Western Kentucky University
**TopSCHOLAR®**

Masters Theses & Specialist Projects | Graduate School

Fall 2017

# Adaptive Variation in Tiger Salamander Populations

Meghan Parsley
*Western Kentucky University*, parsleymeghan1@gmail.com

Follow this and additional works at: https://digitalcommons.wku.edu/theses

Part of the Biology Commons, Ecology and Evolutionary Biology Commons, and the Genomics Commons

Recommended Citation
Parsley, Meghan, "Adaptive Variation in Tiger Salamander Populations" (2017). *Masters Theses & Specialist Projects.* Paper 2056.
https://digitalcommons.wku.edu/theses/2056

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.

# ADAPTIVE VARIATION IN TIGER SALAMANDER POPULATIONS

A Thesis
Presented to
The Faculty of the Department of Biology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Meghan Brooke Parsley

December 2017

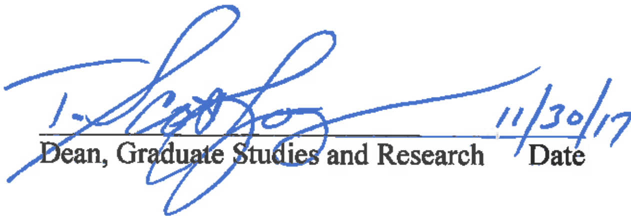
ADAPTIVE VARIATION IN TIGER SALAMANDER POPULATIONS

Date Recommended 15 November 2017

Jarrett R. Johnson, Director of Thesis

T. Keith Philips

Lawrence Alice



Dean, Graduate Studies and Research 11/30/17 Date

I dedicate this thesis to my parents and my sister. They provided me with never-ending support and encouragement and never had a doubt that my thesis was anything other than great.

ACKNOWLEDGMENTS

The completion of this project was made possible only through the help of many people. I would first like to thank my primary advisor Dr. Jarrett Johnson for his guidance and support throughout my thesis. He always had a positive outlook on my project, calmed me down when I thought everything was going wrong, and reminded me “this is how I do it” even when we didn’t know what we were doing. Additionally, I would like to thank my committee members Dr. Keith Philips and Dr. Larry Alice for providing support, feedback and encouragement throughout my project, and Dr. Rodney King for standing in as a committee member and providing feedback during my defense.

I would also like to thank all the biology graduate students that I got to spend countless hours with over the past two years. We experienced the highs and lows of a master’s program together and I knew I could count on them to cheer me up, make me laugh, and also stress out alongside me. I have enjoyed watching us all grow into scientists and look forward to all that we will do in the future. Special thanks to the other members of the Johnson Lab for always being ready to talk herps and for the members of my cohort that put up with me the most.

Lastly, I would like to thank those that have helped me accomplish specific steps along the way and my sources of funding: Evan McCartney-Melstad and the Shaffer Lab at UCLA provided tissue samples and additional sequence data for this project along with near round the clock help for guiding me through the library preparations and bioinformatics steps. Levi Gray and Katlyn Clark helped in the field collecting samples in 2016. Arthur Hunt at the University of Kentucky allowed me to use his lab’s sonicator to shear DNA samples during library preparation. Bobby Gaffney and the WKU

Bioinformatics Center provided computer and bioinformatics support. Naomi Rowland and the WKU Biotechnology center provided guidance on library preparation and access to equipment not available in our lab. Funding for this project was provided by the WKU Graduate School, the National Science Foundation, Kentucky Academy of Sciences, and the WKU Office of Research.

CONTENTS

INTRODUCTION ........................................................................................................ 1

MATERIALS AND METHODS.......................................................................................... 9

RESULTS ............................................................................................................................ 19

DISCUSSION ...................................................................................................................... 22

REFERENCES .................................................................................................................... 29

APPENDIX A: TABLES...................................................................................................... 37

APPENDIX B: FIGURES .................................................................................................... 46

APPENDIX C: BIOINFORMATICS PROGRAMS USED............................................. 63

APPENDIX D: ENVIRONMENTAL DATA ..................................................................... 64

ADAPTIVE VARIATION IN TIGER SALAMANDER POPULATIONS

Meghan Parsley December 2017 67 Pages

Directed by: Jarrett Johnson, Keith Philips, and Lawrence Alice

Department of Biology Western Kentucky University

Amphibians face an unknown future in a time of rapid environmental change due to global climate perturbations. Since amphibians are perceived to be indicators of ecosystem health, understanding the causes of their declines can improve our perception of threats to other species. Molecular techniques have allowed us to explore how environmental change affects genetic variation and to predict evolutionary adaptive potential of amphibian populations. The identification of populations with the greatest potential to respond to changing environmental variables may be an important conservation strategy to aid in future management efforts. I utilized targeted exon capture sequencing to identify adaptive variation in California tiger salamanders (CTS; *Ambystoma californiense*), a species threatened by land use change and hybridization with barred tiger salamanders (*A. mavortium*). I identified 17 and 26 outlier loci for balancing selection in historic and recent samples of CTS respectively. The outlier loci corresponded to genes of various functions, though none of the outliers associated significantly with the change in several tested environmental variables. Despite the lack of environmental correlations detected, it must also be considered that the outlier loci could be involved in epistatic interactions where many genes with small effects influence a single phenotype with fitness benefits. Additional hypotheses to explain the observed changes in allele frequencies and outliers may be the effects of UV-B radiation, pesticide use, or indirect effects of climate change.

# INTRODUCTION

Changes in average global temperatures have been above normal rates for the past six decades, with human greenhouse gas (GHG) emissions at the forefront of causes for this trend (IPCC 2014). Human-mediated climate change has influences on human and natural systems across the globe and includes effects other than rising temperatures. Additional effects may include changes in precipitation, melting snow and ice, increases in extreme weather and climate events, and sea level rise (IPCC 2014). The consequences for biodiversity around the world are staggering as future outlooks on global climate change expect further warming (1.5-2.0$^{o}$C by the end of the 21$^{st}$ century) and dramatic lasting impacts with continued GHG emissions (Sinervo *et al.* 2010; IPCC 2014).

The magnitude of the repercussions on biodiversity caused by global climate change is expected to vary across regions and species. Global climate change is projected to have direct effects on individuals and populations while also altering species interactions and the composition of communities (Hoffmann & Sgrò 2011). Species with small ranges will likely be more severely affected compared to those that are widely distributed, as well as species that are already facing issues relating to fragmentation and other anthropogenic stressors (Hoffmann & Sgrò 2011; Pauls *et al.* 2013). Amphibians are one taxon that are expected to be significantly challenged by changes in climate now and in the future.

Amphibians face many threats due to aspects of their physiology and life history and are strongly affected by temperature and moisture changes (Carey & Alexander 2003; Picco *et al.* 2007). Many amphibians are biphasic and face stressors from both aquatic and terrestrial habitats (Hussain & Pandit 2012; Shaffer *et al.* 2015). Amphibians

also possess permeable skin and lay unshelled eggs that are subject to changes in soil, water and air toxicology (Hussain & Pandit 2012; Shaffer *et al.* 2015). Stressors from the environment can be combined with other threats to amphibians, particularly those posed by humans, to form complex problems for their conservation.

Aside from global climate change, there are five primary hypotheses for the anthropogenic causes of global amphibian declines. These are land use change, introduction of non-indigenous species, commercial over-exploitation, increased application of pesticides, and emerging infectious diseases (Davidson *et al.* 2002; Collins & Storfer 2003; Picco *et al.* 2007). These anthropogenic threats have led to the declaration of over one third of amphibian species as threatened or endangered by the International Union for Conservation of Nature (IUCN), while about one fourth of all species still have insufficient data to be classified as endangered or threatened (Shaffer *et al.* 2015). Scientists first began noting such declines in the 1980's and have since declared the extinction rates of amphibians to have exceeded that of both avian and mammal groups (Collins & Storfer 2003; Picco *et al.* 2007).

Amphibians (especially salamanders) are highly abundant animals in forest and grassland ecosystems and serve important ecological roles. They provide top-down and bottom-up ecosystem services including regulating food webs and providing ecosystem stability (Davic & Welsh 2014). Despite their importance, amphibians are considered to be one of the most threatened vertebrate taxa, with species declines observed across the globe (Hussain & Pandit 2012; Shaffer *et al.* 2015).

The response of organisms to global climate change can occur at the ecological or genetic level. Ecological responses may include shifts in phenology and species ranges,

alterations of trophic interactions, and increases or decreases in parasite and disease infestations (Parmesan 2006). There are also significant ramifications to species' and populations' genetic diversity with a changing climate. Genetic effects of global climate change come through adaptation to changing conditions, selection for phenotypic plasticity, or shifting ranges to avoid extinction (Sinervo *et al.* 2010; Hoffmann & Sgrò 2011; Pauls *et al.* 2013). Evolutionary adaptation will be necessary for species that cannot disperse naturally to desired ranges where climatic conditions are favorable or if human-mediated translocations are not possible (Hoffmann & Sgrò 2011) in the wake of persistent temperature increase.

Evolutionary adaptation is initially a micro-evolutionary process that involves a change in the local selection pressures to favor alleles that are better suited to the new environment (Hoffmann & Sgrò 2011; Pauls *et al.* 2013). This process requires sufficient genetic diversity for natural selection to act upon (Hoffmann & Sgrò 2011; Pauls *et al.* 2013), which is associated with the overall adaptive potential/capacity of a species. Adaptive capacity varies within and between species depending on several factors. Within species variation can be a result of individual populations responding differently to climate change, especially if the selection pressures vary throughout a species range (Pauls *et al.* 2013). Identifying populations and species with the greatest or least adaptive potential can aid in classifying them as either thriving or not surviving in the future of additional changes in climate (Hoffmann & Sgrò 2011). The speed of human-mediated climate change compared to natural variation (IPCC 2014) present additional pressures to populations or species to evolve at a rapid rate to keep up with the pace of observed climatic changes.

