Multiple origins of pyrethroid insecticide resistance across the species complex of a non-target aquatic crustacean, *Hyalella azteca*

Donald P. Weston, Helen C. Poynton, Gary A. Wellborn, Michael J. Lydy, Bonnie J. Blalock, Maria S. Sepulveda, John K. Colbourne

Supporting Information (SI)

SI Materials and Methods

Collection of Hyalella azteca

Wild populations of *H. azteca* were sampled by sweeping a D-net through aquatic vegetation. Collections were made between June and November, 2010, except for Morrison Creek collected June 2011, and a second collection from Pleasant Grove Creek in March, 2013. Surficial sediments (approximately 0-2 cm) were obtained from each site concurrently with the *H. azteca* collections and frozen for pyrethroid analysis.

Toxicity testing

Individuals used for toxicity testing were those that passed through a 600 µm screen, but were retained on a 500 µm screen (except Pleasant Grove Creek in March 2013 when animals passing through a 2000 µm screen but retained on a 1000 µm were used, since smaller individuals were absent from the population in winter). Retention on the 500 µm screen provides individuals 1.8-2.5 mm in length (excluding antennae), which generally corresponds to an age of 7-14 d (1). This assumption is probably reasonable for most populations, but Blodgett Reservoir specimens were particularly small as adults, and thus animals used for testing from this population were of comparable size as all other populations, but may have been of an older age.

Testing was done in lab water, purified through a Milli-Q system (Millipore, Billerica, MA), and then made moderately hard by addition of salts (1). Pyrethroids were added to the water in an acetone carrier, with acetone concentrations kept below 30 µl/L and verified to have no toxic effect in solvent control tests. All populations were tested with cyfluthrin, and many were also tested with bifenthrin when sufficient animals were available. Exposures were done in glass beakers containing 80 ml of water, and a 1 cm^2 square of nylon screen on which the animals cling. Tests were done as a concentration series with five to ten concentrations at 2x steps (e.g., 1, 2, 4, 8, 16 ng/L), and three replicate beakers per step. Beakers were held at 23°C under a 16 h light:8 h dark cycle. After 48 h, 1 ml of yeast-cerophyll-trout food was added to each beaker, and after allowing approximately 6 h for feeding, 80% of the water was removed and replaced with fresh solution. At 96 h, the test was terminated.

All populations were tested within two days of collection, but animals from Grayson and Chualar Creeks were also held in the laboratory for extended periods in pyrethroid-free cultures. Juveniles were collected and tested for cyfluthrin sensitivity after 25 d in

culture (both creeks) and again after 96 d (Chualar only). *H. azteca* grows out of the size class used for testing by 15 d of age (1), a finding confirmed in our own preliminary tests using Mosher Slough animals. Therefore the 25 d tests would be representative of F1 individuals having no prior pyrethroid exposure except through maternal transfer. The 96 d tests could have been either or both F1 and F2 animals.

Toxicity testing statistics were calculated using CETIS (Tidepool Scientific Software, McKinleyville, CA). LC_{50} values were determined by probit analysis.

Chemical analyses

In all toxicity tests, a 400 ml water sample was collected at test initiation from a concentration step in the middle of the range being tested, which typically would be near the LC_{50}. This sample was combined with another 400 ml at the same concentration step from the fresh solution used for water change at 48 h. This composite sample provided the best estimate of initial pyrethroid concentrations, and these results were used to report LC_{50} data on the basis of actual rather than nominal concentrations.

Water samples were preserved by addition of 10 ml of hexane. After adding the 48 h sample to the initial sample, the composite was extracted within 24 h. The surrogates 4,4'-dibromooctafluorobiphenyl (DBOFB) and decachlorobiphenyl (DCBP) were added, and the sample liquid:liquid extracted using U.S. Environmental Protection Agency (EPA) Method 3510C. Three sequential extractions were done with 60 ml dichloromethane (DCM), with one 60 ml aliquot also used to extract the sample bottle. The combined extracts were concentrated to 1 ml in hexane and analyzed following Wang et al. (2). Briefly, the extract was added to a dual layer graphitized black carbon and primary/secondary amine column preconditioned with hexane (Supelclean ENVI™-Carb II/Supelclean™ primary/secondary amine column, 300 mg /600 mg, 6.0 ml; Sigma-Aldrich/Supelco, Bellefonte, PA). The column was eluted with 7.0 ml of 30% dichloromethane in hexane, the eluate reduced to near dryness, and then concentrated to 1.0 ml in 0.1% acetic acid in hexane. The acidification step was used to avoid isomerization of the pyrethroids (3).

Extracts were analyzed on an Agilent 6890 gas chromatograph with a micro-electron capture detector (Agilent Technologies, Palo Alto, CA). Two columns, a RTX 1614 and a DB-608, were used. Calibration was performed using the external standard method. Quality control measures included blanks, lab control spikes, matrix spikes, matrix spike duplicates, and field duplicates, all done with every batch of 20 samples.

Sediment samples were processed following methods detailed in You et al (4). Before extraction, frozen sediment was freeze dried (Labconco Corporation, Kansas City, MO) at -80°C for 24h. Approximately 5 g of dry sediment were mixed with 1 g of silica and 2 g of copper powder, and surrogate standards (DBOFB and DCBP) added. Samples were extracted using a matrix-dispersive accelerated solvent extraction method using a Dionex 200 (Dionex, Sunnyvale, CA) with 33 ml stainless steel cells with 1:1 DCM: acetone (v/v) at 100ºC and 1500 pound-force per square inch for two 5 min static cycles. The extract was collected in 60 ml glass collection vials. Clean-up and analysis of the sediment extract then followed the procedure described for water samples.

Phylogenetic sequencing

Genetic diversity of study populations and laboratory cultures was assessed by evaluating nucleotide sequence variation in mitochondrial and nuclear genes. Specimens were preserved in 100% ethanol, and DNA extracted using chelating resin as described in Wellborn and Broughton (5). A fragment of the mitochondrial gene cytochrome C oxidase subunit I (COI) was PCR-amplified using the primer TrpPar1 (5'—GTTATATAAAC-TATTAGCCTTCCAA—3') paired with either COIaV9 (5'—ACTGCCACAACAGAYAARTAMGA-CCC—3') or COIaV10 (5'—ACAGCAACAACAGATAARTARGACC—3'). The 28S nuclear ribosomal large subunit rRNA gene (*28S* rDNA) was amplified using primers 28S-3311F and 28S-4434R (6). PCR was carried out in 50-μl reactions containing 2.0 μl template DNA, 5 μl 10x PCR buffer, 200 μM of each dNTP, 0.25 μM of each primer, and 1.25 units DNA polymerase. Cycling conditions were 5 min at 94°C, followed by 35 cycles of 45 s at 94°C, 45 s at 51°C, and 60 s at 72°C, followed by 5 min at 72°C. PCR products were gel purified, and sequenced with TrpPar1 (COI) or 28S-3311F (28S) using an ABI 3730 automated sequencer. Sequences were aligned using the Geneious alignment tool implemented in Geneious (7), and evaluation of relationships among aligned sequences employed maximum likelihood (COI) and neighbor-joining (28S) analyses implemented in PhyML (8). For COI, genetic distances among and within groups were quantified as the Kimura 2-parameter distance (9), and calculated in MEGA (10). GenBank accession numbers are KF596727-KF596752 (COI) and KF596753-KF596771 (28S).

