States. <sup>5</sup>Agricultural Research Service, National Agricultural Library, U.S. Department of Agriculture, Beltsville, Maryland 20705 United States. <sup>6</sup>Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas 77030 United States. <sup>7</sup>OmicSoft Corporation, Cary, North Carolina 27513 United States. <sup>8</sup>Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104 United States. <sup>9</sup>School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332 United States. <sup>10</sup>Biology Department, University of Rochester, Rochester, New York 14627 United States. <sup>11</sup>Department of Marine and Environmental Sciences, Marine Science Center, Northeastern University, Nahant, Massachusetts 01908 United States. <sup>12</sup>Department of Biological Sciences, Wayne State University, Detroit Michigan 48202 United States. <sup>13</sup>Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801 United States. <sup>14</sup>Department of Plant and Environmental Sciences, University of Copenhagen, DK-1871 Frederiksberg, Denmark. <sup>15</sup>Center of Rapid Evolution (CORE) and Department of Integrative Biology, University of Wisconsin, Madison, Wisconsin 53706 United States. <sup>16</sup>School of Biosciences, University of Birmingham, Birmingham B15 2TT U.K. <sup>17</sup>National Exposure Research Laboratory, United States Environmental Protection Agency, Cincinnati, Ohio 45268 United States. <sup>18</sup>Department of Entomology, Texas A&M University, College Station, Texas 77843 United States. <sup>19</sup>Department of Biology, University of Oklahoma, Norman, Oklahoma 73019 United States. <sup>20</sup>Laboratory for Evolution and Development, Department of Biochemistry, University of Otago, Dunedin, 9054 New Zealand. <sup>21</sup>Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720 United States. <sup>22</sup>Commonwealth Scientific and

### Supporting Information

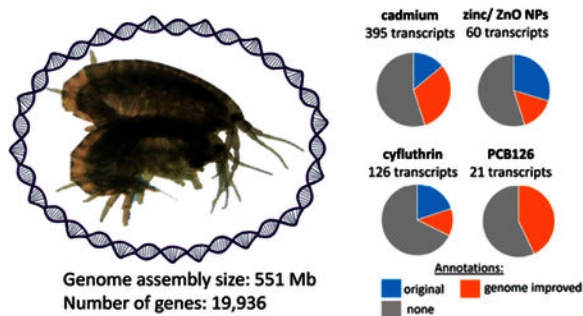
Additional method details (S1), detailed annotation reports (S2–S4), gene expression data tables and figures (S5), supplemental sequence files (S6) and detailed author contributions (S7)

Industrial Research Organisation (CSIRO), Urrbrae SA 5064 Australia. <sup>23</sup>Department of Microbiology & Cell Science, University of Florida, Gainesville, Florida 32611 United States. <sup>24</sup>Center for Environmental and Human Toxicology, Department of Physiological Sciences, University of Florida, Gainesville, Florida 32611 United States. <sup>25</sup>Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, 10617 Taiwan. <sup>26</sup>McConnell Group, Cincinnati, Ohio 45268, United States. <sup>27</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030 United States.

## Abstract

*Hyalella azteca* is a cryptic species complex of epibenthic amphipods of interest to ecotoxicology and evolutionary biology. It is the primary crustacean used in North America for sediment toxicity testing and an emerging model for molecular ecotoxicology. To provide molecular resources for sediment quality assessments and evolutionary studies, we sequenced, assembled, and annotated the genome of the *H. azteca* U.S. Lab Strain. The genome quality and completeness is comparable with other ecotoxicological model species. Through targeted investigation and use of gene expression data sets of *H. azteca* exposed to pesticides, metals, and other emerging contaminants, we annotated and characterized the major gene families involved in sequestration, detoxification, oxidative stress, and toxicant response. Our results revealed gene loss related to light sensing, but a large expansion in chemoreceptors, likely underlying sensory shifts necessary in their low light habitats. Gene family expansions were also noted for cytochrome P450 genes, cuticle proteins, ion transporters, and include recent gene duplications in the metal sequestration protein, metallothionein. Mapping of differentially expressed transcripts to the genome significantly increased the ability to functionally annotate toxicant responsive genes. The *H. azteca* genome will greatly facilitate development of genomic tools for environmental assessments and promote an understanding of how evolution shapes toxicological pathways with implications for environmental and human health.

## Graphical abstract



## Introduction

Sediment quality assessments serve as a metric for overall habitat integrity of freshwater ecosystems. Sediments provide a foundation for aquatic food webs, providing habitat for invertebrate species including crustaceans and insect larvae. However, they also concentrate

pollution over time, especially hydrophobic contaminants that sorb to sediments, leading to bioaccumulation in the food web.(1) While concentrations of a few legacy contaminants are declining in the United States due to regulatory efforts;(2) chemicals designed as their replacements are becoming emerging contaminants and their levels are increasing.(3) This is particularly true of newer generation pesticides, which have become problematic in urban areas,(4) and complex mixtures of pharmaceuticals and personal healthcare products.(5)

*Hyalella azteca* is a freshwater crustacean (Malacostraca: Amphipoda) that lives near the sediment surface, burrowing in sediment and scavenging on leaf litter, algae, and detritus material on the sediment surface.(6) The amphipod's nearly continuous contact with sediment, rapid generation time, and high tolerance to changes in temperature and salinity has made *H. azteca* an ideal species for assessing toxicity(7,8) and the bioavailability of sediment contaminants.(9) Given its ecology and expansive distribution, *H. azteca* provides an important window into sediment toxicant exposure and is a foundational trophic link to vertebrates as prey.(10,11)

The *H. azteca* species complex represents one of the most abundant and broadly distributed amphipods in North America. It was originally characterized as a single cosmopolitan species, but life history and morphological differences of *H. azteca* from different locations suggest that they comprise a species complex.(12–15) Indeed, phylogeographic analyses have resolved several different species,(16–20) which have diverged in North America over the past 11 million years.(19) Figure 1 shows the distribution of seven of the best characterized species within the complex. Even as the species have diverged, convergent evolution appears to be occurring due to similarity between geographically dispersed habitats providing an interesting study system for evolutionary biology.(21) For example, several populations representing multiple species groups have independently evolved genetic resistance to pyrethroid insecticides through mutations to the voltage-gate sodium channel (VGSC).(22) Within ecotoxicology, new strategies are being promoted to address the magnitude and wide range of effects elicited by chemicals and deficiencies in current toxicity testing approaches (e.g., National Research Council (23)). These strategies include developing adverse outcome pathway models that connect “key events” that are predictive of harmful results, from molecular perturbations to ecologically relevant effects. (24,25) In addition, comparative toxicogenomic approaches to identify evolutionarily conserved toxicological pathways(26,27) and target sites(28)enable cross-species predictions of adverse effects.

To ensure that sediment toxicity testing remains current with these emerging approaches in environmental health assessments, we set out to investigate the toxicogenome of *H. azteca* (U.S. Lab Strain), or the complete set of genes involved in toxicological pathways and stress response. As an often-used model, and the most widely used species within the complex for ecotoxicology, a relatively large amount of toxicity data has been amassed for the U.S. Lab Strain of *H. azteca*. Detailed characterization of the toxicogenome creates opportunities to reinterpret and exploit existing toxicity data to better predict risk posed to the environment by chemicals. Here we describe the genome of *H. azteca* with particular emphasis on genes related to key toxicological targets and pathways including detoxification, stress response, developmental and sensory processes, and ion transport. In addition, we begin the process of

creating a “gene ontology” for environmental toxicology using *H. azteca* by annotating the function of genes responsive to model sediment contaminants, thereby shining light on the toxicogenome underlying adverse outcome pathways.

## Materials and Methods

### *Hyalalella azteca* Strain, Inbreeding, And Genomic DNA (gDNA) Extraction

*Hyalalella azteca* (U.S. Laboratory Strain(29)) cultures were reared according to standard test conditions.(10) These organisms have been maintained by the U.S. Environmental Protection Agency since their original collection by A. Nebeker (ca. 1982) from a stream near Corvallis, Oregon.(16) They share highest genetic similarity with populations collected in Florida and Oklahoma (Figure 1, Clade 8)(29,30) and recently in California.(22) Several lines of full sibling matings were maintained for four generations, after which all lines were unable to produce offspring, likely due to inbreeding depression.(31) Twenty animals including both males and females from a single inbred line were collected and gDNA extracted from individual *H. azteca* using the Qiagen DNeasy Blood & Tissue Kit (Qiagen, Germantown, MD) with slight modifications.(22) Because of the gDNA quantities required, multiple individuals were pooled for library construction.

### gDNA Sequencing, Assembly, and Annotation

*H. azteca* is one of 30 arthropod species sequenced as part of a pilot project for the i5K Arthropod Genomes Project at the Baylor College of Medicine Human Genome Sequencing Center, Houston, TX. Sequencing was performed on Illumina HiSeq2000s (Casava v. 1.8.3\_V3) generating 100 bp paired end reads. The amount of sequences generated from each of four libraries (nominal insert sizes 180 bp, 500 bp, 3 kb, and 8 kb) is noted in Supporting Information (SI) Table S1 with NCBI SRA accessions. See SI S1 for more details on library preparation. Reads were assembled using ALLPATHS-LG (v35218)(32) and further scaffolded and gap-filled using in-house tools Atlas-Link (v.1.0) and Atlas gap-fill (v.2.2) (<https://www.hgsc.bcm.edu/software/>). This yielded an initial assembly (HAZT\_1.0; SI Table S1; NCBI Accession GCA\_000764305.1) of 1.18 Gb (596.68 Mb without gaps within scaffolds), compared with genome size of 1.05 Gb determined by flow cytometry (see SI for methods). To improve assembly contiguity, we used the Redundans(33) assembly tool. With Redundans using standard parameters, HAZT\_1.0 scaffolds and all Illumina input reads given to ALLPATHS-LG when producing HAZT\_1.0 as data inputs, generated a new assembly (HAZT\_2.0, SI Table S1; NCBI accession GCA\_000764305.2) of 550.9 Mb (548.3 Mb without gaps within scaffolds).

The HAZT\_1.0 genome assembly was subjected to automatic gene annotation using a Maker 2.0 annotation pipeline tuned specifically for arthropods. The core of the pipeline was a Maker 2 instance,(34) modified slightly to enable efficient running on our computational resources. See SI for additional details. The automated gene sets are available from the BCM-HGSC Web site (<https://hgsc.bcm.edu/arthropods/hyalalella-azteca-genome-project>), Ag Data Commons(35) and the National Agricultural Library ([https://i5k.nal.usda.gov/Hyalalella\\_azteca](https://i5k.nal.usda.gov/Hyalalella_azteca)) where a web-browser of the genome, annotations, and supporting info is accessible. The HAZT\_2.0 assembly was annotated by the automated NCBI Eukaryotic

Genome Annotation Pipeline(36) and is available from NCBI ([https://www.ncbi.nlm.nih.gov/genome/annotation\\_euk/Hyalella\\_azteca/100/](https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Hyalella_azteca/100/)).