Many studies on the effects of global climate change to date have focused on range distributions of species and shifts in phenology, though more empirical evidence on the consequences for genetic diversity is warranted (Pauls *et al.* 2013). Genetic studies on the responses to global warming have focused on neutral variation, despite the fact that coding DNA holds the information to produce proteins that will affect phenotypes and determine the response to selection for increased temperature, which results in the survival of the species (Pauls *et al.* 2013). Recent advances in DNA sequencing technology and computational tools have made the discovery of loci under selection with environmental associations possible, especially in non-model organisms (Pauls *et al.* 2013; Christmas *et al.* 2016). Studies of this type have been conducted on a range of species such as Atlantic cod (*Gadus morhua*, Therkildsen *et al.* 2013), narrow-leaf hopbush (*Dodonaea viscose angustissima*, Christmas *et al.* 2016), and Dall's sheep (*Ovis dalli dalli*, Roffler *et al.* 2016), though studies on amphibians are rare (Bonin *et al.* 2006; Yang *et al.* 2016).

California is a hotspot for declines in amphibian populations by range reductions (Davidson *et al.* 2002). Studies focused on the Great Central Valley of California have shown an overall decline in species distribution of amphibians with the greatest effect in the San Joaquin Valley where the most intensive agriculture is practiced (Fisher & Shaffer 1996). The California tiger salamander (CTS; *Ambystoma californiense*) is one such California endemic experiencing declines. This species occupies a relatively narrow range centered around the Great Central Valley and extends into the surrounding foothills and coastal grasslands (Loredo *et al.* 1996; Shaffer *et al.* 2004). Their range spans from Tulare and San Luis Obispo counties in the south to Sacramento and Solano counties in

the north, with two disjunct populations to the northwest and southwest of the main range (Fig. 1, Shaffer *et al.* 2004).

California tiger salamanders spend the majority of their life underground in California ground squirrel (*Spermophilus beecheyi*) or pocket gopher (*Thomomys bottae*) burrows (Loredo *et al.* 1996; Trenham & Shaffer 2005) which is characteristic of the mole salamander family Ambystomatidae. They migrate mainly to ephemeral ponds for breeding beginning in November, reaching a peak in January and lay eggs in masses where clutch sizes can exceed 1,000 eggs before traveling back to burrows and remaining there for one or more years between breeding (Loredo *et al.* 1996; Trenham 2001; Fitzpatrick *et al.* 2009; Trenham *et al.* 2013). The eggs hatch into fully aquatic larvae, and spend an average of 120-150 days in the aquatic habitat as top predators before undergoing metamorphosis (Fitzpatrick *et al.* 2009; Johnson *et al.* 2013). Metamorphosis occurs in late spring and early summer before juveniles travel to burrows to mature into adults (Loredo *et al.* 1996; Trenham *et al.* 2013).

Recommendations for the federal listing and protection of CTS began in 1989 by Stebbins, who was the first to gather information regarding the species' alarming declines (Shaffer *et al.* 2004). The Santa Barbara and Sonoma county populations were listed as endangered in 2000 and 2003 respectively and the remainder of the CTS range was declared threatened in 2004 under the Endangered Species Act (U.S. Fish and Wildlife Service 2004). These listings were due mainly to range-wide declines, most often associated with habitat destruction as at least 90% of California's original vernal ponds have been destroyed (Fitzpatrick & Shaffer 2007; Fitzpatrick *et al.* 2009; Johnson *et al.* 2013). Additional threats include habitat fragmentation as urbanization has been shown to

have a general negative effect on the abundance of CTS (Loredo *et al.* 1996; Davidson *et al.* 2002; Cook *et al.* 2006) and hybridization with a non-native introduced species (Fitzpatrick & Shaffer 2007).

Barred tiger salamanders (BTS; *A. tigrinum mavortium* or *A. mavortium*) were introduced to the Salinas Valley of California in the 1940's and 50's as bait for bass fisherman (Riley *et al.* 2003; Fitzpatrick *et al.* 2010; Johnson *et al.* 2010b). Thousands of larvae were introduced from Texas, Arizona, Colorado, and New Mexico and released in California ponds in the hopes of establishing fresh bait populations that are available longer into the summer and have larger body sizes than CTS larvae (Riley *et al.* 2003; Fitzpatrick & Shaffer 2007; Ryan *et al.* 2009; Fitzpatrick *et al.* 2010; Johnson *et al.* 2010a). Introductions have since stopped, though after about 5 million years of divergence, CTS and BTS are now sympatric on the California landscape and have hybridized for 10-25 generations (Riley *et al.* 2003; Fitzpatrick *et al.* 2010; Johnson *et al.* 2010a, 2013). Hybrid individuals now occupy at least 20% of the native CTS range in a hybrid swarm located in the Salinas Valley. There is a clear geographic distinction between introgressed populations and native populations of CTS throughout the species range (Fitzpatrick & Shaffer 2007; Fitzpatrick *et al.* 2010; Johnson *et al.* 2013). Hybrid salamanders have a documented impact on the growth and survival of native CTS and other sympatric amphibian species including *Taricha torosa* and *Pseudacris regilla* (Ryan *et al.* 2009).

California tiger salamanders and their hybrids make interesting candidates for the study of their adaptive evolution to a changing environment. Due to their threatened status associated with population declines and alteration of habitat, gaining an

understanding of how anthropogenic changes in environmental variables specifically influence populations can aid in management recommendations for their protection.

Protection of hybridized taxa is largely a grey area of the Endangered Species Act and is currently dealt with on a case by case basis, though management implications with CTS have not been resolved (Fitzpatrick & Shaffer 2007). The Similarity of Appearance provision (Section 4e) of the Endangered Species Act allows non-listed species to be protected if they are difficult to discern from protected species, as CTS may be from their hybrids (US Fish and Wildlife Service 1973; Fitzpatrick & Shaffer 2007). Alternative interpretation of this provision could also result in entire populations being unprotected because of their classification as largely introgressed, though some native CTS may still be present (Fitzpatrick & Shaffer 2007). Strict genetic criteria for protection of the native genotypes may also deliver unwanted consequences of large reductions in protected habitat because of the declaration of hybridization in those areas (Fitzpatrick *et al.* 2010, 2015). Due to difficulties in these classifications and the potential for unwanted negative impacts on the native species, the success of conservation efforts may lie in the comparison of hybrid and native phenotypes and ecological functionality. Their response to different climatic modifications and disturbances can give us insight on how to best conserve CTS in response to a changing climate when dealing with their hybrids (Johnson *et al.* 2013).

I utilized targeted sequence capture genomic techniques to identify genes under selection in response to changing environmental variables over time. This study can help to give a greater understanding to how CTS are adapting to environmental changes and will aid in the conservation efforts and management decisions for this threatened species.

On a larger scale, this understanding can be applied to similar species in predicting the general reaction to anthropogenic environmental change. Gaining a better understanding of the long-term consequences of human-mediated climate change can aid in the protection of amphibian populations.

## MATERIALS AND METHODS

### *Sample Localities*

I selected sample sites using a standing database of CTS tissue samples, looking for ponds that have been sampled at least twice over a time span of at least 10 years. Additionally, I organized sites that fit the following criteria: sites that span the entire latitudinal range of CTS (excluding the Santa Barbara and Sonoma county populations) and sites classified into different levels of hybridization (hybrid swarm, escaped alleles region, and "pure" CTS populations).

I used six to nine tissue samples per sampling location (Fig. 1) per time period (average=8), depending upon the number of samples available from a given year or site and the quality of the tissue samples/extracted DNA for a total of 96 samples for genomic library preparations.

### *DNA Isolation and Genomic Library Preparation*

A simplified schematic of the genetic manipulation is provided in Figure 2. I extracted DNA from all individuals using either an ethanol precipitation protocol (Sambrook & Russell 2001) or DNeasy Blood and Tissue Kit (QIAGEN), utilizing an RNase step in both protocols. I then quantified the extracted DNA using a NanoDrop 2000 (Thermo Scientific) and diluted or concentrated samples to 100ng/µL for sonication depending on DNA concentration. Sonication took place at the University of Kentucky with a BioRuptor NGS (Diagenode) using the following protocol: 28 cycles of 30s on high, 90s off, set at 4°C. Following sonication, I checked fragment sizes using agarose gel

electrophoresis. I found that most samples were sheared to 100-300bp in length, though some appeared not fully sheared and still in large fragments around 2000bp. I sent the samples not fully sheared to the University of Louisville Genomics Core Laboratory for additional sonication on a Covaris Ultrasonicator (Covaris). Following the second round of sonication, samples were bioanalyzed on a high sensitivity DNA1000 chip (Agilent Technologies) to ensure fragments were in the 100-300bp range before further manipulation.

After sonication, I performed a dual SPRI size selection at 0.8X-1.1X with either Agencourt AMPure XP beads (Beckman Coulter) or non-commercial Serapure beads (Rohland & Reich 2012) to remove fragments larger and smaller than desired. I then quantified the fragmented and size selected DNA samples again with a NanoDrop 2000 before library preparations.

I performed half reactions of standard Illuminia library preparations using a Kapa LTP Library Preparation Kit for Illuminia Platforms (Kapa Biosystems) with 278-1890ng of input DNA depending on extraction success. Library preparation included end repair (35µL reaction), A-tailing (25µL reaction) and adapter ligation (25µL reaction) steps. I attached Adapterama dual-indices (iTru5 set 18 and iTru7 set 108, Glenn *et al.* 2016) via limited cycle PCR. After library preparation, I performed a 1.0X single sided size selection with Serapure beads to remove unwanted primer dimers and quantified all libraries with a NanoDrop 2000.