H. azteca transcriptome sequencing

H. azteca chemical exposures and later cDNA library and microarray construction were conducted at Purdue University and Indiana University. In order to capture a wide range of transcriptome responses, we chose to produce libraries and sequences of genes from a combination of adult females and males (35 days of age) and juveniles (3-day old). The animals used were from a culture maintained at Aquatic Research Organisms (Hampton, NH). The animals were exposed to the following environmentally relevant chemicals: 0.5 ppb cadmium chloride; 20 ppb atrazine; and 0.5 ppb PCB Aroclor 1254. These are sublethal concentrations below EC_{50} values known to impact reproduction in H. azteca (11, 12, 13). Animals were exposed for 48 h in glass Petri dishes, one animal per dish with 12 replicates. Each dish was filled with 45 mL of test solution (60:40 dechlorinated tap:ultrapure water). Animals were flash frozen in liquid nitrogen and immediately stored in -80ºC freezer until RNA extraction. Detailed protocols for RNA extraction, cDNA synthesis, as well as 454 transcriptomics sequencing, assembly and sequencing analysis can be found at Meyer et al. (14). Sequencing results including total number of reads and annotated sequences can be found in **Table S6**. Assembled reads including 65,961 contigs and 147,877 singletons were uploaded to GenBank are available as accession numbers GAJQ00000000.1 and GAJP00000000.1.

Voltage-gated sodium channel (vgsc) sequencing

Seven sequences within our transcriptome assembly were identified as partial sequences of the voltage-gated sodium channel through a tblastx homology search to the *Musca domestica* X96668 against all contigs and singletons (Sequence IDs: 317673_2167_1494, contig 22062, seq5340434, contig 09991, 493070_3780_3186,

contig22046, and contig 11860). To clone and sequence target sites of *vgsc* for possible mutations associated with pyrethroid resistance, a combination of long distance PCR to close gaps between short sequences homologous to *M. domestica* VGSC, and degenerate PCR to cover area devoid of homologous sequences was used with gDNA (extracted using Qiagen DNeasy Blood and Tissue kit) and/or cDNA (see below) extracted from the UCB population.

Domain I S6 resides upstream of contig22062. Degenerate primers were designed based on a highly conserved region of the *vsgc* identified by comparing several insect sequences to the *Daphnia pulex* sequence. A degenerate primer: TTYGAYWSNTTYGGNT-GGGC was designed and used the reverse primer: CCATGTGAATTGGTGAGAACC to amplify cDNA using TaqOne DNA polymerase (New England Biolabs), which amplified an approximately 600 bp product.

Domain II resides between contig22062, seq5340434, and contig09991 therefore, long-distance PCR was used. Genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue kit. Long distance PCR (Advantage Genomic LA Polymerase Mix, Clonetech, Mountain View, CA) was used to fill in sequences between: (1) seq5340434 and contig 09991 using forward primer: GTGGGACTGCATGTACGTGGGAGACTT and reverse primer: TAACTTCTCGGATCGGCTGGTTGTTC, and (2) contig22062 and seq5340434 using forward primer: GGAGACGAATCCTTTCATCG and reverse primer: CCGTCTTGAGACCATTTGAA. The resulting partial sequence for *H. azteca* sodium channel that includes all of domain II was deposited in GenBank (JX678289).

Domain III S6 resides between contig 09991 and 493070_3780_3186. Primers were designed and PCR on cDNA was performed using the following primers: Forward: ATACCAACGGGCTGAACAAC and Reverse: TGTCGAACTTTTTGCTCGTG.

PCR products obtained for Domain I S6, Domain II, and Domain III S6 were cloned into the pCR8/GW/TOPO plasmid using TA cloning (Invitrogen), transformed into One Shot Mach1-T1R Chemically Competent *E. coli* (Invitrogen), and appropriate product sizes were determined by restriction digest using ECO RI (New England Biolabs). Plasmids containing the appropriate size PCR product were isolated using Qiagen mini-prep plasmid isolation kit and sequenced at the Massachusetts General Hospital DNA core facility with GW1 and GW2 sequencing primers (Invitrogen).

Analyses of potential mutations in the *vgsc* were conducted using specimens from all the collected populations and the UCB lab culture. Three individuals from the UCB lab culture or from each of the natural populations were selected from animals preserved in RNAlater (Ambion, Grand Island, NY). gDNA was isolated from single individuals using Qiagen DNeasy Blood and Tissue kit. To determine if any of the H. azteca populations contained mutations at in the *vgsc*, which could be responsible for target site resistance, four different segments of the *vgsc* were sequenced (see Figure S4). PCR was performed using primers provided in Table S3 and up to 50 ng of gDNA from each individual. PCR products were then cloned into the pCR8/GW/TOPO plasmid using TA cloning (Invitrogen), transformed into One Shot Mach1-T1R Chemically Competent *E. coli* (Invitrogen), and appropriate product sizes were screened by restriction digest using ECO RI (New England Biolabs) or colony PCR using the appropriate primers. Plasmids containing the desired PCR product were isolated using Qiagen mini-prep plasmid isolation kit and sequenced at the Massachusetts General Hospital DNA core facility with GW1 and GW2 sequencing primers (Invitrogen).

For *vgsc* genotyping of the Pleasant Grove Creek population, a total of 56 individuals collected in 2013, including 16 that survived the 128 ng/L cyfluthrin treatment, and 23 individuals collected in 2010 were genotyped. gDNA was isolated by placing all or part of single individuals into microtubes containing a 5% solution of chelating resin (Chelex 100, Sigma) and heating tubes to 60 °C for 20 min, and then 100 °C for 20 min. Primers were designed which amplified a 187 bp product containing the M918 and L925 sites (Forward: AGGGTGTTCAAGCTCGCTAA and Reverse: AATTCTTGCCGAAGAGTTGC). PCR reactions were carried out using OneTaq DNA polymerase (New England Biolabs) and the following PCR program: 95°C for 5 min.; 35 cycles of 95°C for 30 s, 58°C for 30s, 68°C for 45 s; and a final extension at 68°C for 15 min. PCR products were purified using Qiaquick PCR clean-up kit (Qiagen) and sequenced at the Massachusetts General Hospital DNA core facility.

Transcriptomics

A 135k oligonucleotide microarray was designed using the H. azteca 454 sequence data on a Nimblegen platform (12x135k) (Roche Nimblegen). Analyses were conducted using specimens from the UCB and Grayson Creek populations because they showed a wide variation in pyrethroid sensitivity. Within 48 h of collection, both populations were used for determination of pyrethroid 96-h LC_{50}s, and at the completion of these tests, control animals were preserved from both populations. In addition, animals from both populations were collected from the 0.4 ng/L cyfluthrin exposure (the No-Observed-Effects-Concentration (NOEC) for the UCB population) and Grayson Creek animals were collected from the NOEC for that population (170 ng/L). All animals were preserved immediately at the conclusion of the toxicity tests in RNAlater and frozen until further analysis.