### Manual Annotation and Official Gene Set Generation

Automated gene prediction greatly facilitates the generation of useful genomic annotations for downstream research; however, producing accurate, high-quality genome annotations remains a challenge.(37) Manual correction of gene models generated through automated analyses, also known as manual annotation can provide improved resources for downstream projects.(38) The *Hyalella* genome consortium recruited 25 annotators to improve gene models predicted from the genome assembly scaffolds,(39) adhering to a set of rules and guidelines during the manual annotation process (<https://i5k.nal.usda.gov/content/rules-web-apollo-annotation-i5k-pilot-project>). Manual annotation occurred via the Apollo software, (38) which allows users to annotate collaboratively on the web via the JBrowse genome browser, version 1.0.4(40)(<https://apollo.nal.usda.gov/hyaazt/jbrowse/>). Two transcriptomes (see below) were provided as external evidence for manual annotation and as an additional resource to search for missing genes. After manual annotation, models were exported from Apollo and screened for general formatting and curation errors (see [https://github.com/NAL-i5K/I5KNAL\\_OGS/wiki](https://github.com/NAL-i5K/I5KNAL_OGS/wiki)). Models that overlapped with potential bacterial contaminants were removed. Bacterial contamination included regions identified via the procedure outlined in the SI S1, S2 as well as potential contamination identified by NCBI (Terence Murphy, personal communication). The remaining corrected models were then merged with MAKER gene predictions HAZTv.0.5.3 and miRNA predictions (see SI S3) into a nonredundant gene set, OGSv1.0;(41) for details on the merge procedure see [https://github.com/NAL-i5K/I5KNAL\\_OGS/wiki/Merge-phase](https://github.com/NAL-i5K/I5KNAL_OGS/wiki/Merge-phase)). The manual annotation process generated 911 corrected gene models, including 875 mRNAs and 46 pseudogenes. All annotations are available for download at the i5k Workspace@NAL ([https://i5k.nal.usda.gov/Hyalella\\_azteca](https://i5k.nal.usda.gov/Hyalella_azteca)).(42) Additional details pertaining to the annotation of specific gene families can be found with the annotation reports of the SI S4.

### RNA Sequencing and Transcriptome Libraries

Two sets of transcriptomic data were generated from nonexposed *H. azteca* to assist in gene prediction, and were recently published as part of a de novo transcriptome assembly project to identify peptide hormones(43) (see S1 for details). RNaseq reads from this transcriptome project were aligned to the *H. azteca* genome scaffolds (HAZT\_1.0) using TopHat 2.0.14 with bowtie 2–2.1.0 and SAMtools 1.2. Overall mapping rate was 70.2%. Resulting BAM files were transferred to NAL and added to the Apollo genome browser.

### Gene Expression Data Sets

To assist in the manual annotation of genes related to toxicant stress, additional gene expression data sets were utilized, consisting of differentially expressed genes identified by microarray analysis following exposure to model pollutants (see Table 1). Two of these data sets were previously published(30,44) while a third data set is new. Details related to these microarray experiments can be found in the Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo); Accession number: GPL17458) and SI S1. Contigs corresponding to the differentially expressed microarray probes were aligned to the genome

using Blastn within the Apollo genome browser. Genes were manually annotated in the areas of the genome where the contigs aligned using available MAKER and AUGUSTUS gene models and RNaseq reads to correct exon and intron boundaries if needed.

## Results and Discussion

### Description of the *Hyalella azteca* Genome

Analysis of the *H. azteca* U.S. Lab strain genome size using flow cytometry gave an average genome size of 1.05 Gb (females 1C = 1045 ± 8.7 Mb; males 1C = 1061 ± 10.2 Mb). Our genome size is significantly smaller than recent estimates from *H. azteca* representing different species groups collected throughout North America, which have been shown to vary by a factor of 2. (45) Following the strategy of the i5k pilot project, we generated an assembly of 1.18 Gb (Hazt\_1.0; SI Table S1). However, because of the high gap fraction and an assembly size larger than the experimentally determined genome size, a second assembly of 550.9 Mb was later generated from the same sequencing data using Redundans (Hazt\_2.0).(33) This assembly greatly improved contig N50 and produced a more complete gene set (Figure 2, SI Table S1). Because of the timing of the availability of Hazt\_2.0, most of the manual annotations were performed using Hazt\_1.0; therefore, both assemblies are presented here. RNaseq reads(43) mapped equally well to both genome assemblies (Figure 2E), but significantly more reads mapped to the improved gene models of the Hazt\_2.0 assembly (Figure 2F). BUSCO analysis(46) was performed on both *H. azteca* genome assemblies (Figure 2C) and predicted gene sets (Figure 2D) to assess the completeness of the genome. Hazt\_2.0 contains a higher percentage (Genome = 91.0%, Gene set = 94.2%) of complete BUSCOs in contrast to the Hazt\_1.0 assembly (Genome 85.3%, Gene set = 67.6%). When compared to other genomes of ecotoxicological relevance,(47–53) the Hazt\_2.0 showed comparable quality, ranking fourth in completeness out of the eight genomes assessed.

### Associated Bacteria and Lateral Gene Transfers

Using two complementary approaches to screen the *H. azteca* genome for bacterial contaminants, (54,55) we recovered two draft bacterial genomes, or metagenome assembled genomes (MAGs), and evidence of lateral gene transfers (LGTs) (SI S2). MAG1 is a flavobacterium with distant affinities to currently sequenced bacteria (92% 16S rRNA identity to the most closely identified genera *Chishuiella* and *Empedobacter*). MAG2 is related to bacteria in the genus *Ideonella* (98% 16S rRNA identity to *I. paludis*), of which some are able to degrade plastics, (56) in particular an isolate from the wax moth *Galleria mellonella*.(57) Whether these bacteria are close associates of *H. azteca* or components of their diet is not currently known, but their interesting gene repertoires could be relevant to *H. azteca* ecotoxicology. An analysis of broad functional categories indicates that both genomes contain multiple genes related to metal detoxification and resistance to toxins (i.e., antibiotics) and possibly other organic pollutants (SI Figure S2.2). In addition, strong LGT candidates from Rickettsia-like bacteria were found on five genome scaffolds (SI Table S2.2).

## Genome Methylation and MicroRNAs

We performed two genome-wide analyses to characterize DNA methylation patterns in the genome and characterize the full complement of *H. azteca* microRNAs. DNA methylation is an epigenetic mechanism by which a methyl group (CH<sub>3</sub>) binds to DNA that may alter gene expression.(58) Because methylated cytosines tend to mutate to thymines over evolutionary time, we used signatures of CpG dinucleotide depletion in the *H. azteca* genome to uncover putative patterns of DNA methylation. Our analyses showed that *H. azteca* possesses strong indications of genomic DNA methylation, including CpG depletion of a subset of genes (SI Figure S3.1 A,B) and presence of the key DNA methyltransferase enzymes, DNMT1 and DNMT3 (SI Table S3.1). Furthermore, genes with lower levels of CpG observed/expected (CpG o/e; i.e., putatively methylated) displayed a strong positional bias in CpG depletion (SI Figure S3.1 C, black line), with 5' regions of these genes being considerably depleted of CpGs. In contrast, genes with higher CpG o/e displayed no such positional bias (SI Figure S3.1 C, gray line). Several insects where DNA methylation has been empirically profiled at the single-base level possess such patterns(59)suggesting *H. azteca* has similar patterns of DNA methylation as most insects (but see Glastad et al.(60)).

MicroRNAs (miRNAs) are a family of short noncoding RNAs (~22 nt in length) that play critical roles in post-translational gene regulation.(61) Recent research revealed that miRNAs are involved in aquatic crustaceans' response to environmental stressors (e.g., hypoxia and cadmium exposure),(62,63) which makes miRNAs promising biomarkers for future aquatic toxicological research. We predicted *H. azteca* miRNAs based on sequence homology and hairpin structure identification. A total of 1261 candidate miRNA coding sites were identified by BLAST. After hairpin structure identification, we predicted 148 *H. azteca* miRNAs, which include several highly conserved miRNAs (e.g., miR-9 and let-7 family) (SI Table S3.2, sequences available in SI S6.1). Several Cd-responsive miRNAs in *D. pulex* (miR-210, miR-71 and miR-252)(62) were also predicted in *H. azteca*, suggesting a conserved role of these miRNAs. This number of predicted miRNAs is comparable to what has been reported for other arthropods (SI Figure S3.2).

## Functional Annotation

The manual annotation of the *H. azteca* U.S. Lab strain genome resulted in the characterization of 13 different gene families (see detailed annotation reports in SI S4.1-S4.13). Given its importance in ecological and ecotoxicological studies, a particular focus was given to genes involved in environmental sensing (chemoreceptors and opsins), detoxification and response to stress (cytochrome P450s, cuticle proteins, glutathione peroxidases, glutathione S-transferases, heat shock proteins, and metallothionein proteins), as well as genes involved in important toxicological pathways (ion transporters, early development genes, insecticide target genes, and nuclear receptors). Here we highlight significant findings of gene family expansions or contractions as well as the characterization of genes of particular toxicological importance.

## Environmental Sensing

To better understand how *H. azteca* interacts with its environment, we annotated genes involved in light and chemical sensing. Arthropods deploy a diversity of light sensing



mechanisms including light capture in photoreceptor cells through the expression of opsins: light-sensitive, seven-transmembrane G-protein coupled opsin receptor proteins. Our survey of the *H. azteca* genome revealed only three opsin genes, a middle wavelength-sensitive subfamily 1 (MWS1) opsin and two belonging to the long wavelength-sensitive subfamily (LWS) opsins (SI Figure S4.13.1, Table S4.13.1). Maximum likelihood analysis with a subset of closely related malacostracan LWS opsins moderately supports the two *H. azteca* LWS opsins as 1:1 orthologs of the two LWS opsins previously reported for *Gammarus minus*.(64) Thus, the LWS opsin duplicate pair conserved in *H. azteca* and *G. minus* is likely ancient, predating at least the origin of amphipod Crustacea. The *H. azteca* MWS opsin by contrast, represents the first reported amphipod MWS opsin and is distinct from currently known malacostracan MWS opsins. These findings suggest that amphipod crustaceans are equipped with a minimally diversified set of three opsin genes and implies gene family losses for all of the nonretinal opsin subfamilies. This is in contrast to the 46 opsin genes characterized in *Daphnia pulex*.(47) It remains to be seen whether these candidate gene losses are associated with the adaptation of *H. azteca* to its crepuscular visual ecology or reflect a more ancient trend in amphipods.

In contrast, the *H. azteca* genome reveals gene expansions of chemoreceptors, which may be essential for *H. azteca* given its epibenthic ecology and close association with sediments. (6)Noninsect arthropods have two major families of chemoreceptors: gustatory receptor (GR) family, an ancient lineage extending back to early animals,(49,65–67) and the ionotropic receptors (IRs) that are a variant lineage of the ancient ionotropic glutamate receptor superfamily known only from protostomes.(49,68) These two gene families were manually annotated in the *H. azteca* genome and improved models were generated for two other crustaceans, *D. pulex*(69,70) and *Eurytemora affinis*(49,71) for comparison (SI S4.1, sequences available in SI S6.2). With 155 GR genes, *H. azteca* has over twice the number of GRs compared with *D. pulex* (59)(69) and *E. affinis* (67), although many of the most recent gene duplicates are pseudogenes. Two candidate GR sugar receptors were identified in *H. azteca* and *D. pulex* (independently duplicated in both lineages), but not in *E. affinis*. Otherwise these crustacean GRs form large species-specific expanded clades with no convincing orthology with each other or other conserved insect GRs such as the Gr43a fructose receptor (SI Figure S4.1.1). *H. azteca* has 118 IR genes (two pseudogenic) compared with updated totals of 154 in *D. pulex* (26 pseudogenic) and 22 intact genes for *E. affinis*. All three species contain single copy orthologs of the highly conserved IR genes implicated in perception of salt, amines, amino acids, humidity, and temperature in insects, including Ir25a, Ir8a (missing in *D. pulex*), Ir76b,(72) and Ir93a.(73) The remaining divergent IRs form largely species-specific expanded lineages (SI Figure S4.1.2). The many divergent IRs and GRs in these three crustaceans presumably mediate most of their chemical sense capability, but their great divergence from the proteins of *Drosophila* for which functions are known precludes speculation as to specific roles.