I prepared species-specific $c_0$t-1 (fragments of DNA that contain highly repetitive sequences in the CTS genome) for use in target enrichment to block undesired repetitive sections of the genome that are common in salamanders and prevent nonspecific

hybridization of target probes. To prepare $c_0t$-1 DNA, I extracted genomic DNA from CTS hybrid crosses with an ethanol precipitation protocol, utilizing an RNase step, and adjusted the concentration of the extractions to two 500µL tubes at 1000ng/µL in 1.2X saline sodium citrate buffer (SSC). I sheared one tube (500µL) of extracted DNA on a BioRuptor NGS using the following protocol: 30 cycles of 30s on high, 90s off, set at 4$^o$C, and the other tube (500µL) at the University of Louisville Genomics Core Laboratory on a Covaris Ultrasonicator (due to timing of extractions). I analyzed fragment sizes using agarose gel electrophoresis and a bioanalyzer (high sensitivity DNA1000 chip) for each tubes respectively to ensure fragments were 100-400bp in length. After fragmentation, I treated DNA as described in McCartney-Melstad *et al.* 2016: I denatured fragments for 10 minutes at 95$^o$C and allowed fragments to renature partially for 5 minutes at 60$^o$C before placing them on ice for two minutes. I transferred the partially renatured fragments to a heating block at 42$^o$C and added 250µL of S1 Nuclease in buffer to each of the tubes and incubated them for 1 hour to remove non-repetitive regions of the genome. After incubation, I precipitated $c_0t$-1 DNA with 75µL of 3M sodium acetate and 750µL of 100% isopropanol and centrifuged the samples for 20 minutes at 13,000RPM at 4$^o$C in a microcentrifuge. I removed the supernatant and washed the pellets of DNA with 70% ethanol and centrifuged again at 13,000RPM for 10 minutes at 4$^o$C before removing the ethanol and allowing the pellets to dry. I rehydrated each tube with 50µL 10mM Tris-HCl, pH 8 before combining them and dehydrating in a CentriVap Concentrator (LabConoco) to the appropriate volume prior to use.

I then performed MYbaits (MYcroarray) in-solution sequence capture for targeted high-throughput sequencing with baits designed specifically for CTS. The set includes

40,011 probes that target 5,237 exons from unique genes for a total target region of approximately 1.69 megabases. I first pooled individuals randomly into groups of eight for each of the 12 MYbaits capture reactions with total input DNA from each sample ranging from 114ng to 500ng depending on the yield of each sample from library preparations and followed the MYbaits protocol version 3.01. I utilized 2065ng of the synthesized CTS-specific $c_0t$-1 per capture reaction to replace the human $c_0t$-1 DNA blocker provided in the kit and hybridized probes for 31 hours at 65ºC. Despite preventative measures taken as outlined in the MYbaits protocol, I still experienced evaporation of some samples after the hybridization process. I rehydrated samples with 15µL of nuclease-free water to continue the reaction (approximate volume after hybridization of the rest of the capture reactions). After streptavidin bead binding and washing steps in the protocol, I then amplified enriched fragments (with beads still in solution) with 14 cycles of PCR. With limited Kapa HiFi Hot Start Ready Mix (Kapa Biosystems), I adjusted PCR volumes for each capture reaction to 75% of the recommended volumes, while keeping all reaction concentrations the same as following: 18.75µL of 2X Kapa HiFi Hot Start Ready Mix, 3.75µL of nuclease-free water, 1.875µL of each forward and reverse adapter specific library primers, and 11.25µL of enriched library for a total of 37.5µL. Following post-capture PCR enrichment, I removed the streptavidin beads from each reaction and performed a 1.8X size selection with Serapure beads to remove any primer dimers leftover from the PCR reactions.

With sample manipulation complete, I then quantified libraries with Kapa Library Quantification Kit for Illumina Platforms (Kapa Biosystems) that includes SYBR Fast dye with Illumina specific sequencing primers and a passive ROX dye. I performed

1/1,000 and 1/100,000 dilutions of each capture library to ensure a wide breadth of concentration values to determine where the prepared libraries fall in the range of standards provided, though a 1/100,000 dilution of libraries is recommended in the protocol for post-amplification target capture. I ran 10µL reactions on an Applied Biosystems 7300 (Applied Biosystems) with the recommended qPCR protocol including the optional melt curve analysis and analyzed the output with the Kapa Library Quantification Data Analysis Template following the recommended standards for accuracy. Analysis of the first qPCR revealed that the 1/1,000 dilutions were mostly out of the range of the standards and the melt-curve analysis also showed significant amounts of primer-dimers in the libraries. To improve library quality, I performed another 1.8X size selection with Serapure beads as was conducted after the enrichment process and repeated the qPCR for library quantification with 10µL reactions and dilutions at 1/10,000 and 1/100,000 and saw improvement on primer-dimer contamination.

I sent all capture reactions to Vincent J. Coates Genomics Sequencing Laboratory at University of California Berkeley for additional quantification, fragment analysis, and pooling on a single lane of sequencing on an Illumina HiSeq 2500 for 150bp paired-end reads.

*Bioinformatics*

I checked the general quality of demultiplexed reads (per base sequence quality, average sequence quality, etc.) from the genomics laboratory with fastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) then removed Illumina and index adapters and performed quality trimming with Trimmomatic 0.32 (Bolger *et al.*

2014). I generated index files for each sample containing the iTru5 and iTru7 index sequences and flanking stubs/adapters and used 'grep' to check for accurate adapter composition in the sequences before trimming. I used Trimmomatic settings to remove leading bases with a phred score less than five and trailing bases with a phred score less than 15. Additionally, I utilized the sliding window feature to scan the sequence from the 5' to 3' direction and trim all trailing bases when the average phred score of the four bases in the window dropped below 20. Finally, I discarded all reads less than 40 bases in length and again checked quality of reads with fastQC to ensure sufficient trimming. To merge overlapping paired read outputs from Trimmomatic, I used the program fastq-join (Aronesty 2013) on a Linux operating system.

I mapped individual reads to a combined assembly of targeted regions using Burrows-Wheeler Aligner (BWA) and implemented the BWA-MEM algorithm that is recommended for 70bp or longer reads from Illumina platforms (Li 2013). I then sorted and merged SAM file outputs into a single BAM file for each individual using SAMtools version 1.3.1 (Li *et al.* 2009) before marking duplicate sequence reads that originated from a single DNA molecule using Picard Tools (http://broadinstitute.github.io/picard) 'MarkDuplicates' function. Duplicate reads of a single DNA strand are commonly associated with library preparations involving PCR and can affect variant calling (Van der Auwera *et al.* 2014). To obtain mapping rates and number of marked duplicates, I utilized the SAMtools 'flagstat' function that counts flags in SAM/BAM files.

Prior to SNP calling, I added read groups using Picard Tools 'AddOrReplaceReadGroups' (http://broadinstitute.github.io/picard) to the mapped and duplicate marked BAM files to add an identifying header for each individual. I then

merged BAM files of all individuals using BamTools (Barnett *et al.* 2011) before using FreeBayes v.1.1.0 (Garrison & Marth 2012) to call variants. Individuals in each population and time period were analyzed together as is recommended for increasing the discriminant power of the algorithm. I then filtered the variants of all samples using VCFtools (Danecek *et al.* 2011) based on the following criteria: minor allele frequency greater than 10%, missing data per SNP less than 25%, mean minimum depth of 10 reads per SNP over all individuals, and removed indels. Additionally, I filtered individuals with less than 5X mean depth across the remaining SNPs, removing 8 individuals from further analysis. Finally, I evaluated coverage on the SNPs for each individual and removed those with greater than 10% missing sites (17 individuals) from further analysis.

Due to the poor sequencing and mapping quality of six out of the eight 1991 samples from the Grant 1 location, I received raw sequence data for two additional samples from that site in 1991 from a collaborator at UCLA (Shaffer Lab). The two samples were subjected to the same library preparation protocols and were included in the bioinformatics pipeline as outlined above. For a full list of programs used and their respective purposes in data processing, see Appendix C.

*Environmental Data*

I collected environmental data involving temperature, precipitation, and drought conditions to evaluate changes in climate linked to global warming in my study area. From the Western US Climate Historical Summaries database maintained by the Western Regional Climate Center and the National Centers for Environmental Information maintained by NOAA, I gathered measurements of extreme monthly maximum and

minimum temperatures, average monthly maximum and minimum temperatures, number of rain days per month above trace amounts, and total monthly precipitation from weather stations in close proximity to the six sampling sites (Table 1). Weather station selection priority was placed on availability of data as well as latitudinal and elevational similarities to sample ponds. Additionally, I collected information about drought conditions from the West Wide Drought Tracker (Abatzoglu *et al.* 2017) in the form of monthly Palmer Z-Indices and annual self-calibrated Palmer Drought Severity Indices (sc-PDSI). Palmer drought indices take into account temperature and precipitation data to measure water excess or deficit in the environment. Z-indices measure short term drought conditions on a monthly scale and sc-PDSI measure long-term drought conditions. I narrowed environmental data to the months of April, May, and June for sampling years of each site to target conditions that would have the greatest impact on survival of recently metamorphosed salamanders when dispersing from natal ponds. Though the greatest mortality event likely occurs in the egg stage where survival to hatching has been documented to be as low as 0-10% for *A. tigrinum* (Anderson *et al.* 1971), climatic variables such as temperature and precipitation are likely to influence mortality through changes in pond hydroperiod and potentially directly affect survival when they are crossing the landscape in search of appropriate terrestrial microhabitat.

*Statistical Analysis*

I evaluated the success of library preparations by performing an ANOVA to compare the effect of the samples being pooled into different MYbaits reactions on both the number of sequence reads and the percentage mapped to target regions per individual. Additionally,

I performed linear regressions to determine if the total amount of input DNA to the library preparations as well as the age of the sample in years at the time of extraction were associated with the number of sequence reads and the percentage mapped to targets. Finally, I assessed the relationship between the proportion of joined overlapping reads and the number of sequence reads mapped to target regions using a linear regression. All ANOVAs and linear regressions were calculated in RStudio (RStudio Team, 2015).

To identify outlier loci within my dataset and evaluate population metrics, I used the program BayeScan 2.0 (Foll & Gaggiotti 2008). BayeScan is a Bayesian $F_{ST}$ outlier detection method that calculates the posterior probability of each locus being under selection by comparing a neutral model and a model including selection. I ran BayeScan on historic and recent sampling data separately first including all six study populations, then again after removing hybrid swarm populations (JCL- Pond H and Bluestone Quarry) to see if including those populations significantly affected the loci determined to be under selection. I used the following parameters: 20 pilot runs of 5,000 iterations, 50,000 burn in period, and 5,000 iterations of the model with a thinning interval of 10. Additionally, I set the prior odds of selection for the samples to be 100:1 to reduce the risk of false positives often created by this program.