At the University of Massachusetts Boston, animals were removed from the tubes and washed with Trireagent (Molecular Research Center, Cincinnati, OH) to remove residual RNAlater. Following washing, Trireagent was added to each tube and samples were homogenized using the Tissue Lyser II, bead mill (Qiagen). RNA was extracted according to standard methods and DNAse I treated using Qiagen RNAeasy on-column digestion. RNA was quantified using a Nanodrop Spectrophotometer (Thermo Scientific) and RNA degradation was assessed on a 2% glyoxal-agarose gel (Ambion). All samples were of high quality going into the microarray. RNA samples were reverse transcribed using Sigma TransPlex Whole Transcriptome Amplification (WTA) Kit, which produces amplified cDNA through two steps. First strand cDNA is produced by reverse transcription and cDNA is then amplified by PCR. Amplified cDNA was labeled using Nimblegen's one-color DNA labeling kit (cy3 only). This kit is based on Cy3-random priming and uses a high concentration of exo- Klenow, to incorporate dNTPs while replicating the cDNA. Labeled cDNA was hybridized to the 135k microarray using a Nimblegen hybridization station and a hybridization temperature of 42°C. Microarrays were hybridized for 17 h and washed using Nimblegen wash buffers. Labeling and hybridization proceeded according to the protocols of NimbleGen Arrays User's Guide: Gene Expression Arrays (Version 6.0).

Washed and dried microarrays were scanned at the University of Massachusetts Medical School Genomics Facility within two days of completing the hybridization. They were scanned on an Agilent DNA microarray scanner C with SureScan High-Resolution Technology at 3 µm resolution. Images were processed using Nimblescan software (Roche) that produced "pair" files with raw intensity values for each spot. Raw intensity values were normalized across the set of twenty different microarrays using quantile

normalization with median polish in Arraystar (DNASTAR). Normalized Log2 intensity values were uploaded into MultiExperiment Viewer (Mev 4.5) (part of TM4 microarray software suite; www.tm4.org) and Statistical Analysis for Microarrays (SAM) was used to detect differentially expressed genes between the exposed samples and unexposed samples using a FDR cut-off of 15% (15). Details related to the microarray platform and raw intensity values of the microarray experiments can be found at the Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE48943.

All sequences identified as differentially expressed in the control vs. exposed comparisons (both UCB and GC) were uploaded to Blast2Go. Annotation and gene ontology terms (GO terms) were assigned using Blast2Go (www.blast2go.org) (16). Overrepresented gene functions were identified using enrichment analysis with Blast2Go (17). Comparisons were performed between genes differentially expressed by the pyrethroid exposure in the UCB or GC populations and their respective unexposed controls.

SI Additional Results

Inherited resistance to pyrethroids

While most toxicity testing done in connection with the present study tested juveniles obtained from the various sites within two days of their collection, limited experiments were done to determine if resistance persisted in progeny in the absence of selective pressure from pyrethroids. Animals collected from two highly resistant populations were held in laboratory culture in a pyrethroid-free environment. After 25 d in culture, F1 juveniles of the Chualar Creek population had cyfluthrin LC_{50}s of 546 and 569 ng/L, and 575 ng/L after 96 d in culture, compared to 353 and 535 ng/L in the two tests done at initial collection. Grayson Creek offspring had cyfluthrin LC_{50}s of 769 and >1197 ng/L after 25 d in culture, compared to >638 and >691 at initial collection. Interpretation of the Grayson Creek data is confounded by several "greater than" LC_{50} values, due to the fact that in most tests with this highly resistant population, we were unable to achieve 50% mortality at the highest concentrations tested. However, it is clear that offspring from both Grayson and Chualar Creeks remained insensitive to pyrethroids, and this lack of sensitivity was not conferred through acclimation to the compounds.

Table S1. Kimura 2-parameter genetic distances, expressed as percentage, between major groups identified in maximum likelihood analysis of the COI data. Data are among-group means calculated in MEGA 5 (18). Mean within-group K2P distances (percentage), also calculated in MEGA, were A=0.8, B=0.4, C+0.1, and D=0.8.

	A	B	C	D
A	--			
B	24.7	--		
C	21.9	20.3	--	
D	26.3	24.4	16.7	==

Table S2. Pyrethroid sensitivity of *H. azteca* from the various sources. LC_{50} values typically determined in two separate trials, with both results reported. Bifenthrin LC_{50} values are provided when available, but some populations lacked sufficient numbers for its determination.

Species Group	Collection site	Cyfluthrin 96-h LC_{50}	Bifenthrin 96-h LC_{50}
Lab cultures			
C	UCB	4.8 (3.9-6.2) 2.9 (2.3-3.9)	8.0 (4.8-11.1) 6.5 (5.7-7.5)
	SIU	1.3 (1.0-1.5) 2.6 (2.3-3.0)	No data
	Chesapeake Cultures	1.7 (1.4-2.1) 1.4 (0.92-2.1)	8.3 (7.0-9.6) 6.8 (5.9-7.8)
Wild populations			
A	Laguna Lake	4.8 (3.7-5.8) 3.9 (3.1-4.8)	6.9 (5.1-8.0) 8.5 (7.4 -9.8)
B	Blodgett Reservoir	1.1 (0.89-1.3) 1.3 (1.1-1.5)	No data
B and D	Pleasant Grove Creek	June 2010: 11.8 (8.8-14.7) 10.3 (7.9-12.9) March 2013: 96 (81.0-114)	June 2010: 24.2 (18.7-30.0) 14.8 (5.3-22.3) March 2013: No data
D	Morrison Creek	132 (63.5-174) 91.6 (65.5-128)	No data
	Mosher Slough	193 (158-226) 211 (176-244)	562 (403-785) >842[a]
	Chualar Creek	535 (403-650) 353 (255-441)	No data
	Grayson Creek	>638[b] >691[b]	No data

[a]Maximum concentration tested yielded less than a 50% response (27% mortality).
[b]Maximum concentration tested yielded less than a 50% response (33% mortality both tests).

Table S3: Primers used for sequencing of targeted regions of the *vgsc* from gDNA.

Vgsc segment	Description	Product size	Sequencing primers
Segment 1	Domain I S6	230 bp 152 bp 480 bp	Lt: TCGTGTTCTTCCTGTTCATCA Rt_1: GGTCTCCTCCCAGGTCGT Rt_2: CTCCTCGGCCTCTCTGATG Rt_3: CCATGTGAATTGGTGAGAACC
Segment 2	linker between domains I and II and Domain II S1	151 bp	Lt: GCCTTTAACTGCCAAGATCAA Rt: GATGAAGAGCTCAACGAACG
Segment 3	Domain II S4-S6	376 bp (as cDNA)	Lt: AGGGTGTTCAAGCTCGCTAA Rt: TGTCGGTGCATGAAACTCAC
Segment 4	Domain III S6 and the linker between domains III and IV	188 bp	Lt: TTCGTCTTCTTCATCATCTTCG Rt: GGTATGGCTTTGAGAGGCTT

For segment I, a total of three primer pairs were utilized to produce products in each of the six wild populations. For segment 3, an intron was present within the amplified product, which varied in length between 260-293 bp.

Table S4 – See separate Excel file for this table.