### Cytochrome P450s

The cytochrome P450 superfamily of genes (P450 genes) is ubiquitous and diverse as they have been found in all domains of life and are thought to have originated over 3 billion years ago.(74)P450 genes function in metabolizing a wide range of endogenous and exogenous

compounds, including toxins, drugs, plant metabolites, and signaling molecules.(75–78) In the *H. azteca* genome, we found 70 genes or gene fragments that contained a typical P450 signature (FxxGxxxC), where C is the heme thiolate ligand (SI S4.3, sequences available in SI S6.2). However, only 27 were complete genes. The 70 P450 genes were classifiable into one of four recognized P450 clans, with the CYP2 clan (SI Figure S4.3.1) being the largest with 48 genes. The most notable difference between the P450 complement of *H. azteca* relative to hexapods (insects) was the expansion of the CYP2 clan P450s. Typical of expanded clades, we found several clusters of genes (and gene fragments) of the CYP2 clan. The CYP3 and CYP4 clans in *H. azteca* were represented by eight and seven genes, respectively. The fourth P450 clan is the mitochondrial P450 clan, with at least nine genes in *H. azteca*. The number of P450s found in *H. azteca* was greater to those found in other crustaceans, including the copepods *Tigriopus japonicus* (52)(79) and *Paracyclopsina nana* (46),(80) but somewhat fewer than those in hexapod taxa (including 106 in the mosquito *Anopheles gambiae* and 81 in the silkworm *Bombyx mori*).(81)

### Heat Shock Proteins

The heat shock protein (HSP) molecular chaperones are a highly conserved family of proteins that facilitate the refolding of denatured proteins following stress, including thermal stress, but also in response to metals and other toxicants, oxidative stress, and dehydration. (82) HSPs are divided into several families based on their molecular weight. Of the different families, HSP70, HSP90, and HSP60 play a major role in protein refolding while HSP40/J-protein is a cofactor to HSP70 and delivers nonnative proteins to HSP70.(83) HSPs were identified and annotated for each of these families (SI S4.8). The number of *hsp70* (8 genes), *hsp90* (3), *hsp40* (3), and *hsp60* (1) was well within the expected number found throughout Arthropoda.(84) Of the eight *hsp70* genes, five were found as a gene cluster on scaffold 277, which is similar to gene clusters identified in *Drosophila melanogaster*(85) and *Aedes aegypti*.(86) In agreement with Baringou et al.,(87) the HSP70 proteins described here cannot be easily divided into inducible and cognate forms based on sequence characteristics. We instead decided to compare our eight sequences to sequence motifs described by Baringou et al.(87) and classify the *H. azteca* HSP70s according to their framework (SI Table S4.8.2). According to these motifs and the classification methods described, all *H. azteca* sequences belong to Group A, which agrees with Baringou et al.(87) finding that all amphipod HSP70s characterized to date are Group A proteins. One HSP70 contained slightly different motif characteristics and was grouped with A4 proteins, while the remaining sequences were grouped together in A5.

### Metallothionein Genes

Metallothioneins (MTs) are a group of conserved metalloproteins with a high capacity for binding metal ions. These proteins are characterized by their low molecular weight (<10 kDa), cysteine rich composition (often over 30%), lack of secondary structure in the absence of bound metal ions, and a two domain structure dictated by the bound ions. Although their diversity makes it difficult to assign a specific function by class of MTs, their ability to bind metal ions has provided MTs with a role in detoxification, binding, and sequestration of toxic metals.(88) Four *mt* genes were identified in the *H. azteca* genome by mapping Cd responsive contigs with homology to *Callinectes sapidus*CdMT-1 (AAF08964) to the

HAZT\_1.0 assembly (SI S4.11). These four genes were arranged as repeats on scaffold 460 and each contained three exons, the typical gene structure of *mts* (SI Figure S4.11.1). *Mt-b* and *mt-d* produce identical proteins of 61 amino acids, whereas *mt-c* is missing the downstream splice site on exon 1 and produces a truncated protein of 53 amino acids. *Mt-a* lacks a viable start codon, making it a likely pseudogene. Due to the similarity in the sequences of the remaining three genes, it is not possible to determine if they are all transcribed or regulated differently based solely on the RNaseq mapped reads. However, given their high degree of similarity and arrangement on scaffold 460, these genes are likely the result of recent gene duplications, which may provide an evolutionary advantage against high metal exposure. 1–4 *mt* genes have been identified in at least 35 other Malacostracan species. However, in most cases, the multiple MTs are not identical in amino acid sequence. For example, in the blue crab *C. sapidus* there are three *mt* genes. Two encode for Cd inducible forms with 76% sequence identity, while a third, codes for a longer copper-inducible form.(89) Given that our strategy for identifying the *mt* genes in *H. azteca* relied on using the Cd inducible gene expression set, it is possible that a fourth, copper inducible form also exists in its genome.

### Ion Transport Proteins

In arthropods, a subset of ion transporters are integral in maintaining cellular homeostasis and regulating epithelial transport of common ions such as H<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>, and Cl<sup>-</sup>(90–93) and are likely involved in the toxicity and uptake of metal ions. The proton pump V-type H<sup>+</sup> ATPase (VHA, ATP6) is an evolutionarily conserved molecular machine having a wide range of functions. VHA actively translocates H<sup>+</sup> across the membranes of cells and organelles allowing it to generate electrochemical H<sup>+</sup> gradients(94) that drive H<sup>+</sup>-coupled substrate transport of common bioavailable cations (Na<sup>+</sup>, K<sup>+</sup>, Li<sup>+</sup>). (95) VHA is a large, two domain protein complex (V1 and V0) comprised of 13 subunits, which are ubiquitous in eukaryotes and thought to be expressed in virtually every eukaryotic cell.(96) These 13 VHA subunits were identified in the *H. azteca* genome, but two accessory subunits were not (SI S4.10). A previous comparative analysis identified a wide range of VHA genes in the genomes of *D. melanogaster* (33), human (24), mouse (24), *C. elegans* (19), *Arabidopsis* (28), and *Saccharomyces* (15).(97) The high number of VHA genes identified in those organisms are in stark contrast to the only 13 VHA subunit genes present in *H. azteca*.

The sodium/hydrogen antiporters (NHA, SLC9B2, CPA2) are a subfamily of transmembrane ion transporters, which was only recently discovered in animal genomes and characterized in mosquito larvae.(98–100) In both arthropods and mammals, evidence indicates that NHA is coupled to VHA as a secondary electrogenic transporter for ion uptake against concentration gradients.(99,101–104) The presence of four NHA genes in the *H. azteca* genome was unexpected, as only two NHA paralogs per genome had been found previously in animal genomes (SI S4.10).(98)

The minimal set of VHA subunits and expansion of NHA genes may have toxicological significance, particularly with respect to metal toxicity and transport. Metal speciation, and therefore toxicity, is highly pH-dependent.(105) As a regulator of pH at the epithelial membrane, and thus electrochemical transmembrane gradients, VHA may play an important

role in metal uptake, ion speciation, and solubility.(106–108) The specific substrates for NHA have not yet been fully characterized for most species, but sodium and chloride are likely candidates.(109) The presence of NHA gene duplicates in the *H. azteca* genome suggest a likely adaptation for ion uptake against transmembrane concentration gradients, but may also influence metal uptake and bioavailability (see SI S4.10 for further discussion). As many toxic metal ions are transported across the cell membrane via ion channels and/or ion transporters (e.g., ATPases),(110,111)additional transporters should be explored for their direct roles in metal uptake and indirect influence on metal toxicity.

### The Toxico-Responsive Genome

Contaminants such as heavy metals, organic compounds, and nanoparticles can adversely affect ecologically relevant organisms. In *H. azteca*, heavy metal contaminants (Zn, Cd) have been shown to have negative effects on the development of population growth rate, longevity, and reproduction(112–114) while metal-based nanomaterials (i.e., ZnO NMs) cause increased toxicity which may be related to enhanced bioavailability.(44,115) Organic compounds such as pyrethroid insecticides and polychlorinated biphenyls (PCBs) can cause detrimental effects on behavior, reproduction, and development.(116–118)

A primary goal of the *H. azteca* genome project was to expand the functional gene annotations for transcripts that respond to toxicant stress, referred to as the toxicogenome. We utilized two published gene expression studies (Zn, ZnO, NPs,(44) cyfluthrin(30)) as well as two unpublished gene expression sets for Cd and PCB exposure (Table 1). During our original investigation, we were only able to annotate a small fraction of the differentially expressed transcripts (Cd: 13%; Zn: 29%; PCB:0%; Cyfluthrin: 20%) (Figure 3A). The ability to align these transcripts to the *H. azteca* genome allowed us to identify full length transcripts and more completely assemble transcripts aligning to the same genic region of the genome in a way that was not possible with a de novo transcriptome assembler (i.e., Newbler, Roche). This increased our ability to predict gene function increasing the fraction of annotated transcripts by 10–32% (Figure 3A). However, we also note that for each of these chemical challenges, over half of the genes are still without annotations (Figure 3A), implying that these genes may be lineage specific. This is similar to the finding within the *D. pulex* genome that lineage specific genes were more likely to be differentially expressed following environmental challenges.(47)

### Cadmium

To explore the gene expression response and further annotate genes involved in heavy metal exposure, we conducted a gene expression study at ecologically relevant concentrations of Cd. Compared to controls, 116 genes were up-regulated in expression and 9 were down-regulated by Cd. These genes are related to several cell processes including digestion, oxygen transport, cuticular metabolism, immune function, acid–base balance, visual-sensory perception and signal transduction (SI Table S5.1). Categorizing the genes by biological processes illustrated that the metabolism of *H. azteca* was very broadly affected, with the cellular metabolic processes representing the largest GO term (Figure 3B). Heat shock proteins (general stress response) were significantly upregulated in response to Cd (see SI Table S4.8.1) consistent with other amphipod studies(119,120) and showing a similar

response to other stressors including heat stress, oxidative stress and changes in pH. (121,122) In addition, expression of the newly described MT genes were also significantly induced over 15-fold by Cd (SI Figure S4.11.3). Cd exposure also induced expression of genes involved in oxidative stress including glutathione-*S*-transferase, a commonly used biomarker in toxicity tests of pollutant exposure and oxidative stress, and thioredoxin peroxidase, a gene involved in protection against reactive oxygen species (ROS). Finally, genes involved in regulation of the cuticle were also upregulated. Differential expression of Chitinase and other cuticular proteins has been demonstrated previously in crustaceans in response to stress and has been correlated with impacts to growth and reproduction. (123,124)

## Zinc

We utilized a data set originally published in Poynton et al.(44) that compared the toxicity of ZnO NPs to zinc sulfate (ZnSO<sub>4</sub>) to increase the number of annotated genes that were responsive to metal exposure. Of the 60 differentially expressed genes, we annotated 25, including 15 genes that had not been annotated in the original publication (SI Table S5.2, Figure 3A). For example, chorion peroxidase (contig18799 in Poynton et al.(38)) was induced in both the ZnSO<sub>4</sub> and ZnO NP exposures and acts as an indicator of oxidative stress, as the gene is involved in ROS damage repair. Contig000192 in Poynton et al.(44) is another previously uncharacterized gene that was annotated as asparaginyl beta-hydroxylase-like protein, a regulator of muscle contraction and relaxation; its dysregulation suggests negative impacts to swimming behavior and movement. With the additional annotation results we were able to perform gene ontology analysis (SI Table S5.2) and observed that most genes were mapped to GO:0042221, response to chemical (Figure 3C).