BayeScan outputs include $F_{ST}$ coefficients for every population, $F_{ST}$ coefficients averaged over populations for each locus, alpha values indicating the strength and direction of selection, and the posterior probability for the model including selection for each locus. I used the outputs from the BayeScan runs including all study populations to evaluate changes and direction of population-level and locus-specific $F_{ST}$ coefficients. I created a $\Delta F_{ST}$ variable by subtracting historic $F_{ST}$ values from contemporary values for

each locus and performed a one-sample t-test with a null distribution created by the data (10,000 random resampling) to determine whether the average change in $F_{ST}$ observed differed from that expected at random. I also compared outlier loci detected between historic and contemporary samples and when hybrid swarm populations were excluded and included to determine if hybrid populations had a significant effect on outlier loci detected.

To determine whether unique contemporary outlier loci were under selection pressure due to changes in climate, I utilized the program BayeScEnv (de Villemereuil & Gaggiotti 2015) that detects local adaptation using an $F_{ST}$ model similar to BayeScan given an environmental variable. I ran the program on contemporary SNP data only and transformed environmental data to denote change over time (contemporary values minus historic values for the same variable) and used the following parameters: 20 pilot runs of 5,000 iterations, 50,000 burn in period, and 5,000 iterations of the model with a thinning interval of 10. I also used the default parameters for the prior probabilities (pi = 0.1 and p = 0.5). Loci were determined to be significantly associated with a given environmental parameter when the q-value of the environmental correlation was less than 0.05.

## RESULTS

My sequencing strategy resulted in 146,706,038 total paired end reads across all 96 individuals. The number of reads per sample averaged 1,528,187.9 with a standard deviation of 1,098,132.15, and a minimum and maximum number of reads per individual of 6,235 and 5,265,790 respectively (Table 2). Quality trimming with Trimmomatic resulted in keeping forward and reverse read pairs from all sequences on average 93.17% per sample. Only the forward read was kept on average 3.81%, reverse only reads kept 1.10%, and both read pairs discarded 1.91% for each sample. An average of 72.91% of the paired end reads were merged using fastq-join (Table 2), resulting in an average joined sequence length of 92.97 base pairs. Mapping using BWA resulted in an average of 13.74% of reads mapped to the target assembly per sample (range: 2.89%- 34.39%, Table 2). Of the mapped reads, the proportion of PCR duplicates ranged from 3.34% to 63.11%, with an average of 28.22%.

Overall, my results show no significant effect of library preparation methods on sequencing success based on the ANOVA and linear regression results. The twelve MYbaits capture reactions had no significant effect on the number of sequence reads for each individual (Fig. 3, $F_{(11,84)} = 1.23$, $P = 0.28$), or on the proportion of reads mapped to the desired targets (Fig. 4, $F_{(11,84)} = 0.49$, $P = 0.91$). Despite differences in extraction success, there was no significant relationship between total input DNA to the library preparations and number of sequence reads per individual (Fig. 5, $F_{(1,94)} = 1.60$, $P = 0.21$), or the proportion of sequences mapped to target regions (Fig. 6, $F_{(1,94)} = 0.95$, $P = 0.33$). Additionally, the age of the sample at the time of extraction also had no significant relationship to the number of sequence reads (Fig. 7, $F_{(1,94)} = 1.85$, $P = 0.18$), or the

proportion mapped to targets (Fig. 8, $F_{(1,94)} = 1.96$, $P = 0.16$). There was a minor improvement in the mapping rate of sequences with increasing proportion of joined sequences, though predictability of this was low (Fig. 9, $F_{(1,94)} = 3.73$, $P = 0.06$, r = 0.20, $r^2 = 0.04$).

After filtering, 13,647 SNPs remained for outlier detection with BayeScan. Population-level $F_{ST}$ values varied by population and time period (Fig. 10). Bluestone Quarry and JCL-Pond H, ponds located in the hybrid swarm region, had markedly lower $F_{ST}$ values (0.1-0.4) than the remainder of the populations in this study (0.7-0.9), indicating high levels of heterozygosity in hybrid populations and high levels of homozygosity in pure CTS populations. Differences in $F_{ST}$ coefficients between historic and recent sampling periods for each population were marginal for all populations except Grant 1 exhibited a slight increase in $F_{ST}$ over time, and JCL-Pond H exhibited a considerable decrease over time (Fig. 10). Changes in $F_{ST}$ for each locus across all populations also varied in magnitude and direction throughout the genome (Fig. 11). I observed both increases and decreases in $F_{ST}$ coefficients over time at varying degrees on all chromosomes except chromosome 10, where the only deviations observed were minimal reductions in $F_{ST}$ in a few loci. A one-sample t-test with a null distribution created by randomized resampling indicated that the observed average $\Delta F_{ST}$ across all 13,647 loci of -0.012 differed significantly from that expected by chance (Fig. 12; $P < 0.001$).

BayeScan identified 1,235 and 17 outliers at the false discovery rate (FDR) of 0.05 and 0.01 respectively for the historic samples and 1,549 and 26 outlier SNPs respectively for contemporary samples when all six populations were included in the

analysis. To narrow my focus to loci with the strongest deviation from expectation and very low prevalence of false positives, I focused on those meeting the FDR = 0.01 level for both time periods. Ten out of 17 loci detected as outliers in the historic samples were also identified as outliers in the contemporary samples (Table 3). Additionally, all loci identified as outliers by BayeScan had low $F_{ST}$ values (Figs. 13 and 14), and negative alpha values (Table 3), indicating that they are candidates for balancing or purifying selection instead of diversifying selection.

When removing hybrid populations from the analysis, BayeScan identified 11 and 8 outliers at FDR of 0.05 and 0.01 respectively in historic populations and 10 and 6 outliers respectively for contemporary populations (Table 4). All outliers identified were still candidates for balancing selection with low $F_{ST}$ values (Figs. 15 and 16) and negative alpha values, and three loci overlapped between the two time periods. Additionally, all outlier loci were exclusively a subset of those identified from all six populations. Despite fewer number of outlier SNPs detected, the removal of hybrid swarm populations had little impact on the outlier loci detected in this system.

When analyzing contemporary outliers that were unique compared to the historic samples, BayeScEnv did not identify any significant correlations between the outlier loci and the change in environmental variables related to temperature, precipitation, or drought considered over the course of this study.

# DISCUSSION

This study identified evidence of selection over several decades in a federally endangered amphibian species. Out of 13,647 SNPs, I identified 17 and 26 outlier loci for balancing selection in historic and recent samples of CTS respectively, but did not find significant correlations of the outliers to environmental variables related to global climate change that were tested. The outlier loci identified correspond to a variety of gene functions for each time period classification (historic outliers, recent outliers, or shared outliers, Table 3, Fig. 17). Historic outlier loci were related to catabolic enzymes such as hydrolases and kinases, cell survival and DNA damage response, cytoskeletal motor proteins, mRNA manipulation, and notochord development. Shared outliers still under selection pressures include RNA polymerase subunits and transcription factors, integrin, ring finger, various enzymatic proteins, and several genes associated with microtubule function and structure during mitosis. Finally, outliers detected to have undergone selective pressures during the course of this study correspond to proteolysis, steroidogenesis, centrosome stabilization during mitosis, hydrolase activation, and several genes related to stress response, cell protection, and DNA repair. Although these recent outliers did not significantly associate with the climate change related variables that were  analyzed, it is possible that they could be associated with other environmental stressors that were not tested here.

Stressors such as UV-B radiation, pesticide application, or indirect effects of climate change may be alternative pressures that contribute to the outlier loci detected in this study. Increased production of chlorofluorocarbons (CFCs) has resulted in depletion of the stratospheric ozone layer, causing significant increases in UV-B radiation reaching the earth's surface in the past several decades (Blaustein & Belden 2003; Blaustein *et al.*

2010). UV-B radiation causes mutation and cell death in developing amphibians and can cause individual mortality or have other sub-lethal damage (Blaustein *et al.* 2010). One defense mechanism against UV-B radiation is efficient DNA repair, often involving photolyase enzymes (Blaustein & Belden 2003; Blaustein *et al.* 2010). Though the genetic underpinnings of UV-B damage repair are still generally unknown, it is possible that the outlier loci only detected in contemporary samples that relate to cell stress response or DNA repair are the result of increased exposure to UV-B radiation throughout the time period of the study, though additional information would be necessary to make this conclusion. Additionally, increases in pesticide use may also influenced genetic change. Localities used in this study are located in California's highly agricultural Great Central Valley, meaning that CTS are likely exposed to stress from pesticide runoff or drift by wind from nearby fields (Fisher & Shaffer 1996; Blaustein *et al.* 2010). Finally, I was unable to analyze the influence of indirect effects of climate change such as food availability, predator-prey interactions, competition, diseases, and host-pathogen dynamics (Blaustein *et al.* 2010). These factors make studying the effects of climate change problematic as causative agents are difficult to disentangle and threats to amphibians are incredibly complex.

All of the outlier loci I detected were candidates for balancing selection meaning there was less change in allele frequencies from an ancestral state than what was expected based on background frequencies. Gene flow can be ruled out as the primary reason for low $F_{ST}$ values at these loci because of the distance between sampling sites (45-75 km), providing evidence that selection must be the cause. High differentiation of loci across the study populations likely made it difficult to detect directional selection (loci with $F_{ST}$

values higher than expected) in the samples. Average $F_{ST}$ values were around 0.6 for the majority of SNPs, which is higher than most other studies (Krauss *et al.* 2013; Tsumura *et al.* 2014; Roffler *et al.* 2016) but is similar to that found by Yang et al. (2016). Their study found only candidate outliers for balancing selection among high altitude and geographically distant populations of amphibians (Yang *et al.* 2016).

It must also be considered that the outliers detected could be involved in epistatic interactions where many genes with small effects influence a single phenotype with fitness benefits (Shafer *et al.* 2015). The outlier detection programs available currently do not factor in these effects as it is difficult to disentangle gene interactions. Additionally, there are alternative explanations for outlier loci other than selection to keep in mind. These can include demography or sequence assembly artifacts (Shafer *et al.* 2015) that should also be considered when evaluating results. For example, it is possible that the detection of outlier loci in this study was affected by unequal representation of contigs in the mapping assembly used (Table 5). This may have lead to a potential overrepresentation of linkage groups 6, 8, and 11, or underrepresentation in other linkage groups.