Table S5. Processes significantly enriched in the up and down-regulated gene lists for *H. azteca* exposed to cyfluthrin at their population-specific no observed effects concentrations in comparison to unexposed animals from same populations.

UCB exposed to 0.4 ng/L: Up-regulated processes

GO:0044057: regulation of system process
GO:0031644: regulation of neurological system process
GO:0043038: amino acid activation
GO:0043039: tRNA aminoacylation
GO:0050804: regulation of synaptic transmission
GO:0050808: synapse organization
GO:0051969: regulation of transmission of nerve impulse
GO:0006399: tRNA metabolic processes
GO:0051049: regulation of transport
GO:0006418: tRNA aminoacylation for protein translation
GO:0046903: secretion
GO:0043062: extracellular structure organization
GO:0034660: ncRNA metabolic processes
GO:0019725: cellular homeostasis
GO:0007268: synaptic transmission
GO:0010646: regulation of cell communication
GO:0019226: transmission of nerve impulse

Table S5 continued:

UCB exposed to 0.4 ng/L: Down-regulated processes

No significant gene ontology terms

Grayson Creek exposed to 170 ng/L: Up-regulated processes

GO:0005975: carbohydrate metabolic process
GO:0055114: oxidation reduction
GO:0008152: metabolic process
GO:0044262: cellular carbohydrate metabolic process
GO:0006066: alcohol metabolic process
GO:0006091: generation of precursor metabolites and energy
GO:0019318: hexose metabolic process
GO:0005996: monosaccharide metabolic process

Grayson Creek exposed to 170 ng/L: Down-regulated processes

GO:0060255: regulation of macromolecule metabolic process
GO:0007165: signal transduction
GO:0050789: regulation of biological process
GO:0050794: regulation of cellular process
GO:0080090: regulation of primary metabolic process
GO:0051716: cellular response to stimuli
GO:0019222: regulation of metabolic process
GO:0023052: signaling
GO:0065007: biological regulation
GO:0000278: mitotic cell cycle
GO:0031323: regulation of cellular metabolic process
GO:0006468: protein amino acid phosphorylation
GO:0019219: regulation of nucleic acid metabolic process
GO:0051171: regulation of nitrogen compound metabolic process
GO:0051246: regulation of protein metabolic process
GO:0000017: alpha-glucoside transport
GO:0042946: glucoside transport
GO:0015766: disaccharide transport
GO:0015771: trehalose transport
GO:0015772: oligosaccharide transport
GO:0007243: intracellular protein kinase cascade
GO:0010605: negative regulation of metabolic process
GO:0051252: regulation of RNA metabolic process
GO:0022402: cell cycle process
GO:0006355: regulation of transcription, DNA-dependent
GO:0010556: regulation of macromolecule biosynthetic process

Table S6. Summary of sequencing reads and annotation of the *H. azteca* transcriptome.

Total number raw sequencing reads	1,039,832
Total number of bases	205 Mb
Number of aligned reads	891,955
Number of bases in aligned reads	147 Mb
Number of contigs	65,961
Number of large contigs	13,113
Average size large contigs (bases)	896
Number of singletons	147,877
Total number of unique sequences	213,838
Average coverage of transcriptome	4X
Number of annotated sequences (with BLAST matches, evalue ≤1E-4)	15,830
Number of sequences with assigned GO terms	5,167

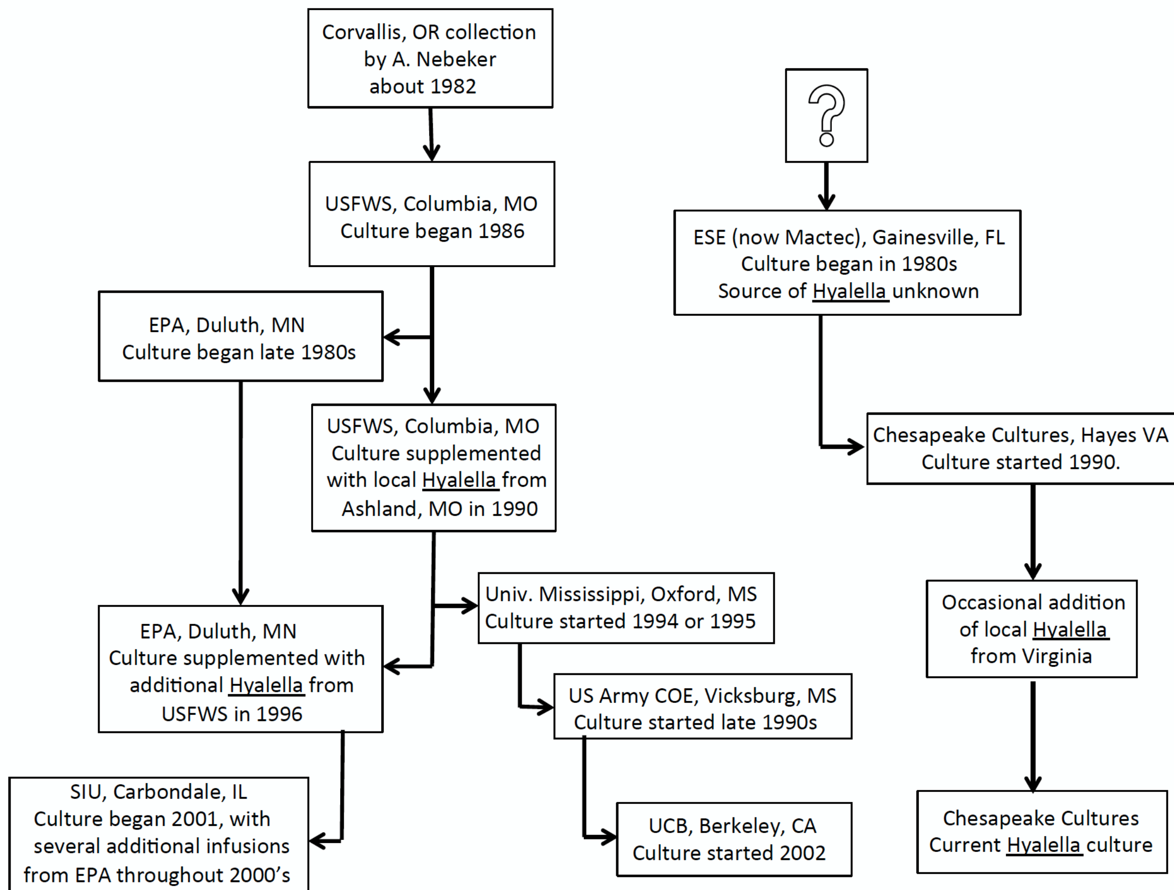


Figure S1. Origin of the laboratory cultures of *H. azteca* used in the present study. History of these cultures has been reconstructed from published literature (19) and personal communications (T. Norberg-King, EPA-Duluth; C. Ingersoll, USGS; E. Wilkins, Chesapeake Cultures; J. Steevens, D. Farrar and T. Bridges, U.S. Army Corps of Engineers)