## Cyfluthrin

The pyrethroid insecticide cyfluthrin is one the most widely applied insecticides worldwide(125,126) and has been shown to be highly toxic to *H. azteca*. (EC<sub>50</sub> < 1 ng/L). (116) We previously showed that cyfluthrin exposure at 1 ng/L caused differential expression of 127 sequences.(30) Through the reanalysis of this data set, we were able to annotate 33 genes and successfully mapped them to GO terms (SI Table S5.3). Many affected genes were consistent with the known mechanism of pyrethroid toxicity, showing involvement in neurological system processes, synapse organization, and transmission of nerve impulses, but also stress response such as oxidization processes, damage repair, maintaining of homeostasis, and immune response (Figure 3D).

## PCB126

For *H. azteca*, PCBs represent a major exposure and accumulation threat due to their habitat and feeding behavior in benthic areas. In fishes, dioxin-like PCBs (e.g., PCB126) are highly toxic as they bind to the aryl hydrocarbon receptor and induce the expression of CYP1 genes.(127) Much less is known about this mechanism in crustaceans,(128) but in general they appear more tolerant of PCBs. Following exposure of *H. azteca* to PCB126, the most potent and ubiquitous of the PCB congeners,(129) we identified 21 differentially expressed sequences, representing seven genes, of which five were annotated (SI Table S5.4). Three of the five characterized genes are transmembrane proteins, while two are involved in

endocrine processes (growth hormone, thyroid hormone). Neuroendocrine disruption of PCBs was described in crustaceans previously (see review in(130)); however, most investigations on neuroendocrine disruption to date focus on effects in vertebrates. Our study demonstrates that potential impacts of neuroendocrine disruption on invertebrates deserves further attention.

In summary, with a total of 19936 genes including 911 manually curated genes, the genome of the *Hyalella azteca* U.S. Lab strain provides a foundational tool for understanding the molecular ecology of benthic invertebrates as well as the mechanisms of toxicity of sediment associated pollutants. The critical gene families annotated here will serve as basis for studying toxicologically conserved pathways in invertebrates and developing adverse outcome pathways for sediment dwelling organisms. Overall, our results illustrate the advantage of applying a genome assembly to ecotoxicogenomic studies including the improved ability to annotate genes of interest and we strongly encourage the expansion of genomic, not just transcriptomic, resources for other species of ecotoxicological relevance.

The ever-growing list of chemical contaminants entering the environment poses a significant challenge in terms of risk assessment. The low-throughput and high cost of traditional toxicity testing, suggests that the need for alternative means to assess risk.(23) The characterization of ‘omics responses has emerged as a potential alternative.(25) Measures on the cellular level provide valuable information on the mode of action of uncharacterized chemicals, the health status of exposed organisms, and can act as a means to extrapolate beyond model organisms, and can be integrated into predictive risk models. The interpretation of these omics responses within the context of the well-defined *H. azteca* genome described herein will greatly expand the utility and applications of omics responses to sediment ecotoxicology and risk assessment.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank the staff at the Baylor College of Medicine Human Genome Sequencing Center for their contributions, Terence Murphy at NCBI for assisting in screening for bacterial contamination in the assembly and with removing contaminants from the genome, and Dawn Oliver at Purdue University for assistance with the distribution map (Figure 1). Additional thanks to Mitch Kostich of the U.S. EPA for his thoughtful comments on the manuscript. Funding for genome sequencing, assembly and automated annotation was provided by NHGRI grant U54 HG003273 to RAG. Funding for transcriptome sequencing was provided by NSF grant CBET-1437409 to HCP. Additional funding was provided by the State and Federal Contractors Water Agency (contract no. 15–14) to Richard Connon, UC Davis (supported S.H.), University of Cincinnati Faculty Development Research grant and partial support from NSF grant DEB165441 to JBB, and NSF grant DEB1257053 to J.H.W. M.M.T.’s work with Apollo was supported by NIH grants 5R01GM080203 from NIGMS, and 5R01HG004483 from NHGRI, and by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02–05CH11231. The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

## References

1. Chapman PM Current approaches to developing sediment quality criteria. *Environ. Toxicol. Chem* 1989,8 (7), 589–599, 10.1002/etc.5620080706

2. US EPA. The Incidence and Severity of Sediment Contamination in Surface Waters of the United States, EPA-823-R-04-007; Office of Science and Technology, US Environmental Protection Agency: Washington D.C., 11, 2004.
3. Gavrilescu M ; Demnerová K ; Aamand J ; Agathos S ; Fava F Emerging pollutants in the environment: present and future challenges in biomonitoring, ecological risks and bioremediation. *New Biotechnol* 2015,32 (1), 147–156, 10.1016/j.nbt.2014.01.001
4. Stone WW ; Gilliom RJ ; Ryberg KR Pesticides in US Streams and Rivers: Occurrence and Trends During 1992–2011 In ACS Publications, 2014.
5. Lee KE ; Langer SK ; Menheer MA ; Foreman WT ; Furlong ET ; Smith SG Chemicals of emerging concern in water and bottom sediment in Great Lakes areas of concern, 2010 to 2011-Collection methods, analyses methods, quality assurance, and data; 723; Reston, VA, 2012; p 36.
6. Wang F ; Goulet RR ; Chapman PM Testing sediment biological effects with the freshwater amphipod *Hyaella azteca*: the gap between laboratory and nature. *Chemosphere* 2004, 57, 1713–1724, 10.1016/j.chemosphere.2004.07.05015519418
7. Norberg-King TJ ; Sibley PK ; Burton GA ; Ingersoll CG ; Kemble NE ; Ireland S ; Mount DR ; Rowland CD Interlaboratory evaluation of *Hyaella azteca* and *Chironomus tentans* short-term and long-term sediment toxicity tests. *Environ. Toxicol. Chem* 2006, 25 (10), 2662–2674, 10.1897/05-044R2.117022407
8. Dussault EB ; Balakrishnan VK ; Sverko E ; Solomon KR ; Sibley PK Toxicity of human pharmaceuticals and personal care products to benthic invertebrates. *Environ. Toxicol. Chem* 2008, 27 (2),425–432, 10.1897/07-354R.118348646
9. You J ; Pehkonen S ; Weston DP ; Lydy MJ Chemical availability and sediment toxicity of pyrethroid insecticides to *Hyaella azteca*: Application to field sediment with unexpectedly low toxicity. *Environ. Toxicol. Chem* 2008, 27 (10), 2124–2130, 10.1897/08-016.118419174
10. U.S. EPA. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates; US Environmental Protection Agency: 2000.
11. Taylor LN ; Novak L ; Rendas M ; Antunes PMC ; Scroggins RP Validation of a new standardized test method for the freshwater amphipod *Hyaella azteca*: Determining the chronic effects of silver in sediment. *Environ. Toxicol. Chem* 2016, 35 (10), 2430–2438, 10.1002/etc.345327062160
12. Strong DR Life history variation among populations of an Amphipod (*Hyaella azteca*). *Ecology* 1972, 53(6), 1103–1111, 10.2307/1935422
13. Wellborn GA The Mechanistic Basis of Body-Size Differences between 2 *Hyaella* (Amphipoda) Species. *J. Freshwater Ecol* 1994, 9 (2), 159–168, 10.1080/02705060.1994.9664443
14. Wellborn GA Size-Biased Predation and Prey Life-Histories - a Comparative-Study of Fresh-Water Amphipod Populations. *Ecology* 1994, 75 (7), 2104–2117, 10.2307/1941614
15. Wellborn GA Predator Community Composition and Patterns of Variation in Life-History and Morphology among *Hyaella* (Amphipoda) Populations in Southeast Michigan. *Am. Midl. Nat* 1995, 133 (2), 322–332, 10.2307/2426397
16. Duan YH ; Guttman SI ; Oris JT Genetic differentiation among laboratory populations of *Hyaella azteca*: Implications for toxicology. *Environ. Toxicol. Chem* 1997, 16 (4), 691–695, 10.1897/1551-5028(1997)016<0691:GDALPO>2.3.CO;2
17. Thomas PE ; Blinn DW ; Keim P Genetic and behavioural divergence among desert spring amphipod populations. *Freshwater Biol* 1997, 38 (1), 137–143, 10.1046/j.1365-2427.1997.00199.x
18. Hogg ID ; Larose C ; de Lafontaine Y ; Doe KG Genetic evidence for a *Hyaella* species complex within the Great Lakes St Lawrence River drainage basin: implications for ecotoxicology and conservation biology. *Can. J. Zool* 1998, 76 (6), 1134–1140, 10.1139/z98-034
19. Witt JD ; Hebert PD Cryptic species diversity and evolution in the amphipod genus *Hyaella* within central glaciated North America: a molecular phylogenetic approach. *Can. J. Fish. Aquat. Sci* 2000, 57 (4),687–698, 10.1139/f99-285
20. Duan YH ; Guttman SI ; Oris JT ; Bailer AJ Genotype and toxicity relationships among *Hyaella azteca*: I. Acute exposure to metals or low pH. *Environ. Toxicol. Chem* 2000, 19 (5), 1414–1421, 10.1897/1551-5028(2000)019<1414:GATRAH>2.3.CO;2