*Conservation Implications*

The lack of significant environmental correlations to the outlier loci detected in this study reveals that the direct impacts of climate change on this species are still unknown. It is possible that the temporal range of sampling used in this study (12-29 years) was not sufficient to capture significant changes in climate or allele frequencies in the populations. Despite an average overall increase in global temperatures in the past

century, changes in climate vary by year and location in their direction and magnitude (IPCC 2014). Changes in environmental variables in the study area were not consistent in direction and magnitude (see Appendix D) and may have not captured climatic change as expected. Additionally, sufficient time may not have passed during the study to observe significant changes in allele frequencies in response to climatic changes. Another tiger salamander species (*A. t. melanostictum*) has a generation time estimated to be four years (McMenamin & Hadly 2012), implying that this study only captured an estimated 3-6 generations.

However, the information identified here regarding outlier loci can still be considered and investigated for future conservation action for this species. Information on outlier loci, neutral loci, and phenotypic data should all be combined to make educated decisions regarding what management actions should be taken to best conserve CTS (Funk *et al.* 2012; Shafer *et al.* 2015). Additionally, further analysis on the effects of a changing climate on CTS should be investigated to best prepare for the future of this species.

*Utility of Target Capture and Temporal Sampling*

Target capture and other reduced representation sequencing methods continue to be more widely utilized in ecological and evolutionary biology. Reference genomes for model organisms are now widely available and resources for non-model organisms are increasing, giving scientists the ability to obtain thousands of SNPs with relative ease (Manel *et al.* 2015). Currently, one of the main objectives of population genomics is identifying areas of the genome that are under selection (Foll & Gaggiotti 2008) for a variety of organisms. The selective enrichment of target capture approaches, that used in

this study, is especially useful for disease research in biomedical fields but is also advantageous for use in population genomics studies (Jones & Good 2016). Target capture has been used for genetic mapping of phenotypic traits, phylogenetic studies, sequencing ancient DNA, performing metagenomics analyses from environmental samples, and identifying signatures of selection throughout the genome (Jones & Good 2016).

Using target capture to identify signatures of selection comes with several advantages over other reduced representation sequencing methods. One advantage is that coding regions of the genome can be selected for enrichment when baits are designed to target exons (exome capture, Jones & Good 2016). This reduces the need for additional validation to determine whether candidate loci for selection are in coding regions or in close proximity to be linked to coding regions that can arise when using other sequencing methods such as restriction site associated DNA sequencing (RAD-seq, Manel *et al.* 2015). In the case of this study, target capture was an ideal method to target exons throughout the genome and avoid sequencing large amounts of repetitive non-coding DNA that is prevalent in the genome of salamanders. Another advantage of using target capture sequencing to detect signatures of selection is the ability to target *a priori* candidate genes using a completely annotated, closely related reference genome. For example, bighorn sheep (*Ovis canadensis*) and domestic sheep (*Ovis aries*) genome annotations were used to test associations between candidate genes involved in immunological, metabolic, and growth functions and environmental variables related to latitude, longitude, precipitation, and temperature in Dall's sheep (*Ovis dalli dalli*,

Roffler *et al.* 2016). Though these resources are not yet available for a wide variety of organisms, they present an promising way to set up hypothesis driven experiments.

Detecting contemporary evolution in response to climate change previously required DNA samples throughout history to draw conclusions. However, temporal ranges of genetic data across decades are nonexistent for most species (Balanyá *et al.* 2006). These challenges can be overcome with new computational and statistical programs to detect outlier loci, but does not rule out the efficacy of utilizing museum specimens or other preserved organisms to study changes through time in response to climate change (Pauls *et al.* 2013; Urban *et al.* 2014). Additionally, the pace of observed climatic changes makes understanding the expected time scales of evolutionary responses to environmental change increasingly important and the use of temporally spaced sampling is the only way to facilitate that.

*Conclusions*

Identifying signatures of adaptive evolution and the selective pressures behind them are only the first steps for conservation of focal species in a changing climate. Next, experimental validation of phenotypes and fitness implications of outlier loci is required to move information from identifying outliers to making conservation and management decisions (Vitti *et al.* 2013; Manel *et al.* 2015; Shafer *et al.* 2015). These next steps may take time as greater genomic annotation is likely required for in depth validation studies, therefore, some degree of uncertainty is expected when searching for adaptive loci at this point, especially when trying to apply information to conservation measures (Shafer *et al.* 2015). Though I was unable to identify the selective pressures behind the outlier loci

detected, further research on this system is needed to validate the outlier loci detected in this study and determine if climate change affects amphibian species in predictable ways. Additionally, phenotype and fitness consequences of those loci should also be investigated to better recommend conservation action for CTS.

REFERENCES

Abatzoglu JT, McEvoy DJ, Redmond KT (2017) The West Wide Drought Tracker: Drought monitoring at fine spatial scales. *Bulletin of the American Meteorological Society*. In press.

Anderson JD, Hassinger DD, Dalryn GH (1971) Natural mortality of eggs and larvae of *Ambystoma t. tigrinum*. *Ecology,* **52**, 6, 1107-1112.

Aronesty E (2013) Comparison of sequencing utility programs. *The Open Bioinformatics Journal*, **7**, 1–8.

Balanyá J, Oller JM, Huey RB *et al.* (2006) Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science*, **313**, 1773–1775.

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT (2011) BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics*, **27**, 1691–1692.

Blaustein AR, Belden LK (2003) Amphibian defenses against UV-B radiation. *Evolution & Development*, **5**, 89–97.

Blaustein AR, Walls SC, Bancroft BA *et al.* (2010) Direct and indirect effects of climate change on amphibian populations. *Diversity*, **2**, 281–313.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution*, **23**, 773–783.

Carey C, Alexander M (2003) Climate change and amphibian declines: is there a link ? *Diversity and Distributions*, **9**, 111–121.

Christmas MJ, Biffin E, Breed MF, Lowe AJ (2016) Finding needles in a genomic haystack: targeted capture identifies clear signatures of selection in a nonmodel plant species. *Molecular Ecology*, **25**, 4216–4233.

Collins JP, Storfer A (2003) Global amphibian declines: Sorting the hypotheses. *Diversity and Distributions*, **9**, 89–98.

Cook DG, Trenham PC, Northen PT (2006) Demography and breeding phenology of the California tiger salamander (*Ambystoma californiense*) in an urban landscape. *Northwestern Naturalist*, **87**, 215.

Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Davic D, Welsh H (2014) On the ecological roles of salamanders. *Annual Review of Ecology, Evolution and Systematics*, **35**, 405–434.

Davidson C, Shaffer HB, Jennings MR (2002) Spatial tests of the pesticide drift, habitat destruction, UV-B, and climate-change hypotheses for California amphibian declines. *Conservation Biology*, **16**, 1588–1601.

Fisher RN, Shaffer HB (1996) The decline of amphibians in California's Great Central Valley. *Conservation Biology*, **10**, 1387–1397.

Fitzpatrick BM, Johnson JR, Kump DK *et al.* (2009) Rapid fixation of non-native alleles revealed by genome-wide SNP analysis of hybrid tiger salamanders. *BMC Evolutionary Biology*, **9**, 176.

Fitzpatrick BM, Johnson JR, Kump DK *et al.* (2010) Rapid spread of invasive genes into a threatened native species. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 3606–3610.

Fitzpatrick BM, Ryan ME, Johnson JR, Corush J, Carter ET (2015) Hybridization and the species problem in conservation. *Current Zoology*, **61**, 206–216.

Fitzpatrick BM, Shaffer HB (2007) Introduction history and habitat variation explain the landscape genetics of hybrid tiger salamanders. *Ecological Applications*, **17**, 598–608.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A bayesian perspective. *Genetics*, **180**, 977–993.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, **27**, 489–496.

Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. *ArXiv e-prints,* 1207:3907.

Glenn TC, Nilsen RA, Kieran TJ *et al.* (2016) Adapterama I: Universal stubs and primers for thousands of dual-indexed Illumina libraries (iTru & iNext). *BioRxiv,* 049114.

Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. *Nature*, **470**, 479–485.

Hussain QA, Pandit AK (2012) Global amphibian declines: A review. *International Journal of Biodiversity and Conservation*, **4**, 348–357.

IPCC (2014) Climate Change 2014: Synthesis Report. Contribution of working groups I, II, and III to the Fifth Assessment Report of the Intergovernmental Panel on

Climate Change [Core Writing Team, R.K. Pachauri and L.A. Mayer (eds.)]. Geneva, Switzerland, 151 pp.

Johnson JR, Fitzpatrick BM, Shaffer HB (2010a) Retention of low-fitness genotypes over six decades of admixture between native and introduced tiger salamanders. *BMC Evolutionary Biology*, **10**, 147.

Johnson JR, Johnson BB, Shaffer HB (2010b) Genotype and temperature affect locomotor performance in a tiger salamander hybrid swarm. *Functional Ecology*, **24**, 1073–1080.

Johnson JR, Ryan ME, Micheletti SJ, Shaffer HB (2013) Short pond hydroperiod decreases fitness of nonnative hybrid salamanders in California. *Animal Conservation*, **16**, 556–565.

Jones MR, Good JM (2016) Targeted capture in evolutionary and ecological genomics. *Molecular Ecology*, **25**, 185–202.

Krauss SL, Sinclair EA, Bussell JD, Hobbs RJ (2013) An ecological genetic delineation of local seed-source provenance for ecological restoration. *Ecology and Evolution*, **3**, 2138–2149.

Li H (2013) Aligining sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv e-prints,*1303.3997.

Li H, Handsaker B, Wysoker A *et al.* (2009) The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Loredo I, Vuren D Van, Morrison ML (1996) Habitat use and migration behavior of the California tiger salamander. *Journal of Herpetology*, **30**, 282–285.

Manel S, Perrier C, Pratlong M *et al.* (2015) Genomic resources and their influence on the detection of the signal of positive selection in genome scans. *Molecular Ecology*, **25**, 170–184.