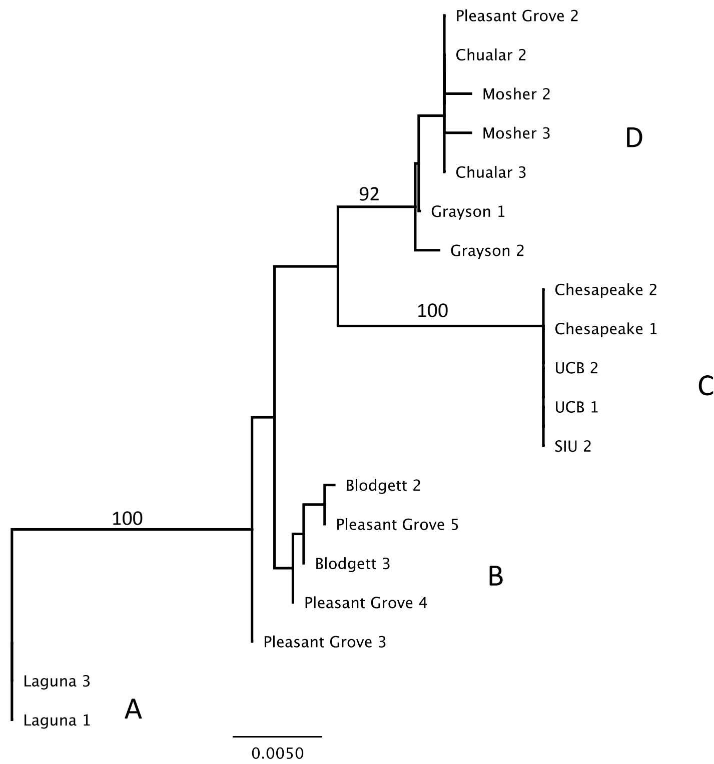


Figure S2. Maximum likelihood phylogram based on *28S* rDNA sequence data. Numbers on branches indicate branch support based on 1000 bootstrap replicates. Group designations A-D indicate putative species. Analysis was conducted with PhyML implemented in Geneious using a JC69 model of nucleotide substitution. Scale bar is number of substitutions per site.

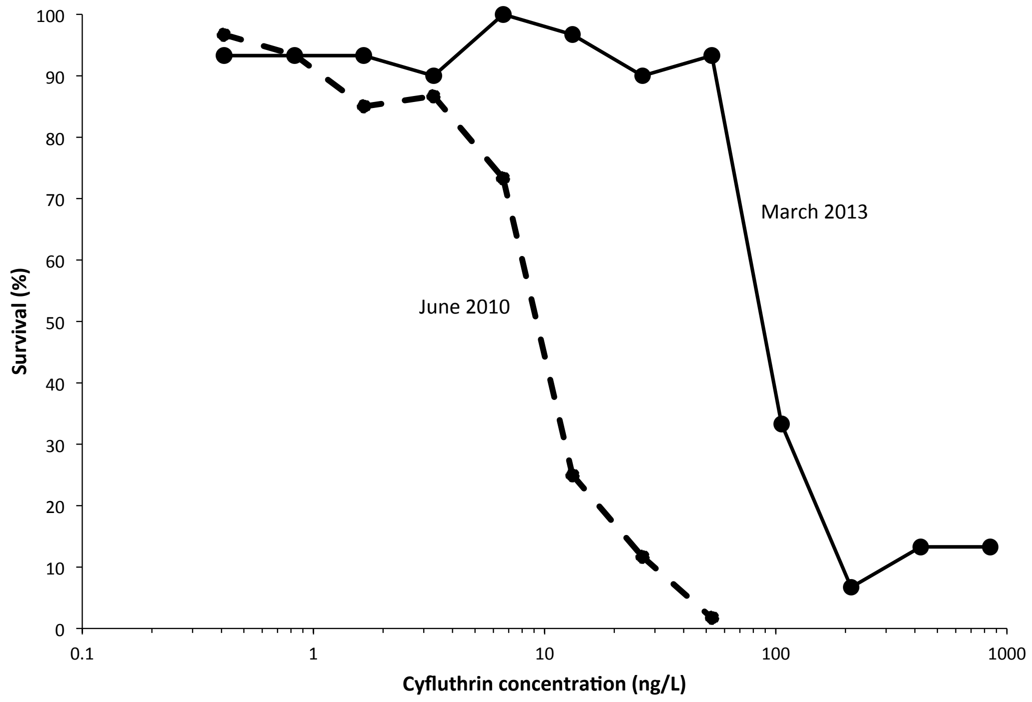
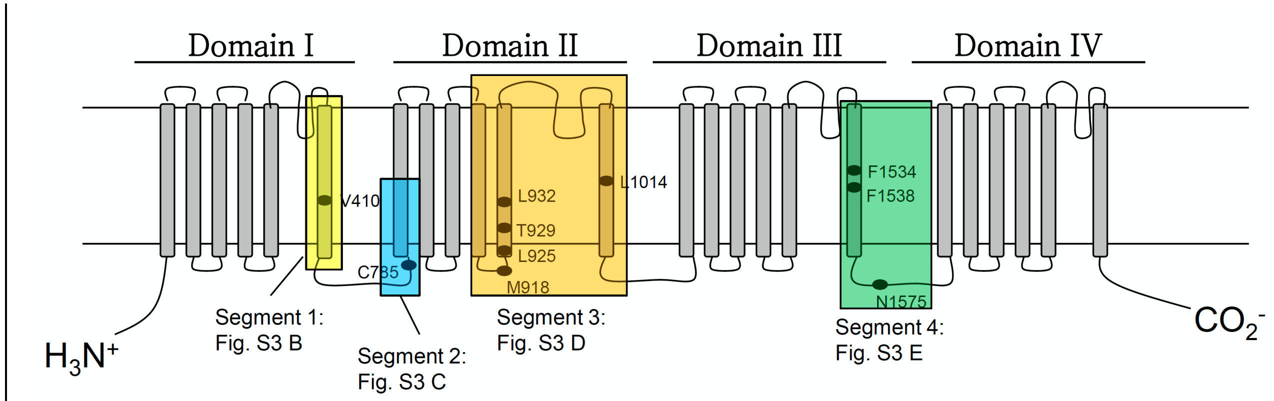


Figure S3. *H. azteca* survival as a function of cyfluthrin water concentration in the toxicity tests, comparing animals collected in 2010 and 2013 at the same location in Pleasant Grove Creek.

Figure S4:

A.