21. Wellborn GA ; Cothran R ; Bartholf S Life history and allozyme diversification in regional ecomorphs of the *Hyalella azteca* (Crustacea: Amphipoda) species complex. *Biol. J. Linn. Soc* 2005, 84 (2), 161–175,10.1111/j.1095-8312.2005.00422.x
22. Major KM ; Weston DP ; Lydy MJ ; Wellborn GA ; Poynton HC , Unintentional exposure to terrestrial pesticides drives widespread and predictable evolution of resistance in freshwater crustaceans. *Evol. Appl* 2017, (in press), 1–14. 10.1111/eva.1258410.1111/eva.12584. DOI: 10.1111/eva.12584
23. National Research Council. *Toxicity Testing in the 21st Century: A Vision and a Strategy*: National Academies Press: 2007.
24. Ankley GT ; Bennett RS ; Erickson RJ ; Hoff DJ ; Hornung MW ; Johnson RD ; Mount DR ;Nichols JW ; Russom CL ; Schmieder PK ; Serrano JA ; Tietge JE ; Villeneuve DL Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem* 2010, 29 (3), 730–41, 10.1002/etc.3420821501
25. Villeneuve DL ; Garcia-Reyero N Vision & strategy: Predictive ecotoxicology in the 21st century. *Environ. Toxicol. Chem* 2011, 30 (1), 1–8, 10.1002/etc.39621182100
26. Goldstone J ; Hamdoun A ; Cole B ; Howard-Ashby M ; Nebert D ; Scally M ; Dean M ; Epel D ; Hahn M ; Stegeman J The chemical defensome: environmental sensing and response genes in the *Strongylocentrotus purpuratus* genome. *Dev. Biol* 2006, 300 (1), 366–384, 10.1016/j.ydbio.2006.08.06617097629
27. Perkins EJ ; Ankley GT ; Crofton KM ; Garcia-Reyero N ; LaLone CA ; Johnson MS ; Tietge JE ;Villeneuve DL Current perspectives on the use of alternative species in human health and ecological hazard assessments. *Environ. Health Perspect* 2013, 121 (9), 1002, 10.1289/ehp.130663823771518
28. LaLone CA ; Villeneuve DL ; Burgoon LD ; Russom CL ; Helgen HW ; Berninger JP ;Tietge JE ; Severson MN ; Cavallin JE ; Ankley GT Molecular target sequence similarity as a basis for species extrapolation to assess the ecological risk of chemicals with known modes of action. *Aquat. Toxicol* 2013, 144, 141–154, 10.1016/j.aquatox.2013.09.00424177217
29. Major K ; Soucek DJ ; Giordano R ; Wetzel MJ ; Soto-Adames F The common ecotoxicology laboratory strain of *Hyalella azteca* is genetically distinct from most wild strains sampled in Eastern North America. *Environ. Toxicol. Chem* 2013, 32 (11), 2637–47, 10.1002/etc.235523955770
30. Weston DP ; Poynton HC ; Wellborn GA ; Lydy MJ ; Blalock BJ ; Sepulveda MS ;Colbourne JK Multiple origins of pyrethroid insecticide resistance across the species complex of a nontarget aquatic crustacean. *Proc. Natl. Acad. Sci. U. S. A* 2013, 110 (41), 16532–7, 10.1073/pnas.130202311024065824
31. Powell JR ; Evans BR How Much Does Inbreeding Reduce Heterozygosity? Empirical Results from *Aedes aegypti*. *Am. J. Trop. Med. Hyg* 2017, 96 (1), 157–158, 10.4269/ajtmh.16-069327799643
32. Gnerre S ; MacCallum I ; Przybylski D ; Ribeiro FJ ; Burton JN ; Walker BJ ; Sharpe T ; Hall G ;Shea TP ; Sykes S High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci. U. S. A* 2011, 108 (4), 1513–1518, 10.1073/pnas.101735110821187386
33. Przyczek LP ; Gabaldón T Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Res* 2016, 44 (12), e113–e113, 10.1093/nar/gkw29427131372
34. Cantarel BL ; Korf I ; Robb SM ; Parra G ; Ross E ; Moore B ; Holt C ; Alvarado AS ; Yandell M MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* 2008, 18 (1), 188–196, 10.1101/gr.674390718025269
35. Hughes DST ; Richards S ; Poynton H *Hyalella azteca* Genome Annotations v0.5.3. *Ag Data Commons*, 2018, 10.15482/USDA.ADC/1415993.
36. Thibaud-Nissen F ; Souvorov A ; Murphy T ; DiCuccio M ; Kitts P , Eukaryotic Genome Annotation Pipeline. In *The NCBI Handbook*, 2nd ed.; National Center for Biotechnology Information (US): Bethesda, MD, 2013.
37. Yandell M ; Ence D A beginner’s guide to eukaryotic genome annotation. *Nat. Rev. Genet* 2012, 13 (5),329, 10.1038/nrg317422510764



38. Lee E ; Helt GA ; Reese JT ; Munoz-Torres MC ; Childers CP ; Buels RM ; Stein L ; Holmes IH ;Elsik CG ; Lewis SE Web Apollo: a web-based genomic annotation editing platform. *Genome Biol*2013, 14 (8), R93, 10.1186/gb-2013-14-8-r9324000942
39. Poynton H ; Gibbs RA ; Worley KC ; Murali SC ; Lee SL ; Muzny DM ; Hughes DST ; Chao H ;Dinh H ; Doddapaneni H ; Qu J ; Dugan S ; Han Y ; Richards S Hyalella azteca Genome Assembly 1.0. Ag Data Commons, 2018, 10.15482/USDA.ADC/1415994.
40. Skinner ME ; Uzilov AV ; Stein LD ; Mungall CJ ; Holmes IH JBrowse: a next-generation genome browser. *Genome Res* 2009, 19 (9), 1630–1638, 10.1101/gr.094607.10919570905
41. Poynton H ; Bain P ; Benoit J ; Boucher P ; Chen M-JM ; Chen S ; Christie AE ; Connolly J ;Cridge A ; Etxabe A ; Glazier A ; Hasenbein S ; Hughes D ; Jennings E ; Jones J ; Kearns P ;Lambert F ; Lin Y-Y ; Major C ; Manny A ; Mathers N ; Mechler-Hickson A ; Munoz-Torres M ;Poelchau M ; Robertson H ; Scanlan J ; Szuter E ; Werren J ; Wood SA ; Winchell K ; Richards S Hyalella azteca Official Gene Set v1.0. Ag Data Commons, 2018, 10.15482/USDA.ADC/1415995.
42. Poelchau M ; Childers C ; Moore G ; Tsavatapalli V ; Evans J ; Lee C-Y ; Lin H ; Lin J-W ; Hackett K The i5k Workspace@ NAL—enabling genomic data access, visualization and curation of arthropod genomes. *Nucleic Acids Res* 2015, 43 (D1), D714–D719, 10.1093/nar/gku98325332403
43. Christie AE ; Cieslak MC ; Roncalli V ; Lenz PH ; Major KM ; Poynton HC Prediction of a peptidome for the ecotoxicological model *Hyalella azteca* (Crustacea; Amphipoda) using a de novo assembled transcriptome. *Mar. Genom* 2018, 38, 67–88, 10.1016/j.margen.2017.12.003
44. Poynton HC ; Lazorchak JM ; Impellitteri CA ; Blalock B ; Smith ME ; Struewing K ; Unrine J ;Roose D Toxicity and transcriptomic analysis in *Hyalella azteca* suggests increased exposure and susceptibility of epibenthic organisms to zinc oxide nanoparticles. *Environ. Sci. Technol* 2013, 47 (16),9453–60, 10.1021/es401396t23889737
45. Vergilino R ; Dionne K ; Nozais C ; Dufresne F ; Belzile C Genome size differences in *Hyalella* cryptic species. *Genome* 2012, 55 (2), 134–139, 10.1139/g11-08522263854
46. Simão FA ; Waterhouse RM ; Ioannidis P ; Kriventseva EV ; Zdobnov EM BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015, 31 (19), 3210–3212, 10.1093/bioinformatics/btv35126059717
47. Colbourne JK ; Pfrender ME ; Gilbert D ; Thomas WK ; Tucker A ; Oakley TH ; Tokishita S ;Aerts A ; Arnold GJ ; Basu MK ; Bauer DJ ; Caceres CE ; Carmel L ; Casola C ; Choi JH ;Detter JC ; Dong Q ; Dusheyko S ; Eads BD ; Frohlich T ; Geiler-Samerotte KA ; Gerlach D ;Hatcher P ; Jogdeo S ; Krijgsveld J ; Kriventseva EV ; Kultz D ; Laforsch C ; Lindquist E ; Lopez J ;Manak JR ; Muller J ; Pangilinan J ; Patwardhan RP ; Pitluck S ; Pritham EJ ; Rechtsteiner A ;Rho M ; Rogozin IB ; Sakarya O ; Salamov A ; Schaack S ; Shapiro H ; Shiga Y ; Skalitzky C ;Smith Z ; Souvorov A ; Sung W ; Tang Z ; Tsuchiya D ; Tu H ; Vos H ; Wang M ; Wolf YI ;Yamagata H ; Yamada T ; Ye Y ; Shaw JR ; Andrews J ; Crease TJ ; Tang H ; Lucas SM ;Robertson HM ; Bork P ; Koonin EV ; Zdobnov EM ; Grigoriev IV ; Lynch M ; Boore JL The ecoresponsive genome of *Daphnia pulex*. *Science* 2011, 331 (6017), 555–61, 10.1126/science.119776121292972
48. Kao D ; Lai AG ; Stamatakis E ; Rosic S ; Konstantinides N ; Jarvis E ; Di Donfrancesco A ;Pouchkina-Stancheva N ; Semon M ; Grillo M The genome of the crustacean *Parhyale hawaiiensis*, a model for animal development, regeneration, immunity and lignocellulose digestion. *eLife* 2016, 5,e2006227849518
49. Eyun S.-i. ; Young Soh H ; Posavi M ; Munro JB ; Hughes DS ; Murali SC ; Qu J ; Dugan S ;Lee SL ; Chao H Evolutionary history of chemosensory-related gene families across the Arthropoda. *Mol. Biol. Evol* 2017, 34, 10.1093/molbev/msx147
50. Kutsenko A ; Svensson T ; Nystedt B ; Lundeberg J ; Björk P ; Sonnhammer E ; Giacomello S ; Visa N ; Wieslander L The *Chironomus tentans* genome sequence and the organization of the Balbiani ring genes. *BMC Genomics* 2014, 15 (1), 819, 10.1186/1471-2164-15-81925261295
51. Burns FR ; Cogburn AL ; Ankley GT ; Villeneuve DL ; Waits E ; Chang YJ ; Llaca V ;Deschamps SD ; Jackson RE ; Hoke RA Sequencing and de novo draft assemblies of a fathead minnow (*Pimephales promelas*) reference genome. *Environ. Toxicol. Chem* 2016, 35 (1), 212–217, 10.1002/etc.318626513338