McCartney-Melstad E, Mount GG, Shaffer HB (2016) Exon capture optimization in large-genome amphibians. *Molecular Ecology Resources*, **16**, 1084–1094.

McMenamin SK, Hadly EA (2012) Ancient DNA assessment of tiger salamander population in Yellowstone National Park. *PLoS ONE*, **7**, 1–6.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology and Systematics*, **37**, 637–669.

Pauls SU, Nowak C, Bálint M, Pfenninger M (2013) The impact of global climate change on genetic diversity within populations and species. *Molecular Ecology*, **22**, 925–946.

Picco AM, Brunner JL, Collins JP (2007) Susceptibility of the endangered California tiger salamander, *Ambystoma californiense*, to ranavirus infection. *Journal of Wildlife Diseases*, **43**, 286–290.

Riley SPD, Shaffer HB, Voss SR, Fitzpatrick BM (2003) Hybridization between a rare, native tiger salamander (*Ambystoma californiense*) and its introduced congener. *Ecological Applications*, **13**, 1263–1275.

Roffler GH, Amish SJ, Smith S *et al.* (2016) SNP discovery in candidate adaptive genes using exon capture in a free-ranging alpine ungulate. *Molecular Ecology Resources*, **16**, 1147–1164.

Rohland N, Reich D (2012) Cost-effective , high-throughput DNA sequencing. *Genome Research*, **22**, 939–946.

Ryan ME, Johnson JR, Fitzpatrick BM (2009) Invasive hybrid tiger salamander genotypes impact native amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 11166–11171.

Sambrook J, Russell D (2001) *Molecular Cloning: A Laboratory Manual (3-Volume Set)*. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Shafer ABA, Wolf JBW, Alves PC *et al.* (2015) Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution*, **30**, 78–87.

Shaffer HB, Gidiş M, McCartney-Melstad E *et al.* (2015) Conservation genetics and genomics of amphibians and reptiles. *Annual Review of Animal Biosciences*, **3**, 113–138.

Shaffer HB, Pauly GB, Oliver JC, Trenham PC (2004) The molecular phylogenetics of endangerment: Cryptic variation and historical phylogeography of the California tiger salamander, *Ambystoma californiense*. *Molecular Ecology*, **13**, 3033–3049.

Sinervo B, Mendez-de-la-Cruz F, Miles DB *et al.* (2010) Erosion of lizard diversity by climate change and altered thermal niches. *Science*, 328, 894–899.

Therkildsen NO, Hemmer-Hansen J, Hedeholm RB *et al.* (2013) Spatiotemporal SNP analysis reveals pronounced biocomplexity at the northern range margin of Atlantic cod *Gadus morhua*. *Evolutionary Applications*, 6, 690–705.

Trenham P (2001) Terrestrial habitat use by adult California tiger salamanders. *Journal of Herpetology*, **35**, 343–346.

Trenham PC, Shaffer HB (2005) Amphibian upland habitat use and its consequences for population viability. *Ecological Applications*, **15**, 1158–1168.

Trenham PC, Shaffer HB, Koenig WD *et al.* (2013) Life history and demographic variation in the California tiger salamander (*Ambystoma californiense*). *Copeia*, **2000**, 365–377.

Tsumura Y, Uchiyama K, Moriguchi Y *et al.* (2014) Genetic differentiation and evolutionary adaptation in *Cryptomeria japonica*. *G3: Genes|Genomes|Genetics*, **4**, 2389–2402.

Urban MC, Richardson JL, Freidenfelds NA (2014) Plasticity and genetic adaptation mediate amphibian and reptile responses to climate change. *Evolutionary Applications*, **7**, 88–103.

US Fish and Wildlife Service (1973) Endangered Species Act of 1973. *Federal Register*, 43098-43105.

US Fish and Wildlife Service (2004) Determination of threatened status for the California tiger salamander; and special rule exemption for existing routine ranching activities; final rule. *Federal Register*, **69**, 47212–47248.

Van der Auwera GA, Carneiro MO, Hartl C *et al.* (2014) From FastQ data to high confidence varant calls: the Genome Analysis Toolkit best practices pipeline. *Currrent Protocols in Bioinformatics*, **11,** 1110.

de Villemereuil P, Gaggiotti OE (2015) A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, **6,** 1248–1258.

Vitti JJ, Grossman SR, Sabeti PC (2013) Detecting natural selection in genomic data. *Annual Review of Genetics*, **47**, 97–120.

Yang W, Qi Y, Fu J (2016) Genetic signals of high-altitude adaptation in amphibians : a comparative transcriptome analysis. *BMC Genetics*, **17**, 1–10.

## APPENDIX A: TABLES

**Table 1.** Sample localities with their associated weather stations used for environmental data and station IDs. (Listed in north-south order, see Fig. 1)

| **Sample Site** | **Weather Station** | **Station ID** |
|---|---|---|
| Dunnigan | Woodland 1 WNW, CA US | USC00049781 |
| Jepson-Olcott | Fairfield, CA US | USC00042934 |
| Camino Diablo 2 | Tracy Pumping Plant, CA US | USC00049001 |
| Grant 1 | Los Gatos, CA US | USC00045123 |
| Bluestone Quarry | Gilroy, CA US | USC00043417 |
| JCL-Pond H | Salinas Municipal Airport, CA US | USW00023233 |

**Table 2.** Summary of library preparation and sequencing data for each individual used in this study including total input DNA to library preparations, capture (MYbaits) reaction, number of sequence reads per individual, proportion of overlapping sequences joined, and the proportion of sequences mapped to target regions of the genome.

| Sample | Site | Year | Total input DNA (ng) | MYbaits reaction | Number of reads | Percent of reads joined | Percent of reads mapped to targets |
|---|---|---|---|---|---|---|---|
| 6686 | Jepson-Olcott | 1986 | 230 | 7 | 312,964 | 71.26 | 12.45 |
| 6688 | Jepson-Olcott | 1986 | 154 | 11 | 11,667 | 61.65 | 8.61 |
| 6689 | Jepson-Olcott | 1986 | 168 | 9 | 6,235 | 50.92 | 8.13 |
| 6691 | Jepson-Olcott | 1986 | 500 | 10 | 2,336,238 | 77.21 | 32.47 |
| 6692 | Jepson-Olcott | 1986 | 148 | 9 | 8,329 | 63.57 | 6.92 |
| 6693 | Jepson-Olcott | 1986 | 500 | 8 | 695,222 | 68.83 | 22.72 |
| 6695 | Jepson-Olcott | 1986 | 500 | 11 | 1,767,129 | 79.45 | 11.17 |
| 6696 | Jepson-Olcott | 1986 | 500 | 7 | 2,276,033 | 71.27 | 10.97 |
| 6697 | Jepson-Olcott | 1986 | 500 | 2 | 2,763,835 | 73.24 | 11.45 |
| 8270 | Camino Diablo 2 | 1988 | 500 | 2 | 1,901,987 | 79.22 | 3.65 |
| 8272 | Camino Diablo 2 | 1988 | 126 | 4 | 796,824 | 76.17 | 9.97 |
| 8273 | Camino Diablo 2 | 1988 | 406 | 11 | 2,370,089 | 72.12 | 8.99 |
| 8274 | Camino Diablo 2 | 1988 | 354 | 12 | 23,951 | 66.25 | 6.31 |
| 8412 | Camino Diablo 2 | 1988 | 178 | 7 | 1,541,233 | 71.82 | 12.96 |
| 8413 | Camino Diablo 2 | 1988 | 282 | 2 | 1,066,622 | 73.17 | 31.13 |
| 8417 | Camino Diablo 2 | 1988 | 236 | 11 | 3,690,803 | 78.16 | 11.74 |
| 8419 | Camino Diablo 2 | 1988 | 114 | 7 | 854,452 | 72.79 | 10.32 |
| 12499* | Grant 1 | 1991 | - | - | 2,441,247 | 60.61 | 23.79 |
| 12505 | Grant 1 | 1991 | 500 | 3 | 639,032 | 66.25 | 12.01 |
| 12508 | Grant 1 | 1991 | 500 | 12 | 30,409 | 63.60 | 8.35 |
| 12509 | Grant 1 | 1991 | 500 | 1 | 2,065,962 | 58.61 | 3.53 |
| 12510 | Grant 1 | 1991 | 212 | 8 | 1,506,703 | 74.31 | 25.74 |
| 12511 | Grant 1 | 1991 | 470 | 1 | 953,891 | 59.40 | 4.03 |
| 12514* | Grant 1 | 1991 | - | - | 3,562,933 | 64.09 | 24.03 |
| 12515 | Grant 1 | 1991 | 388 | 8 | 11,036 | 59.04 | 10.61 |
| 12516 | Grant 1 | 1991 | 500 | 4 | 3,107,951 | 46.15 | 4.20 |
| 12518 | Grant 1 | 1991 | 500 | 2 | 187,953 | 59.28 | 13.53 |
| 14380 | Dunnigan | 1992 | 500 | 2 | 788,890 | 60.85 | 6.01 |
| 14381 | Dunnigan | 1992 | 500 | 7 | 1,019,453 | 50.11 | 6.45 |
| 14382 | Dunnigan | 1992 | 432 | 4 | 988,243 | 73.62 | 10.10 |
| 14383 | Dunnigan | 1992 | 500 | 1 | 480,962 | 75.29 | 18.06 |