B. Segment 1: Domain I S6

```
UCB           VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
BR            VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
PGC-B-a       VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
PGC-B-b       VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
PGC-D         VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
MS            VFFLFIIFLGPYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
MO            VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
CH-b          VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
CH-a          VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
GC            VFFLFIIFLGSYYLVNLILAIVAMSYDQLQKKAEEEEEAALAEEEAIREAEE
```

C. Segment 2: Linker I&II, Domain II S1

```
UCB           PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
BR            PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
PGC-D         PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
PGC-B-a       PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
PGC-B-b       PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
MS            PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
MO            PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
CH-b          PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
CH-a          PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
GC            PLTAKIKKKVIEYGKRLIEILCVWDCCWLWLKIQHILGLIVFDPFVELFI
```

D. Segment 3: Domain II S4-S6

```
UCB           RVFKLAKSWPTLNLLISIMGKTVGALGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
BR            RVFKLAKSWPTLNLLISIMGKTVGALGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
PGC-D         RVFKLAKSWPTLNLLISIMGKTVGALGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
PGC-B-a.1     RVFKLAKSWPTLNLLISIMGKTVGALGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
PGC-B-a.2     RVFKLAKSWPTLNLLISIMGKTVGALGNLTFVLCIIFFIFAVMGMQLFGKNYTEKVTKFKWSQ
PGC-B-b       RVFKLAKSWPTLNLLISIMGKTVGAIGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
MS            RVFKLAKSWPTLNLLISIMGKTVGAIGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
MO            RVFKLAKSWPTLNLLISIMGKTVGAIGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
CH-b          RVFKLAKSWPTLNLLISIMGKTVGAIGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
CH-a          RVFKLAKSWPTLNLLISILGKTVGALGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
GC            RVFKLAKSWPTLNLLISILGKTVGALGNLTFVLCIIIFIFAVMGMQLFGKNYTEKVTKFKWSQ
```

Figure S4 continued:

```
UCB         DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
BR          DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
PGC-D       DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
PGC-B-a     DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
PGC-B-b     DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
MS          DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
MO          DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
CH-b        DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
CH-a        DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
GC          DGQMPRWNFVDFFHSFMIVFRVLCGEWIESMWDCMYVGDFSCVPFFLATVVIGNLVVSFMHR
```

E. Segment 4: Domain III S6, Linker III&IV

```
UCB       FVFFIIFGSFFTLNLFIGVIIDNFNEQKKKAGGSLEMFMTDDQKKYYNAMKKMGGKKPLK
BR        FVFFIIFGSFFTLNLFIGVIIDNFNEQKKKAGGSLEMFMTDDQKKYYNAMKKMGGKKPLK
PGC-D     FVFFIIFGSFFTLNLFIGVIIDNFNEQKKRAGGSLEMFMTDDQKKYYNAMKKMGGKKPLK
PGC-B-a   FVFFIIFGSFFTLNLFIGVIIDNFNEQKKKAGGSLEMFMTDDQKKYYNAMKKMGGKKPLK
PGC-B-b   FVFFIIFGSFFTLNLFIGVIIDNFNEQKKKAGGSLEMFMTDDQKKYYNAMKKMGGKKPLK
MS        FVFFIIFGSFFTLNLFIGVIIDNFNEQKKKAGGSLEMFMTDDQKKYYNAMNKMGGKKPLK
MO        FVFFIIFGSFFTLNLFIGVIIDNFNEQKKKAGGSLEMFMTDDQKKYYNAMKKMGGKKPLK
CH        FVFFIIFGSFFTLNLFIGVIIDNFNEQKKKAGGSLEMFMTDDQKKYYNAMKKMGGKKPLK
GC        FVFFIIFGSFFTLNLFIGVIIDNFNEQKKKAGGSLEMFMTDDQKKYYNAMKKMGGKKPLK
```

Figure S4. Partial amino acid sequences of VGSC of *H. azteca*. (A) Illustration of the four domains of the *H. azteca* VGSC highlighting amino acid residues, which have been correlated with pyrethroid resistance in multiple insect species (Rinkevich et al., 2013) and were targeted by our sequencing efforts. Colored boxes indicate areas of the VGSC that were sequenced in the lab and wild populations of *H. azteca*. All four sequences were obtained from the same 2-3 individuals representing each population or genotype within the population. (B) Amino acid sequence from Segment 1: Domain I S6. (C) Amino acid sequence from Segment 1 including the linker between domains I and II and Domain II S1. (D) Amino acid sequence from Segment 3 including Domain II S4-S6. The M918L and L925I mutations are highlighted in black. An additional amino acid substitution was found in some individuals of PGC-B, I936F, and is highlighted in grey. (E) Amino acid sequence from Segment 4 including Domain III S6 and the linker between domains III and IV. Two amino acid substitutions were noted and highlighted in grey: K1557R, which was only found in the PGC population, and K1578N which was the major sequence in MS, but was also found in one individual at BR, PGC, MO, and CH. Abbreviations: UCB=lab culture; GC=Grayson Creek; CH=Chualar Creek; MO=Morrison Creek; MS=Mosher Slough population; PGC=Pleasant Grove Creek; BR=Blodgett Reservoir.

Figure S5:

	R	V	F	K	L	A	K	S	W/R	P	T	L	N	L	L	I
UCB	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
BR	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
PGC-D	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	AGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
PGC-B-*a*.1	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTG	AAC	TTG	CTC	ATC
PGC-B-*a*.2	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
PGC-B-*b*	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
MS	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
MO	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
CH-*b*	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
CH-*a*	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATC
GC	AGG	GTG	TTC	AAG	CTC	GCT	AAG	TCC	TGG	CCC	ACG	CTC	AAC	TTG	CTC	ATT

	S	I	M/L	G	K	T	V	G	A	L/I	G	N	L	T	F	V
UCB	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	CTC	GGC	AAC	TTG	ACC	TTC	GTA
BR	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	CTC	GGC	AAC	TTG	ACC	TTC	GTG
PGC-D	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	CTC	GGC	AAC	TTG	ACC	TTC	GTG
PGC-B-*a*.1	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	CTC	GGC	AAC	TTG	ACC	TTC	GTG
PGC-B-*a*.2	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	CTC	GGC	AAC	TTG	ACC	TTC	GTG
PGC-B-*b*	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	ATC	GGC	AAC	TTG	ACC	TTC	GTG
MS	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	ATC	GGC	AAC	TTG	ACC	TTC	GTG
MO	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	ATC	GGC	AAC	CTG	ACC	TTC	GTG
CH-*b*	TCC	ATC	ATG	GGC	AAG	ACG	GTG	GGT	GCC	ATC	GGC	AAC	CTG	ACC	TTC	GTG
CH-*a*	TCC	ATC	TTG	GGC	AAG	ACG	GTG	GGT	GCC	CTC	GGC	AAC	CTG	ACC	TTC	GTG
GC	TCC	ATC	TTG	GGC	AAG	ACG	GTG	GGT	GCC	CTC	GGC	AAC	CTG	ACG	TTC	GTG

	L	C	I	I	I/F	F	I	F	A	V	M	G	M	Q	L	F
UCB	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
BR	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
PGC-D	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
PGC-B-*a*.1	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
PGC-B-*a*.2	CTC	TGC	ATT	ATT	TTC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
PGC-B-*b*	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
MS	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
MO	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
CH-*b*	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
CH-*a*	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC
GC	CTC	TGC	ATT	ATC	ATC	TTC	ATC	TTC	GCC	GTG	ATG	GGC	ATG	CAA	CTC	TTC