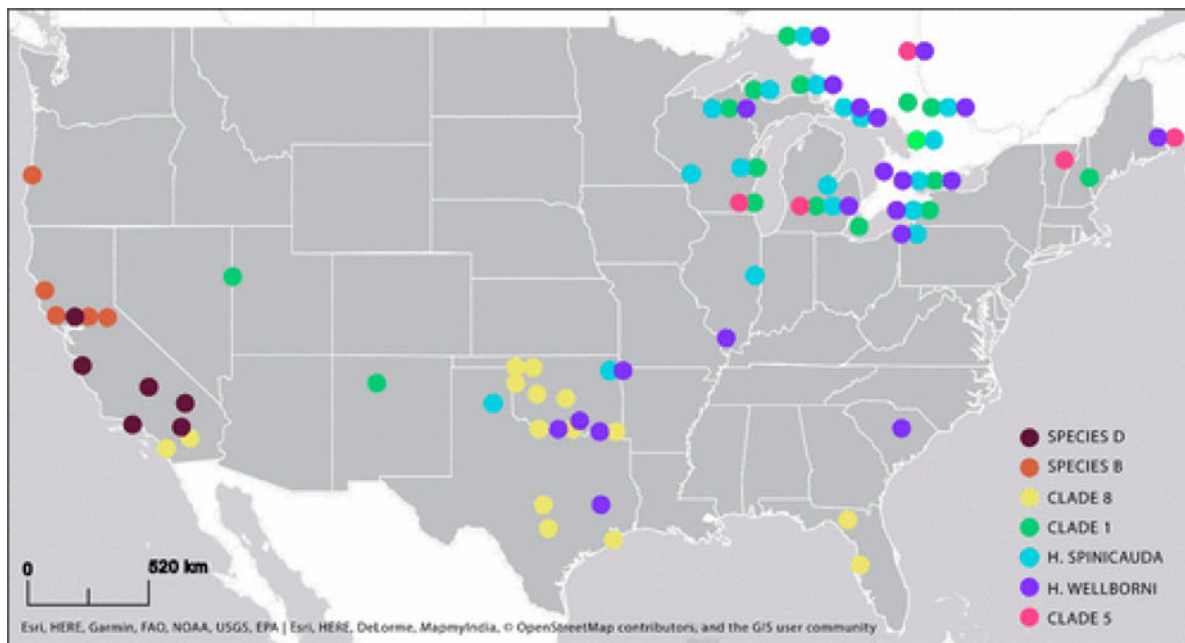
52. Reid NM ; Jackson CE ; Gilbert D ; Minx P ; Montague MJ ; Hampton TH ; Helfrich LW ; King BL ; Nacci DE ; Aluru N The landscape of extreme genomic variation in the highly adaptable Atlantic killifish. *Genome Biol. Evol* 2017, 9 (3), 00–00, 10.1093/gbe/evx023
53. Faddeeva-Vakhrusheva A ; Kraaijeveld K ; Derks MF ; Anvar SY ; Agamennone V ; Suring W ; Kampfraath AA ; Ellers J ; Le Ngoc G ; Gestel CA Coping with living in the soil: the genome of the parthenogenetic springtail *Folsomia candida*. *BMC Genomics* 2017, 18 (1), 493, 10.1186/s12864-017-3852-x28659179
54. Wheeler D ; Redding AJ ; Werren JH Characterization of an ancient lepidopteran lateral gene transfer. *PLoS One* 2013, 8 (3), e59262, 10.1371/journal.pone.005926223533610
55. Alneberg J ; Bjarnason BS ; de Bruijn I ; Schirmer M ; Quick J ; Ijaz UZ ; Lahti L ; Loman NJ ; Andersson AF ; Quince C Binning metagenomic contigs by coverage and composition. *Nat. Methods* 2014, 11 (11), 1144–1146, 10.1038/nmeth.310325218180
56. Yoshida S ; Hiraga K ; Takehana T ; Taniguchi I ; Yamaji H ; Maeda Y ; Toyohara K ; Miyamoto K ; Kimura Y ; Oda K A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* 2016, 351 (6278), 1196–1199, 10.1126/science.aad635926965627
57. Bombelli P ; Howe CJ ; Bertocchini F , Polyethylene bio-degradation by caterpillars of the wax moth *Galleria mellonella*. *Curr. Biol* 2017 27, (8), R292–R293. 10.1016/j.cub.2017.02.06028441558
58. Glastad K ; Hunt B ; Yi S ; Goodisman M DNA methylation in insects: on the brink of the epigenomic era. *Insect Mol. Biol* 2011, 20 (5), 553–565, 10.1111/j.1365-2583.2011.01092.x21699596
59. Hunt BG ; Glastad KM ; Yi SV ; Goodisman MA The Function of Intragenic DNA Methylation: Insights from Insect Epigenomes In *Oxford University Press*, 2013.
60. Glastad KM ; Gokhale K ; Liebig J ; Goodisman MA The caste- and sex-specific DNA methylome of the termite *Zootermopsis nevadensis*. *Sci. Rep* 2016, 6, 37110, 10.1038/srep3711027848993
61. Choudhuri S Small noncoding RNAs: biogenesis, function, and emerging significance in toxicology. *J. Biochem. Mol. Toxicol* 2010, 24 (3), 195–216, 10.1002/jbt.2032520143452
62. Chen S ; McKinney GJ ; Nichols KM ; Colbourne JK ; Sepúlveda MS Novel cadmium responsive microRNAs in *Daphnia pulex*. *Environ. Sci. Technol* 2015, 49 (24), 14605–14613, 10.1021/acs.est.5b0398826550707
63. Chen S ; Nichols KM ; Poynton HC ; Sepúlveda MS MicroRNAs are involved in cadmium tolerance in *Daphnia pulex*. *Aquat. Toxicol* 2016, 175, 241–248, 10.1016/j.aquatox.2016.03.02327078211
64. Carlini DB ; Satish S ; Fong DW Parallel reduction in expression, but no loss of functional constraint, in two opsin paralogs within cave populations of *Gammarus minus* (Crustacea: Amphipoda). *BMC Evol. Biol* 2013, 13 (1), 89, 10.1186/1471-2148-13-8923617561
65. Eyun S.-i. ; Soh HY ; Posavi M ; Munro JB ; Hughes DST ; Murali SC ; Qu J ; Dugan S ; Lee SL ; Chao H ; Dinh H ; Han Y ; Doddapaneni H ; Worley KC ; Muzny DM ; Park E-O ; Silva JC ; Gibbs RA ; Richards S ; Lee CE Evolutionary History of Chemosensory-Related Gene Families across the Arthropoda. *Mol. Biol. Evol* 2017, 34 (8), 1838–1862, 10.1093/molbev/msx14728460028
66. Saina M ; Busengdal H ; Sinigaglia C ; Petrone L ; Oliveri P ; Rentzsch F ; Benton R A cnidarian homologue of an insect gustatory receptor functions in developmental body patterning. *Nat. Commun* 2015, 6:6243, 1–12, 10.1038/ncomms7243
67. Robertson HM The insect chemoreceptor superfamily is ancient in animals. *Chem. Senses* 2015, 40 (9), 609–614, 10.1093/chemse/bjv04626354932
68. Rytz R ; Croset V ; Benton R Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. *Insect Biochem. Mol. Biol* 2013, 43 (9), 888–897, 10.1016/j.ibmb.2013.02.00723459169
69. Peñalva-Arana DC ; Lynch M ; Robertson HM The chemoreceptor genes of the waterflea *Daphnia pulex*: many Grs but no Ors. *BMC Evol. Biol* 2009, 9 (1), 79, 10.1186/1471-2148-9-7919383158
70. Croset V ; Rytz R ; Cummins SF ; Budd A ; Brawand D ; Kaessmann H ; Gibson TJ ; Benton R Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet* 2010, 6 (8), e1001064, 10.1371/journal.pgen.100106420808886

71. Robertson HM Noncanonical GA and GG 5' Intron Donor Splice Sites Are Common in the Copepod *Eurytemora affinis*. *G3: Genes, Genomes, Genet* 2017, 7 (12), 3967–3969, 10.1534/g3.117.300189
72. Croset V ; Schleyer M ; Arguello JR ; Gerber B ; Benton R A molecular and neuronal basis for amino acid sensing in the *Drosophila* larva. *Sci. Rep* 2016, 6, 34871, 10.1038/srep3487127982028
73. Knecht ZA ; Silbering AF ; Cruz J ; Yang L ; Croset V ; Benton R ; Garrity PA Ionotropic Receptor-dependent moist and dry cells control hygro-sensation in *Drosophila*. *eLife* 2017, 6, e26654, 10.7554/eLife.2665428621663
74. Danielson P The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. *Curr. Drug Metab* 2002, 3 (6), 561–597, 10.2174/138920002333705412369887
75. Miller WL ; Chung B.-c . The First Defect in Electron Transfer to Mitochondrial P450 Enzymes In Oxford University Press, 2016.
76. Morale MC ; L'Episcopo F ; Tirolo C ; Giaquinta G ; Caniglia S ; Testa N ; Arcieri P ; Serra P-A ; Lupo G ; Alberghina M Loss of aromatase cytochrome P450 function as a risk factor for Parkinson's disease?. *Brain Res. Rev* 2008, 57 (2), 431–443, 10.1016/j.brainresrev.2007.10.01118063054
77. Sigel A ; Sigel H ; Sigel RK The Ubiquitous Roles of Cytochrome P450 Proteins. Wiley Online Library:2007; Vol. 3.
78. Wahlang B ; Falkner KC ; Cave MC ; Prough RA Role of Cytochrome P450 Monooxygenase in Carcinogen and Chemotherapeutic Drug Metabolism. *Adv. Pharmacol* 2015, 74, 1–33, 10.1016/bs.apha.2015.04.00426233902
79. Han J ; Won E-J ; Hwang D-S ; Shin K-H ; Lee YS ; Leung KM-Y ; Lee S-J ; Lee J-S Crude oil exposure results in oxidative stress-mediated dysfunctional development and reproduction in the copepod *Tigriopus japonicus* and modulates expression of cytochrome P450 (CYP) genes. *Aquat. Toxicol* 2014, 152,308–317, 10.1016/j.aquatox.2014.04.02724813263
80. Han J ; Won E-J ; Kim H-S ; Nelson DR ; Lee S-J ; Park HG ; Lee J-S Identification of the full 46 cytochrome P450 (CYP) complement and modulation of CYP expression in response to water-accommodated fractions of crude oil in the cyclopoid copepod *Paracyclopsina nana*. *Environ. Sci. Technol* 2015, 49 (11), 6982–6992, 10.1021/acs.est.5b0124425942333
81. Feyereisen R Evolution of insect P450. *Biochem. Soc. Trans* 2006, 34 (6), 1252–1255, 10.1042/BST034125217073796
82. Zhao L ; Jones W Expression of heat shock protein genes in insect stress responses. *Invertebrate Survival Journal* 2012, 90, 93–101
83. Richter K ; Haslbeck M ; Buchner J The heat shock response: life on the verge of death. *Mol. Cell* 2010,40 (2), 253–266, 10.1016/j.molcel.2010.10.00620965420
84. Nguyen AD ; Gotelli NJ ; Cahan SH The evolution of heat shock protein sequences, cis-regulatory elements, and expression profiles in the eusocial Hymenoptera. *BMC Evol. Biol* 2016, 16 (1), 15, 10.1186/s12862-015-0573-026787420
85. Bettencourt BR ; Feder ME Hsp70 duplication in the *Drosophila melanogaster* species group: how and when did two become five? *Mol. Biol. Evol* 2001, 18 (7), 1272–1282, 10.1093/oxfordjournals.molbev.a00391211420366
86. Gross TL ; Myles KM ; Adelman ZN Identification and characterization of heat shock 70 genes in *Aedes aegypti* (Diptera: Culicidae). *J. Med. Entomol* 2014, 46 (3), 496–504, 10.1603/033.046.0313
87. Baringou S ; Rouault J-D ; Koken M ; Hardivillier Y ; Hurtado L ; Leignel V Diversity of cytosolic HSP70 Heat Shock Protein from decapods and their phylogenetic placement within Arthropoda. *Gene* 2016, 591(1), 97–107, 10.1016/j.gene.2016.06.06127374152
88. Amiard J-C ; Amiard-Triquet C ; Barka S ; Pellerin J ; Rainbow P Metallothioneins in aquatic invertebrates: their role in metal detoxification and their use as biomarkers. *Aquat. Toxicol* 2006, 76 (2),160–202, 10.1016/j.aquatox.2005.08.01516289342
89. Syring RA ; Brouwer TH ; Brouwer M Cloning and sequencing of cDNAs encoding for a novel copper-specific metallothionein and two cadmium-inducible metallothioneins from the blue crab *Callinectes sapidus*. *Comp. Biochem. Physiol., Part C: Pharmacol., Toxicol. Endocrinol* 2000, 125 (3), 325–332, 10.1016/S0742-8413(99)00114-0