| Sample | Site | Year | Total input DNA (ng) | MYbaits reaction | Number of reads | Percent of reads joined | Percent of reads mapped to targets |
|---|---|---|---|---|---|---|---|
| 14385 | Dunnigan | 1992 | 284 | 6 | 4,961,668 | 72.01 | 13.07 |
| 14388 | Dunnigan | 1992 | 500 | 11 | 2,619,520 | 64.96 | 6.39 |
| 21550 | JCL-Pond H | 1998 | 230 | 10 | 2,068,568 | 72.56 | 3.33 |
| 21551 | JCL-Pond H | 1998 | 428 | 12 | 494,933 | 71.13 | 12.30 |
| 21554 | JCL-Pond H | 1998 | 228 | 3 | 2,530,778 | 74.16 | 11.91 |
| 21555 | JCL-Pond H | 1998 | 500 | 4 | 305,143 | 66.50 | 28.61 |
| 21558 | JCL-Pond H | 1998 | 484 | 6 | 1,861,058 | 60.94 | 28.36 |
| 21559 | JCL-Pond H | 1998 | 224 | 6 | 640,547 | 67.38 | 3.18 |
| 21560 | JCL-Pond H | 1998 | 486 | 9 | 2,419,577 | 71.61 | 6.38 |
| 26983 | Bluestone Quarry | 1999 | 500 | 10 | 686,189 | 69.15 | 10.54 |
| 26984 | Bluestone Quarry | 1999 | 500 | 9 | 378,631 | 73.34 | 10.75 |
| 26986 | Bluestone Quarry | 1999 | 500 | 3 | 1,307,911 | 77.99 | 7.36 |
| 26987 | Bluestone Quarry | 1999 | 500 | 7 | 1,569,074 | 78.48 | 27.64 |
| 26990 | Bluestone Quarry | 1999 | 500 | 12 | 1,699,242 | 77.41 | 6.42 |
| 26992 | Bluestone Quarry | 1999 | 264 | 5 | 1,114,140 | 74.20 | 3.46 |
| 28782 | Bluestone Quarry | 1999 | 282 | 8 | 4,757,598 | 73.04 | 8.29 |
| 28791 | Bluestone Quarry | 1999 | 500 | 11 | 2,211,142 | 78.58 | 31.98 |
| 119832 | Bluestone Quarry | 2011 | 256 | 6 | 1,492,407 | 71.39 | 23.66 |
| 119836 | Bluestone Quarry | 2011 | 500 | 5 | 924,081 | 73.81 | 17.39 |
| 119837 | Bluestone Quarry | 2011 | 426 | 1 | 507,546 | 48.80 | 10.28 |
| 119839 | Bluestone Quarry | 2011 | 286 | 8 | 444,454 | 73.11 | 22.43 |
| 119840 | Bluestone Quarry | 2011 | 500 | 1 | 2,218,848 | 76.30 | 11.30 |
| 119841 | Bluestone Quarry | 2011 | 278 | 2 | 346,342 | 71.50 | 12.64 |
| 119842 | Bluestone Quarry | 2011 | 476 | 3 | 766,288 | 72.36 | 13.86 |
| 119843 | Bluestone Quarry | 2011 | 500 | 10 | 1,061,423 | 64.85 | 3.27 |
| 121218 | JCL-Pond H | 2010 | 360 | 1 | 408,425 | 76.20 | 22.52 |
| 121220 | JCL-Pond H | 2010 | 500 | 10 | 1,620,488 | 73.91 | 3.24 |
| 121221 | JCL-Pond H | 2010 | 500 | 12 | 1,155,908 | 77.48 | 10.29 |
| 121223 | JCL-Pond H | 2010 | 500 | 5 | 1,166,339 | 75.00 | 3.29 |
| 121224 | JCL-Pond H | 2010 | 500 | 9 | 1,509,932 | 78.52 | 30.84 |
| 121225 | JCL-Pond H | 2010 | 500 | 3 | 517,361 | 76.65 | 25.26 |
| 121226 | JCL-Pond H | 2010 | 500 | 12 | 1,326,319 | 76.78 | 2.89 |
| 121227 | JCL-Pond H | 2010 | 500 | 11 | 1,560,229 | 78.38 | 31.81 |
| 121228 | JCL-Pond H | 2010 | 500 | 10 | 2,289,981 | 74.94 | 31.34 |
| 126227 | Jepson-Olcott | 2015 | 500 | 6 | 669,412 | 75.49 | 15.37 |
| 126230 | Jepson-Olcott | 2015 | 500 | 11 | 1,628,141 | 79.30 | 8.38 |
| 126231 | Jepson-Olcott | 2015 | 500 | 2 | 1,488,919 | 76.87 | 3.25 |
| 126232 | Jepson-Olcott | 2015 | 500 | 4 | 439,287 | 60.18 | 19.65 |

| Sample | Site | Year | Total input DNA (ng) | MYbaits reaction | Number of reads | Percent of reads joined | Percent of reads mapped to targets |
|---|---|---|---|---|---|---|---|
| 126233 | Jepson-Olcott | 2015 | 500 | 9 | 2,012,544 | 74.33 | 29.89 |
| 126234 | Jepson-Olcott | 2015 | 500 | 10 | 1,325,605 | 78.57 | 11.05 |
| 126235 | Jepson-Olcott | 2015 | 486 | 3 | 1,622,939 | 85.56 | 31.88 |
| 126237 | Jepson-Olcott | 2015 | 500 | 5 | 1,293,992 | 83.49 | 12.97 |
| 126238 | Jepson-Olcott | 2015 | 500 | 12 | 2,075,313 | 81.02 | 13.88 |
| 126247 | Dunnigan | 2015 | 500 | 4 | 2,405,868 | 85.15 | 11.73 |
| 126248 | Dunnigan | 2015 | 432 | 8 | 3,345,038 | 80.44 | 6.90 |
| 126251 | Dunnigan | 2015 | 500 | 7 | 1,812,401 | 83.31 | 9.15 |
| 126252 | Dunnigan | 2015 | 500 | 1 | 1,915,171 | 79.79 | 33.34 |
| 126253 | Dunnigan | 2015 | 500 | 10 | 1,281,963 | 85.64 | 3.76 |
| 126255 | Dunnigan | 2015 | 500 | 3 | 687,129 | 82.08 | 17.65 |
| 126256 | Dunnigan | 2015 | 500 | 6 | 1,687,457 | 77.93 | 31.70 |
| 126257 | Dunnigan | 2015 | 500 | 4 | 3,398,363 | 78.53 | 6.81 |
| 126293 | Camino Diablo 2 | 2015 | 500 | 9 | 588,811 | 59.85 | 27.55 |
| 126294 | Camino Diablo 2 | 2015 | 500 | 6 | 5,265,790 | 79.62 | 13.20 |
| 126295 | Camino Diablo 2 | 2015 | 500 | 8 | 2,344,048 | 74.59 | 3.27 |
| 126296 | Camino Diablo 2 | 2015 | 500 | 5 | 2,356,442 | 75.36 | 34.39 |
| 126297 | Camino Diablo 2 | 2015 | 500 | 6 | 876,461 | 64.78 | 3.25 |
| 126298 | Camino Diablo 2 | 2015 | 500 | 9 | 1,162,270 | 78.41 | 13.11 |
| 126299 | Camino Diablo 2 | 2015 | 500 | 4 | 3,319,504 | 74.09 | 6.40 |
| 126300 | Camino Diablo 2 | 2015 | 500 | 3 | 2,146,383 | 78.44 | 8.75 |
| 127903 | Grant 1 | 2016 | 500 | 2 | 1,053,830 | 81.72 | 13.80 |
| 127904 | Grant 1 | 2016 | 500 | 1 | 607,901 | 79.94 | 25.73 |
| 127905 | Grant 1 | 2016 | 424 | 5 | 795,369 | 83.91 | 17.70 |
| 127906 | Grant 1 | 2016 | 500 | 12 | 1,062,287 | 81.31 | 14.08 |
| 127907 | Grant 1 | 2016 | 500 | 5 | 2,207,875 | 82.65 | 9.11 |
| 127908 | Grant 1 | 2016 | 500 | 8 | 3,811,703 | 79.35 | 6.75 |
| 127909 | Grant 1 | 2016 | 500 | 5 | 3,338,932 | 80.09 | 12.61 |
| 127910 | Grant 1 | 2016 | 500 | 7 | 1,532,702 | 80.21 | 3.35 |
| **Average** | | | | | **1,528,187.9** | **72.91** | **13.74** |

*Raw sequence reads obtained from another study, data not included in averages

**Table 3.** Historic (H), contemporary (C), and shared (S) outlier loci identified by BayeScan 2.0 from all six populations when FDR = 0.01 with corresponding $\Delta F_{ST}$ and alpha values from the program. Outlier SNPs were mapped back to contigs and linkage groups (LG) corresponding to those in the linkage map of *A. mexicanum* and respective GenBank and XenBase BLAST searches for gene values and potential functions. (Alpha values presented as historic | contemporary when locus is an outlier for both periods)

| Outlier SNP Index | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| H | S | C | LG | Contig | Position | $\Delta F_{ST}$ | Alpha | Gene | Function |
| | | 765 | 11 | Contig125111 | 160 | -0.010 | -2.208 | HERPUD2 | Membrane associated ER protein involved in stress response |
| | 1129 | | 8 | Contig138515 | 487 | -0.002 | -2.729 \| -2.779 | POLR2K | RNA polymerase subunit |
| | 1489 | | 11 | Contig145190 | 47 | 0.008 | -2.852 \| -2.811 | RNF145 | Ring finger protein |
| | | 1767 | 6 | Contig156100 | 353 | -0.009 | -2.198 | RABGAP1 | GTPase activating protein |
| | 1940 | | 6 | Contig157828 | 73 | -0.003 | -2.148 \| -2.190 | PHF6 | Transcription factor |
| | | 2193 | 11 | Contig171591 | 477 | -0.009 | -2.208 | TIGAR | Negative regulator of glycolysis, contributes to the protection of cells from oxidative or metabolic stress-induced cell death, may be involved in mitophagy inhibition |

| H | S | C | LG | Contig | Position | $\Delta F_{ST}$ | Alpha | Gene | Function |
|---|---|---|---|---|---|---|---|---|---|
| | | 2524 | 8 | Contig189377 | 263 | -0.067 | -1.860 | ZBTB26 | Zinc finger protein |
| | | 3032 | 8 | Contig194081 | 611 | -0.083 | -2.097 | CXorf40A | May be involved in cell protection from inflammation |
| | 3449 | | 6 | Contig202512 | 489 | 0.048 | -2.940 \| -2.522 | FAM96A | May be involved in chromosome segregation |
| 3471 | | | 6 | Contig202588 | 325 | -0.002 | -2.146 | FDFT1 | Squalene synthetase |
| | | 4905 | 6 | Contig211872 | 105 | -0.011 | -2.198 | CCDC47 | May be involved in RNA, calcium, or protein binding |
| | | 5088 | 11 | Contig215601 | 321 | -0.023 | -2.207 | PSMB4 | Multicatalytic proteinase complex subunit |
| | | 5312 | 11 | Contig222474 | 624 | -0.006 | -2.214 | TEX30 | - |
| 5316 | | | 11 | Contig222474 | 658 | -0.001 | -2.168 | TEX30 | Predicted alpha/beta hydrolase |
| | | 5324 | 11 | Contig222474 | 729 | 0.061 | -2.635 | TEX30 | - |
| | | 5383 | 6 | Contig227314 | 553 | -0.011 | -2.205 | RPS27A | Creates a ubiquitin compound that targets cellular proteins for degradation |
| | | 6262 | 6 | Contig314965 | 359 | -0.025 | -1.677 | ACTR1A | Cytoskeletal constituent |