	G	K	N	Y	T	E	K	V	T	K	F	K	W	S	Q	D
UCB	GGC	AAG	AAT	TAT	ACT	GAA	AAG	GTC	ACA	AAG	TTC	AAA	TGG	TCT	CAA	GAC
BR	GGC	AAG	AAT	TAT	ACT	GAA	AAA	GTC	ACG	AAG	TTC	AAA	TGG	TCT	CAA	GAC
PGC-D	GGC	AAG	AAT	TAT	ACT	GAA	AAA	GTC	ACG	AAG	TTC	AAA	TGG	TCT	CAA	GAC
PGC-B-*a*.1	GGC	AAG	AAT	TAT	ACT	GAA	AAA	GTC	ACG	AAG	TTC	AAA	TGG	TCT	CAA	GAC
PGC-B-*a*.2	GGC	AAG	AAT	TAT	ACT	GAA	AAA	GTC	ACG	AAG	TTC	AAA	TGG	TCT	CAA	GAC
PGC-B-*b*	GGC	AAG	AAT	TAT	ACT	GAA	AAA	GTC	ACG	AAG	TTC	AAA	TGG	TCT	CAA	GAC
MS	GGC	AAG	AAT	TAT	ACT	GAA	AAA	GTC	ACG	AAG	TTC	AAA	TGG	TCT	CAA	GAC
MO	GGC	AAG	AAT	TAT	ACT	GAA	AAG	GTC	ACA	AAG	TTC	AAA	TGG	TCT	CAA	GAC
CH-*b*	GGC	AAG	AAT	TAT	ACT	GAA	AAG	GTC	ACA	AAG	TTC	AAA	TGG	TCT	CAA	GAC
CH-*a*	GGC	AAG	AAT	TAT	ACT	GAA	AAG	GTC	ACA	AAG	TTC	AAA	TGG	TCT	CAA	GAC
GC	GGC	AAG	AAT	TAT	ACT	GAA	AAG	GTC	ACA	AAG	TTC	AAA	TGG	TCT	CAA	GAC

Figure S5 continued:

	G	Q	M	P	R	W	N	F	V	D	F	F	H	S	F	M
UCB	GGC	CAG	ATG	CCG	AGG	TGG	AAT	TTC	GTG	GAC	TTC	TTC	CAC	TCG	TTC	ATG
BR	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTC	TTC	CAC	TCG	TTC	ATG
PGC-D	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTT	TTC	CAC	TCG	TTC	ATG
PGC-B-*a*.1	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTT	TTC	CAC	TCG	TTC	ATG
PGC-B-*a*.2	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTT	TTC	CAC	TCG	TTC	ATG
PGC-B-*b*	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTC	TTC	CAC	TCG	TTC	ATG
MS	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTC	TTC	CAC	TCG	TTC	ATG
MO	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTC	TTC	CAC	TCG	TTC	ATG
CH-*b*	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTC	TTC	CAC	TCG	TTC	ATG
CH-*a*	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTC	TTC	CAC	TCG	TTC	ATG
GC	GGC	CAG	ATG	CCG	CGG	TGG	AAT	TTC	GTG	GAC	TTC	TTC	CAC	TCG	TTC	ATG

	I	V	F	R	V	L	C	G	E	W	I	E	S	M	W	D
UCB	ATC	GTG	TTC	CGC	GTG	CTG	TGC	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
BR	ATC	GTG	TTC	CGC	GTG	CTG	TGC	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
PGC-D	ATC	GTG	TTC	CGC	GTG	CTG	TGC	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
PGC-B-*a*.1	ATC	GTG	TTC	CGC	GTG	CTG	TGC	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
PGC-B-*a*.2	ATC	GTG	TTC	CGC	GTG	CTG	TGC	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
PGC-B-*b*	ATC	GTG	TTC	CGC	GTG	CTG	TGC	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
MS	ATC	GTG	TTC	CGC	GTG	CTG	TGC	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
MO	ATC	GTG	TTC	CGC	GTG	CTG	TGT	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
CH-*b*	ATC	GTG	TTC	CGC	GTG	CTG	TGT	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
CH-*a*	ATC	GTG	TTC	CGC	GTG	CTG	TGC	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC
GC	ATC	GTG	TTC	CGC	GTG	CTG	TGT	GGG	GAG	TGG	ATC	GAG	AGC	ATG	TGG	GAC

	C	M	Y	V	G	D	F	S	C	V	P	F	F	L	A	T
UCB	TGT	ATG	TAC	GTG	GGA	GAC	TTC	TCG	TGC	GTG	CCT	TTC	TTC	CTG	GCC	ACG
BR	TGC	ATG	TAC	GTA	GGG	GAC	TTC	TCG	TGC	GTA	CCT	TTC	TTC	CTG	GCC	ACG
PGC-D	TGC	ATG	TAC	GTA	GGG	GAC	TTC	TCT	TGC	GTG	CCT	TTC	TTC	CTG	GCC	ACG
PGC-B-*a*.1	TGC	ATG	TAC	GTA	GGG	GAC	TTC	TCT	TGC	GTG	CCT	TTC	TTC	CTG	GCC	ACG
PGC-B-*a*.2	TGC	ATG	TAC	GTA	GGG	GAC	TTC	TCG	TGC	GTA	CCT	TTC	TTC	CTG	GCC	ACG
PGC-B-*b*	TGC	ATG	TAC	GTA	GGG	GAC	TTC	TCG	TGC	GTA	CCT	TTC	TTC	CTG	GCC	ACG
MS	TGC	ATG	TAC	GTA	GGG	GAC	TTC	TCG	TGC	GTA	CCT	TTC	TTC	CTG	GCC	ACG
MO	TGT	ATG	TAC	GTG	GGA	GAC	TTC	TCG	TGC	GTG	CCT	TTC	TTC	CTG	GCC	ACG
CH-*b*	TGT	ATG	TAC	GTG	GGA	GAC	TTC	TCG	TGC	GTG	CCT	TTC	TTC	CTG	GCC	ACG
CH-*a*	TGT	ATG	TAC	GTG	GGA	GAC	TTC	TCG	TGC	GTG	CCT	TTC	TTC	CTG	GCC	ACG
GC	TGT	ATG	TAC	GTG	GGA	GAC	TTC	TCG	TGC	GTG	CCT	TTC	TTC	CTG	GCC	ACG

	V	V	I	G	N	L	V	V	S	F	M	H	R
UCB	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
BR	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
PGC-D	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
PGC-B-*a*.1	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
PGC-B-*a*.2	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
PGC-B-*b*	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
MS	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
MO	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
CH-*b*	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
CH-*a*	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA
GC	GTC	GTC	ATC	GGC	AAC	TTG	GTC	GTG	AGT	TTC	ATG	CAC	CGA

Figure S5. Partial sequence of *vgsc* Domain II of *H. azteca*. Translated amino acid sequence, shown at the top, indicates two resistance mutations in wild populations previously characterized in insects (20), for which the codon is highlighted in black. Two additional amino acid substitutions, highlighted in grey, are found in the Pleasant Grove Creek populations, one in species group D at W908R and one in the species group B at I936F, but their role, if any, in pyrethroid resistance has not been established. Single nucleotide substitutions not affecting the translated sequence and are also highlighted. Two resistant alleles are present in the Chualar Creek population. UCB=lab culture; GC=Grayson Creek; CH=Chualar Creek; MO=Morrison Creek; MS=Mosher Slough population; PGC=Pleasant Grove Creek; BR=Blodgett Reservoir.