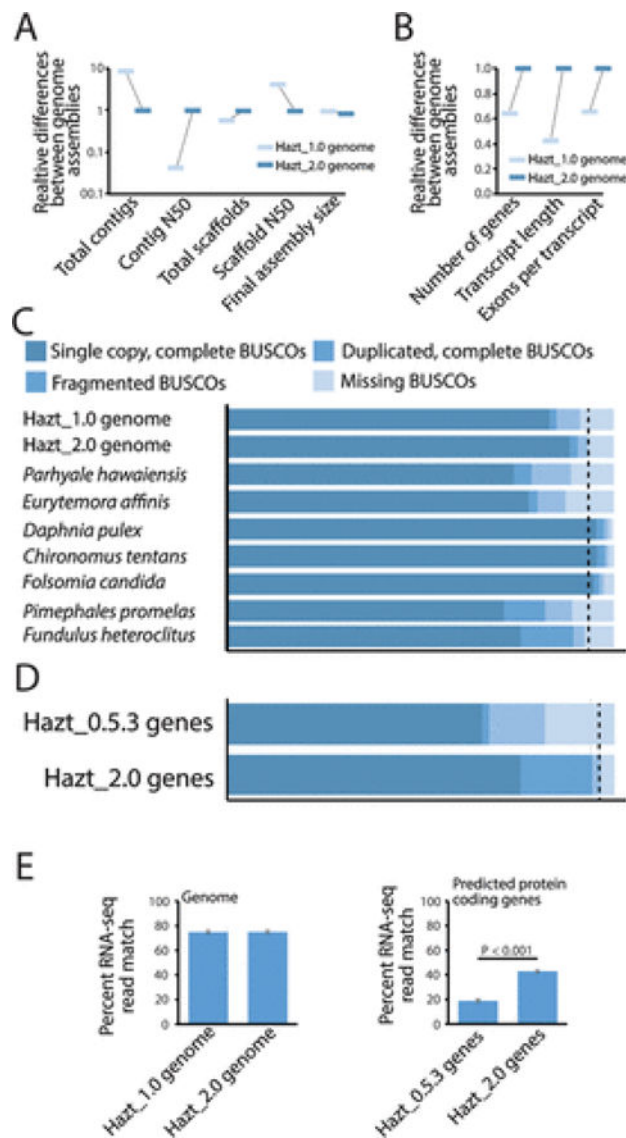
90. Wiczorek H ; Beyenbach KW ; Huss M ; Vitavska O Vacuolar-type proton pumps in insect epithelia. *J. Exp. Biol* 2009, 212 (11), 1611–1619, 10.1242/jeb.03000719448071
91. Lee CE ; Kiergaard M ; Gelembiuk GW ; Eads BD ; Posavi M Pumping ions: rapid parallel evolution of ionic regulation following habitat invasions. *Evolution* 2011, 65 (8), 2229–2244, 10.1111/j.1558-5646.2011.01308.x21790571
92. McNamara JC ; Faria SC Evolution of osmoregulatory patterns and gill ion transport mechanisms in the decapod Crustacea: a review. *J. Comp. Physiol., B* 2012, 182 (8), 997–1014, 10.1007/s00360-012-0665-822534792
93. Lee CE Evolutionary mechanisms of habitat invasions, using the copepod *Eurytemora affinis* as a model system. *Evol. Appl* 2016, 9 (1), 248–270, 10.1111/eva.1233427087851
94. Wiczorek H ; Putzenlechner M ; Zeiske W ; Klein U A vacuolar-type proton pump energizes K<sup>+</sup>/H<sup>+</sup> antiport in an animal plasma membrane. *J. Biol. Chem* 1991, 266 (23), 15340–153471831202
95. Maxson ME ; Grinstein S The vacuolar-type H<sup>+</sup>-ATPase at a glance—more than a proton pump; The Company of Biologists Ltd, 2014.
96. Beyenbach KW ; Wiczorek H The V-type H<sup>+</sup> ATPase: molecular structure and function, physiological roles and regulation. *J. Exp. Biol* 2006, 209 (4), 577–589, 10.1242/jeb.0201416449553
97. Allan AK ; Du J ; Davies SA ; Dow JA Genome-wide survey of V-ATPase genes in *Drosophila* reveals a conserved renal phenotype for lethal alleles. *Physiol. Genomics* 2005, 22 (2), 128–138, 10.1152/physiolgenomics.00233.200415855386
98. Brett CL ; Donowitz M ; Rao R Evolutionary origins of eukaryotic sodium/proton exchangers. *Am. J. Physiol.-Cell Ph* 2005, 288 (2), C223–C239, 10.1152/ajpcell.00360.2004
99. Rheault MR ; Okech BA ; Keen SB ; Miller MM ; Meleshkevitch EA ; Linser PJ ; Boudko DY ; Harvey WR Molecular cloning, phylogeny and localization of AgNHA1: the first Na<sup>+</sup>/H<sup>+</sup> antiporter (NHA) from a metazoan. *J. Exp. Biol* 2007, 210 (21), 3848–3861, 10.1242/jeb.00787217951426
100. Donowitz M ; Tse CM ; Fuster D SLC9/NHE gene family, a plasma membrane and organellar family of Na<sup>+</sup>/H<sup>+</sup> exchangers. *Mol. Aspects Med* 2013, 34 (2), 236–251, 10.1016/j.mam.2012.05.00123506868
101. Harvey WR Voltage coupling of primary H<sup>+</sup> V-ATPases to secondary Na<sup>+</sup>-or K<sup>+</sup>-dependent transporters. *J. Exp. Biol* 2009, 212 (11), 1620–1629, 10.1242/jeb.03153419448072
102. Xiang MA ; Linser PJ ; Price DA ; Harvey WR Localization of two Na<sup>+</sup>-or K<sup>+</sup>-H<sup>+</sup> antiporters, AgNHA1 and AgNHA2, in *Anopheles gambiae* larval Malpighian tubules and the functional expression of AgNHA2 in yeast. *J. Insect Physiol* 2012, 58 (4), 570–579, 10.1016/j.jinsphys.2011.12.00922206887
103. Kondapalli KC ; Kallay LM ; Muszelik M ; Rao R Unconventional chemiosmotic coupling of NHA2, a mammalian Na<sup>+</sup>/H<sup>+</sup> antiporter, to a plasma membrane H<sup>+</sup> gradient. *J. Biol. Chem* 2012, 287 (43), 36239–36250, 10.1074/jbc.M112.40355022948142
104. Battaglini RA ; Pham L ; Morse LR ; Vokes M ; Sharma A ; Odgren PR ; Yang M ; Sasaki H ; Stashenko P NHA-oc/NHA2: a mitochondrial cation–proton antiporter selectively expressed in osteoclasts. *Bone* 2008, 42 (1), 180–192, 10.1016/j.bone.2007.09.04617988971
105. Schubauer-Berigan MK ; Dierkes JR ; Monson PD ; Ankley GT pH-Dependent toxicity of Cd, Cu, Ni, Pb and Zn to *Ceriodaphnia dubia*, *Pimephales promelas*, *Hyalella azteca* and *Lumbriculus variegatus*. *Environ. Toxicol. Chem* 1993, 12 (7), 1261–1266, 10.1897/1552-8618(1993)12[1261:PTOCCN]2.0.CO;2
106. Jackson B ; Lasier P ; Miller W ; Winger P Effects of calcium, magnesium, and sodium on alleviating cadmium toxicity to *Hyalella azteca*. *Bull. Environ. Contam. Toxicol* 2000, 64 (2), 279–286, 10.1007/s00128991004110656896
107. Wang Z ; Meador JP ; Leung KM Metal toxicity to freshwater organisms as a function of pH: A meta-analysis. *Chemosphere* 2016, 144, 1544–1552, 10.1016/j.chemosphere.2015.10.03226498103

108. de Schampelaere KA ; Janssen CR A biotic ligand model predicting acute copper toxicity for *Daphnia magna*: the effects of calcium, magnesium, sodium, potassium, and pH. *Environ. Sci. Technol* 2002, 36 (1),48–54, 10.1021/es000253s11817370
109. Chintapalli VR ; Kato A ; Henderson L ; Hirata T ; Woods DJ ; Overend G ; Davies SA ;Romero MF ; Dow JA Transport proteins NHA1 and NHA2 are essential for survival, but have distinct transport modalities. *Proc. Natl. Acad. Sci. U. S. A* 2015, 112 (37), 11720–11725, 10.1073/pnas.150803111226324901
110. Foulkes EC Transport of Toxic Heavy Metals Across Cell Membranes (44486). *Proc. Soc. Exp. Biol. Med* 2000, 223 (3), 234–240, 10.1046/j.1525-1373.2000.22334.x10719835
111. Argüello JM ; Raimunda D ; González-Guerrero M Metal Transport across Biomembranes: Emerging Models for a Distinct Chemistry. *J. Biol. Chem* 2012, 287 (17), 13510–13517, 10.1074/jbc.R111.31934322389499
112. Liber K ; Doig LE ; White-Sobey SL Toxicity of uranium, molybdenum, nickel, and arsenic to *Hyalella azteca* and *Chironomus dilutus* in water-only and spiked-sediment toxicity tests. *Ecotoxicol. Environ. Saf*2011, 74 (5), 1171–1179, 10.1016/j.ecoenv.2011.02.01421529943
113. Nguyen LTH ; Muysen BTA ; Janssen CR Single versus combined exposure of *Hyalella azteca* to zinc contaminated sediment and food. *Chemosphere* 2012, 87 (1), 84–90, 10.1016/j.chemosphere.2011.11.06622197310
114. Ball AL ; Borgmann U ; Dixon DG Toxicity of a cadmium-contaminated diet to *Hyalella azteca*. *Environ. Toxicol. Chem* 2006, 25 (9), 2526–2532, 10.1897/05-650R.116986809
115. Wallis LK ; Diamond SA ; Ma H ; Hoff DJ ; Al-Abed SR ; Li S Chronic TiO<sub>2</sub> nanoparticle exposure to a benthic organism, *Hyalella azteca*: impact of solar UV radiation and material surface coatings on toxicity. *Sci. Total Environ* 2014, 499 (Supplement C), 356–362, 10.1016/j.scitotenv.2014.08.068
116. Hasenbein S ; Connon RE ; Lawler SP ; Geist J A comparison of the sublethal and lethal toxicity of four pesticides in *Hyalella azteca* and *Chironomus dilutus*. *Environ. Sci. Pollut. Res* 2015, 22 (15), 11327–11339, 10.1007/s11356-015-4374-1
117. Hasenbein S ; Lawler SP ; Geist J ; Connon RE A long-term assessment of pesticide mixture effects on aquatic invertebrate communities. *Environ. Toxicol. Chem* 2016, 35 (1), 218–232, 10.1002/etc.318726565581
118. Ingersoll CG ; Brunson EL ; Dwyer FJ ; Hardesty DK ; Kemble NE Use of sublethal endpoints in sediment toxicity tests with the amphipod *Hyalella azteca*. *Environ. Toxicol. Chem* 1998, 17 (8), 1508–1523, 10.1897/1551-5028(1998)017<1508:UOSEIS>2.3.CO;2
119. Hook SE ; Twine NA ; Simpson SL ; Spadaro DA ; Moncuquet P ; Wilkins MR 454 pyrosequencing-based analysis of gene expression profiles in the amphipod *Melita plumulosa*: Transcriptome assembly and toxicant induced changes. *Aquat. Toxicol* 2014, 153 (Sp. Iss. SI), 73–88,10.1016/j.aquatox.2013.11.022
120. Werner I ; Nagel R Stress proteins HSP60 and HSP70 in three species of amphipods exposed to cadmium, diazinon, dieldrin and fluoranthene. *Environ. Toxicol. Chem* 1997, 16 (11), 2393–2403, 10.1002/etc.5620161127
121. Liang P ; MacRae TH Molecular chaperones and the cytoskeleton. *J. Cell Science* 1997, 110 (13), 1431–14409224761
122. Watanabe M ; Suzuki T Cadmium-induced synthesis of HSP70 and a role of glutathione in *Euglena gracilis*. *Redox Rep* 2004, 9 (6), 349–353, 10.1179/13510000422500686715720831
123. Poynton HC ; Varshavsky JR ; Chang B ; Cavigliolo G ; Chan S ; Holman PS ; Loguinov AV ;Bauer DJ ; Komachi K ; Theil EC ; Perkins EJ ; Hughes O ; Vulpe CD *Daphnia magna* Ecotoxicogenomics Provides Mechanistic Insights into Metal Toxicity. *Environ. Sci. Technol* 2007, 41 (3),1044–1050, 10.1021/es061557317328222
124. Poynton HC ; Loguinov AV ; Varshavsky JR ; Chan S ; Perkins EJ ; Vulpe CD Gene expression profiling in *Daphnia magna* part I: concentration-dependent profiles provide support for the no observed transcriptional effect level. *Environ. Sci. Technol* 2008, 42 (16), 6250–6256, 10.1021/es801078318767695
125. Spurlock F ; Lee M Synthetic Pyrethroid Use Patterns, Properties, and Environmental Effects. In *Synthetic Pyrethroids*. ACS Symposium Series 2008, 991, 3–25, 10.1021/bk-2008-0991.ch001

126. Stehle S ; Schulz R Agricultural insecticides threaten surface waters at the global scale. *Proc. Natl. Acad. Sci. U. S. A* 2015, 112 (18), 5750–5755, 10.1073/pnas.150023211225870271
127. Nebert DW Aryl hydrocarbon receptor (AHR): "pioneer member" of the basic-helix/loop/helix per-Arnt-sim (bHLH/PAS) family of "sensors" of foreign and endogenous signals. *Prog. Lipid Res* 2017, 67, 38–57, 10.1016/j.plipres.2017.06.00128606467
128. Kim B-M ; Rhee J-S ; Hwang U-K ; Seo JS ; Shin K-H ; Lee J-S Dose- and time-dependent expression of aryl hydrocarbon receptor (AhR) and aryl hydrocarbon receptor nuclear translocator (ARNT) in PCB-, B[a]P-, and TBT-exposed intertidal copepod *Tigriopus japonicus*. *Chemosphere* 2015, 120, 398–406, 10.1016/j.chemosphere.2014.07.09925216468
129. Lee HG ; Yang JH PCB126 induces apoptosis of chondrocytes via ROS-dependent pathways. *Osteoarthr. Cartilage* 2012, 20 (10), 1179–1185, 10.1016/j.joca.2012.06.004
130. Wayne A ; Trudeau VL Neuroendocrine Disruption: More than Hormones are Upset. *J. Toxicol. Environ. Health, Part B* 2011, 14 (5–7), 270–291, 10.1080/10937404.2011.578273
131. Soucek DJ ; Lazo-Wasem EA ; Taylor CA ; Major KM Description of Two New Species of *Hyaella* (Amphipoda: Hyaellidae) from Eastern North America with a Revised Key to North American Members of the Genus. *J. Crustacean Biol* 2015, 35 (6), 814–829, 10.1163/1937240X-00002373
132. Dionne K ; Vergilino R ; Dufresne F ; Charles F ; Nozais C No Evidence for Temporal Variation in a Cryptic Species Community of Freshwater Amphipods of the *Hyaella azteca* Species Complex. *Diversity* 2011, 3 (3), 390, 10.3390/d3030390
133. Wellborn GA ; Broughton RE Diversification on an ecologically constrained adaptive landscape. *Mol. Ecol* 2008, 17 (12), 2927–2936, 10.1111/j.1365-294X.2008.03805.x18522695
134. Galbreath JGS ; Smith JE ; Becnel JJ ; Butlin RK ; Dunn AM Reduction in post-invasion genetic diversity in *Crangonyx pseudogracilis* (Amphipoda: Crustacea): a genetic bottleneck or the work of hitchhiking vertically transmitted microparasites?. *Biol. Invasions* 2010, 12 (1), 191–209, 10.1007/s10530-009-9442-3
135. Baird DJ ; Pascoe TJ ; Zhou X ; Hajibabaei M Building freshwater macroinvertebrate DNA-barcode libraries from reference collection material: formalin preservation vs specimen age. *J. N. Am. Benthol. Soc* 2011, 30 (1), 125–130, 10.1899/10-013.1
136. Witt JD ; Threlloff DL ; Hebert PD DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. *Mol. Ecol* 2006, 15 (10), 3073–3082, 10.1111/j.1365-294X.2006.02999.x16911222
137. Rosendale AJ ; Romick-Rosendale LE ; Watanabe M ; Dunlevy ME ; Benoit JB Mechanistic underpinnings of dehydration stress in the American dog tick revealed through RNA-Seq and metabolomics. *J. Exp. Biol* 2016, 219 (12), 1808–1819, 10.1242/jeb.13731527307540
138. Schoville SD ; Chen YH ; Andersson MN ; Benoit JB ; Bhandari A ; Bowsher JH ; Brevik K ; Cappelle K ; Chen M-JM ; Childers AK A model species for agricultural pest genomics: the genome of the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). *bioRxiv* 2017, 192641



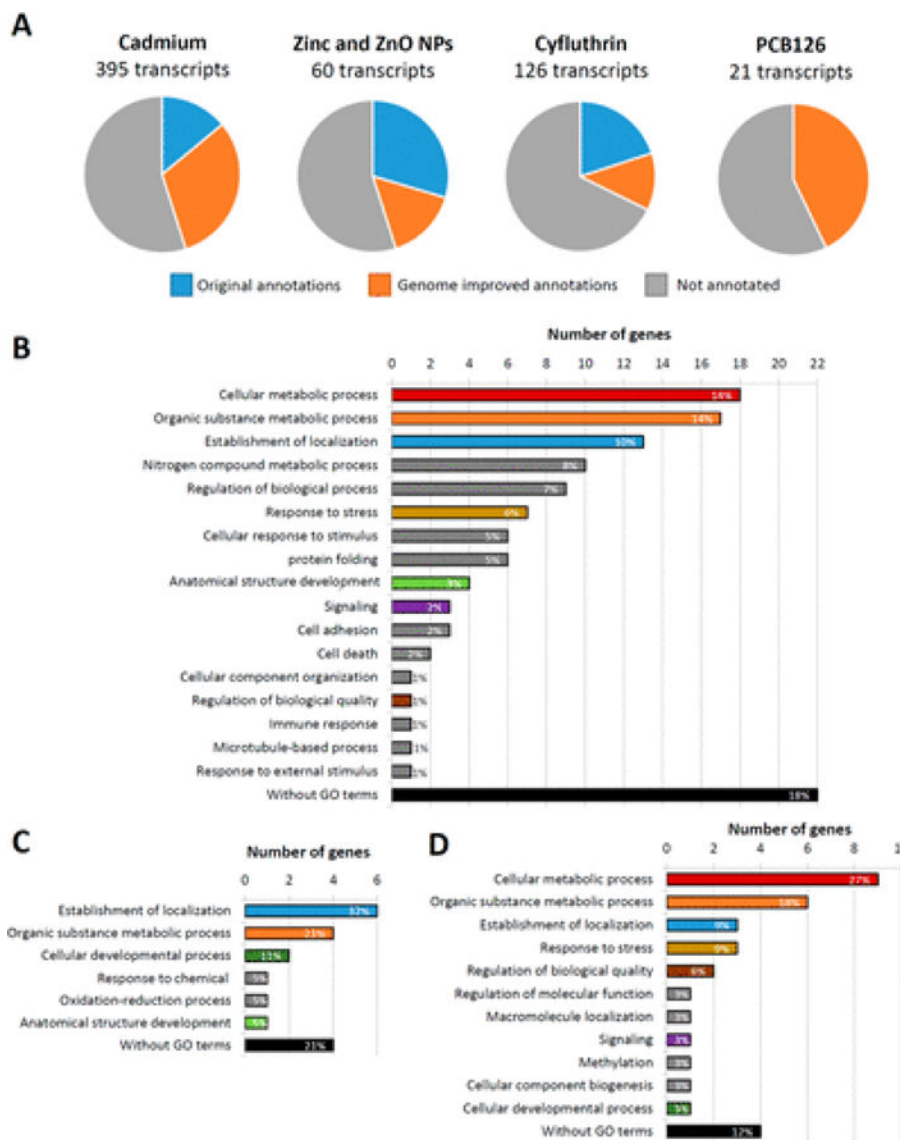
**Figure 1.** Geographical distribution of seven of the most well characterized *Hyalella azteca* species groups in the United States and Canada. The two species that have been taxonomically described include *Hyalella spinicauda* and *Hyalella wellborni*.(131) The distribution shown here was described in several publications for *H. wellborni*(19,29,132) with additional collections by R. Cothran and G. Wellborn; and *H. spinicauda*.(19,29,133) The remaining five species have species level divergence in the cytochrome oxidase-I gene, but without taxonomic descriptions, they are named with the designations applied in publications describing the collections. These include Clade 8 (also referred to as U.S. Lab strain,(29) Species C,(22) OK-L(133)) with additional collections shown here by G. Wellborn, R. Cothran, M. Worsham, A. Kuzmic; Clade 1;(19,29,132,134,135) Clade 5;(19)Species B; (22,30,133) and Species D.(22,30,136) The genome sequence described here represents Clade 8 or the U.S. Lab strain.(29) The other commonly used laboratory strain, primarily from Canada, belongs to Clade 1.(29)



**Figure 2.** Summary of the quality and completeness of the two *H. azteca* genome assemblies. (A) Comparison of the contig, scaffold, and assembly size between the two assemblies. (B) Comparison of predicted gene sets from the two assemblies. The original MAKER gene set for Hazt\_1.0 is Hazt\_0.5.3 while the gene set developed by NCBI for Hazt\_2.0 is referred to as Hazt\_2.0. (C) BUSCO analysis compared to other genomes of ecotoxicological relevance including the amphipod *Parhyale hawaiensis*,(48) the copepod *Eurytemora affinis*,(49) *Daphnia pulex*,(47) the aquatic midge *Chironomus tentans*,(50) the terrestrial springtail *Folsomia candida*,(53) the fathead minnow *Pimephales promelas*,(51) and the killifish *Fundulus heteroclitus*.(52) Dotted line corresponds to the total number of BUSCOs (single copy, duplicated, or fragmented) for the Hazt\_2.0 assembly. (D) BUSCO comparison for the predicted gene sets. Dotted line corresponds to the total number of BUSCOs (single copy, duplicated, or fragmented) for the Hazt\_2.0 gene set. (E) Percentage of RNaseq reads that mapped to the genome (left) and the predicted protein coding genes (right). Illumina data



sets were acquitted from the NCBI Bioproject (PRJNA312414). Reads were mapped according to methods described in Rosendale et al.(137) and Schoville et al.(138)



**Figure 3.** Annotation of toxicant responsive genes. (A) The annotation of differentially expressed transcripts was significantly improved using the *H. azteca* genome. The total number of unique differentially expressed transcripts is listed below each treatment. Pie graphs represent differentially expressed contigs and illustrate the percentage of original annotations (blue) and the additional annotations that were added when contigs were aligned to the genome (orange). For PCB126, none of the contigs were annotated prior to alignment to the genome. In many cases more than one contig aligned to the same transcript; therefore, the total number of contigs is greater than the number of transcripts. (B–D) Biological processes gene ontology (GO) terms representing the differentially expressed transcripts from Cd (B), Zn, and ZnO NPs (C), and cyfluthrin (D). The number of genes mapped to each of the GO terms is shown by the length of the bars, while the percentage of total transcripts is marked at the end of each bar. For PCB126, none of the 12 annotated

transcripts were mapped to biological processes GO terms. Similar graphs for molecular function GO terms can be found in the SI Figure S5.

**Table 1.**

Model Toxicants and Exposure Concentrations for Microarray Gene Expression Analysis

<b>pollutant class</b>	<b>chemical</b>	<b>end points</b>	<b>concentrations</b>	<b>references</b>
heavy metals	cadmium	LC <sub>10</sub>	5.5 µg/L	this study
	zinc	1/10 LC <sub>50</sub>	25 µg/L	Poynton et al. (2013)(44)
		LC <sub>25</sub>	104 µg/L	
insecticide	cyfluthrin	NOEC	1 ng/L	Weston et al. (2013)(30)
nanomaterial	ZnO NP	1/10 LC <sub>50</sub>	18 µg/L	Poynton et al. (2013)(44)
		LC <sub>25</sub>	65 µg/L	
organic pollutant	PCB126		7.0 µg/L	this study