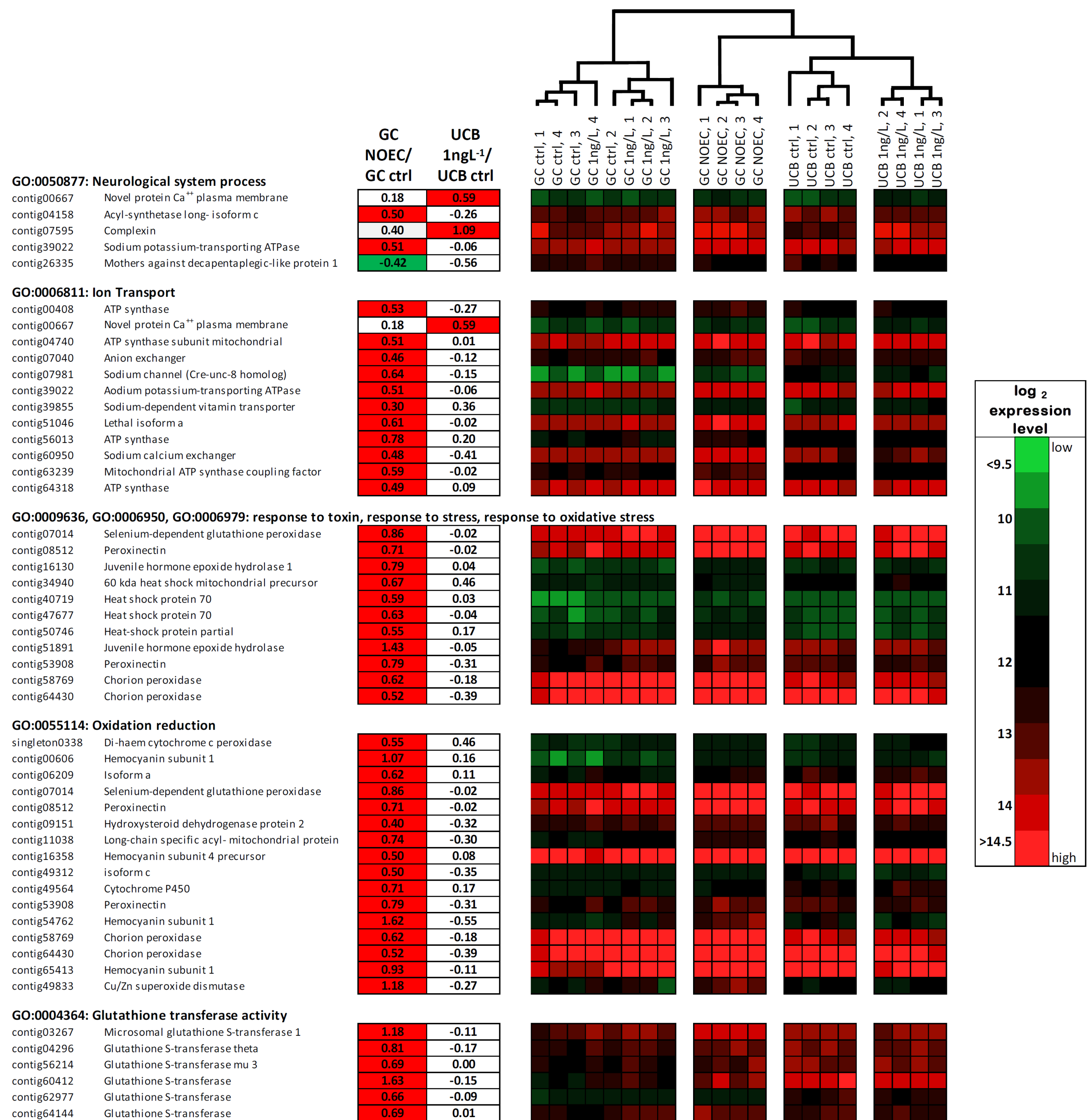


Figure S6. Gene ontology terms associated with neuronal function and stress response that were enriched in the genes by pyrethroid exposure in lab (UCB) and resistant wild (GC) populations. Contigs are shown which were mapped to enriched GO terms along with their log2 ratio (exposed vs. control) (in the middle with significantly differentially expressed genes highlighted in red, up-regulated, or green, down-regulated) and their log2 normalized expression level in all experiments (on the right). Hierarchical clustering analysis was performed to show the relationships between gene expression patterns in each replicate. UCB=University of California Berkeley lab culture; GC=Grayson Creek.

Supporting Information References

1. USEPA (2000) *Methods for Measuring the Toxicity and Bioaccumulation of Sediment-Associated Contaminants with Freshwater Invertebrates*. EPA Publication 600/R-99/064. U.S. Environmental Protection Agency, Washington, D.C.

2. Wang D, Weston DP, Lydy MJ (2009) Method development for the analysis of organophosphate and pyrethroid insecticides at low parts per trillion in water. *Talanta* 78:1345-1351.

3. You J, Lydy MJ (2007) A solution for isomerization of pyrethroid insecticides in gas chromatography. *J Chromatogr A* 1166:181-190.

4. You J, Weston DP, Lydy MJ (2008) Quantification of pyrethroid insecticides at sub-ppb levels in sediment using matrix-dispersive accelerated solvent extraction with tandem SPE cleanup. In Gan J, Spurlock F, Hendley P, Weston DP, eds, *Synthetic Pyrethroids: Occurrence and Behavior in Aquatic Environment.* ACS Symposium Series 991. Oxford University Press, New York, NY, USA, pp 87-113.

5. Wellborn GA, Broughton RE (2008) Diversification on an ecologically constrained adaptive landscape. *Molec Ecol* 17:2927-2936.

6. Witt JDS, Threloff DL, Hebert PDN (2006) DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. *Molec Ecol* 15:3073-3082.

7. Drummond AJ, et al. (2012) Geneious v5.6 (http://www.geneious.com).

8. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Sys Biol* 59:307-21.

9. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotid sequences. *J Mol Evol* 15: 111–120.

10. Tamura K, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731-2739.

11. Borgmann U, Ralph KM, Norwood WP (1989) Toxicity test procedures for *Hyalella azteca*, and chronic toxicity of cadmium and pentachlorophenol to *H. azteca*, *Gammarus fasciatus*, and *Daphnia magna*. *Arch Environ Contam Toxicol* 18:756-764.

12. Borgmann U, Norwood WP, Ralph KM (1990) Chronic toxicity and bioaccumulation of 2,5,2',5'- and 3,4,3',4'-tetrachlorobiphenyl and Aroclor 1242 in the amphipod *Hyalella azteca*. *Arch Environ Contam Toxicol* 19:558-564.

13. Ralston-Hooper K, et al. (2009) Acute and chronic toxicity of atrazine and its metabolites deethylatrazine and deisopropylatrazine on aquatic organisms. *Ecoto*xicology 18:899-905.

14. Meyer E, et al. (2009) Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics* 10:219.

15. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98:5116-5121.

16. Conesa A, et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676.

17. Bluthgen N, et al. (2005) Biological profiling of gene groups utilizing gene ontology. *Genome Inform* 16:106-115.

18. Tamura K, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731-2739.

19. Duan Y, Guttman SI, Oris JT (1997) Genetic differentiation among laboratory populations of *Hyalella azteca*: implications for toxicology. *Environ Toxicol Chem* 16:691-695.

20. Rinkevich FD, Du Y, Dong K (2013) Diversity and convergence of sodium channel mutations involved in resistance to pyrethroids. *Pestic. Biochem. Physiol.* 106:93-100.