| H | S | C | LG | Contig | Position | $\Delta F_{ST}$ | Alpha | Gene | Function |
|---|---|---|---|---|---|---|---|---|---|
| | 6662 | | 12 | Contig315407 | 251 | -0.066 | -2.118 | -2.655 | APCS | May control degradation of chromatin |
| | 7042 | | 8 | Contig315943 | 648 | -0.055 | -2.236 | -2.704 | ITGBL1 | Integrin beta subunit |
| 7086 | | | 11 | Contig315991 | 674 | 0.001 | -2.147 | SCYL3 | Serine/threonine kinase, may be involved in cell adhesion and migration |
| | | 7813 | 11 | Contig316997 | 704 | -0.007 | -2.201 | Unknown | - |
| 8307 | | | 12 | Contig317800 | 68 | -0.0002 | -2.159 | UFSP2 | Cysteine protease |
| | | 8388 | 11 | Contig317897 | 469 | -0.012 | -2.212 | FANCA | DNA repair protein |
| 9611 | | | 11 | Contig321148 | 286 | -0.002 | -2.151 | RNF146 | Mediator of ubiquitin ligase activity, may regulate cell survival and DNA damage response |
| | | 11310 | 6 | Contig333238 | 731 | -0.006 | -2.199 | PARL | Integral membrane protease involved in intramembrane proteolysis |
| | | 11551 | 8 | Contig338415 | 666 | 0.081 | -2.728 | HSD17B11 | Metabolizes secondary alcohols and ketones in steroidogenesis |

| H | S | C | LG | Contig | Position | $\Delta F_{ST}$ | Alpha | Gene | Function |
|---|---|---|---|---|---|---|---|---|---|
| | 11790 | | 6 | Contig345632 | 733 | -0.045 | -2.557 | -3.012 | NUP37 | Component of nuclear pore complex, required for normal microtubule-kinetochore interaction during mitosis |
| | 11869 | | 11 | Contig347756 | 556 | -0.002 | -2.222 | -2.266 | TEKT4 | Microtubule constituent, structural component of filamentous polymers |
| | 12269 | | 11 | Contig387709 | 293 | -0.079 | -2.324 | -3.057 | MTHFS | Contributes to tetrahydrofolate metabolism |
| | | 12464 | 8 | Contig43717 | 35 | 0.066 | -2.805 | KIZ | Centrosomal protein, strengthens and stabilizes pericentriolar region prior to spindle formation |
| 12514 | | | 6 | Contig452440 | 575 | 0.074 | -2.468 | DYNLRB2 | Dynein-associated protein roadblock |
| | 12686 | | 11 | Contig539011 | 507 | 0.030 | -3.003 | -2.714 | COX7A2 | Oxidoreductase |
| 12956 | | | 6 | Contig75581 | 484 | 0.002 | -2.158 | QKI | Regulates mRNA splicing, export, and stability, translation repressor, essential for notochord development |

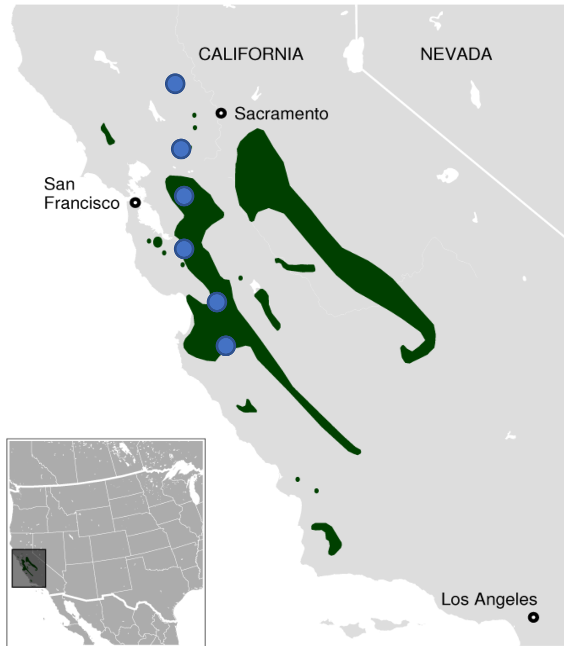
**Table 4.** Historic and contemporary outlier loci identified by BayeScan 2.0 when FDR = 0.01 for non-hybrid swarm populations. Loci identified were exclusively a subset of those identified when all six populaitons were included (See Table 3).

| Outlier SNP Index | | |
|---|---|---|
| **Historic** | **Both** | **Contemporary** |
| | 1129 | |
| 1489 | | |
| 3449 | | |
| 5324 | | |
| | 6662 | |
| | 11551 | |
| | | 11790 |
| | | 12269 |
| 12464 | | |
| | 12686 | |

**Table 5.** Number and proportion of contigs assembled to each linkage group in the assembly used to map sequences for this study.

| **Linkage Group** | **Number of Contigs** | **Percentage of Total** |
|---|---|---|
| 1 | 33 | 0.6% |
| 2 | 62 | 1.2% |
| 3 | 174 | 3.3% |
| 4 | 27 | 0.5% |
| 5 | 22 | 0.4% |
| 6 | 1529 | 29.2% |
| 7 | 203 | 3.9% |
| 8 | 890 | 17.0% |
| 9 | 206 | 3.9% |
| 10 | 21 | 0.4% |
| 11 | 1769 | 33.8% |
| 12 | 222 | 4.2% |
| OPA | 79 | 1.5% |
| **Total** | **5237** | |

APPENDIX B: FIGURES




| Locality | Sampling Years | | Number of Years Between Sampling |
|---|---|---|---|
| Dunnigan | 1992 | 2015 | 23 |
| Jepson-Olcott | 1986 | 2015 | 29 |
| Camino Diablo 2 | 1988 | 2015 | 27 |
| Grant 1 | 1991 | 2016 | 25 |
| Bluestone Quarry | 1999 | 2011 | 12 |
| JCL-Pond H | 1998 | 2010 | 12 |

**Figure 1.** Map of sampling locations in California shown with blue circles, overlaid on the range of CTS in green (IUCN). Site names are listed in the north-south order they appear on the map, including sampling years utilized for this study and the number of years between sampling.

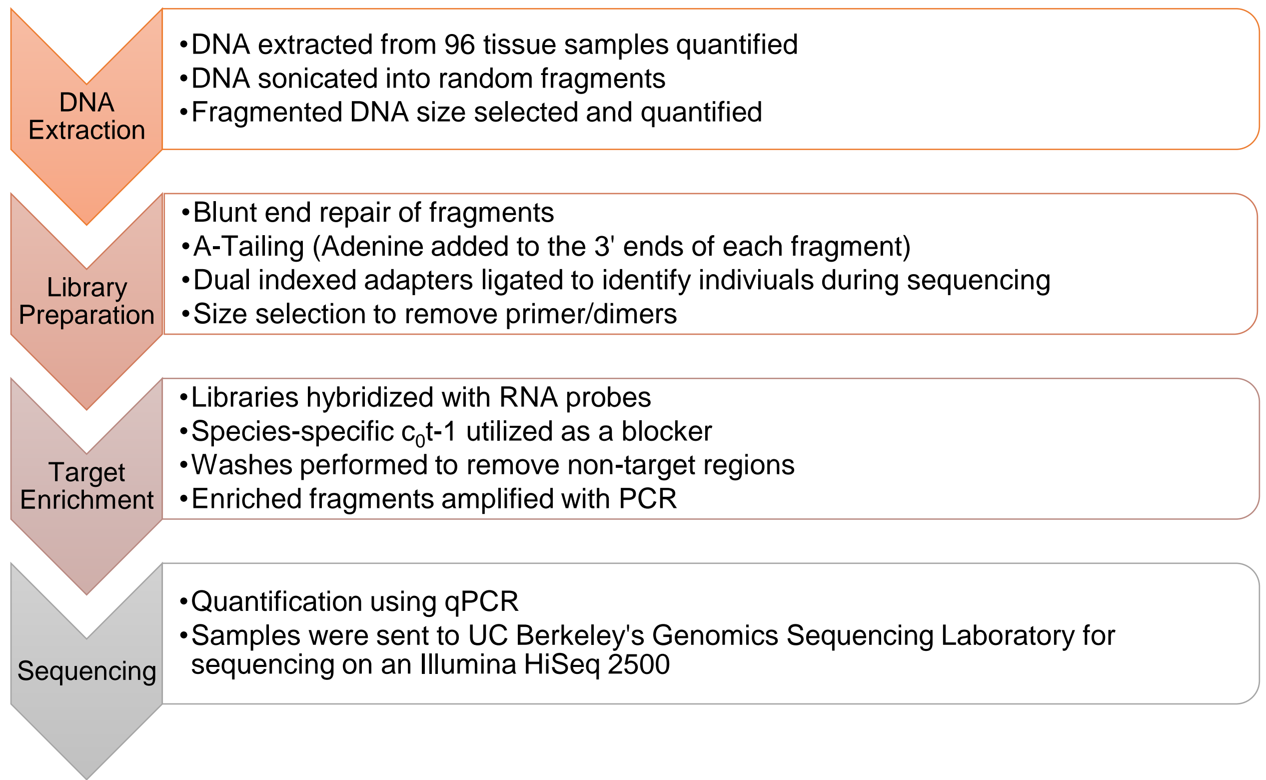



**Figure 2.** Workflow diagram of the genetic manipulation for each sample.

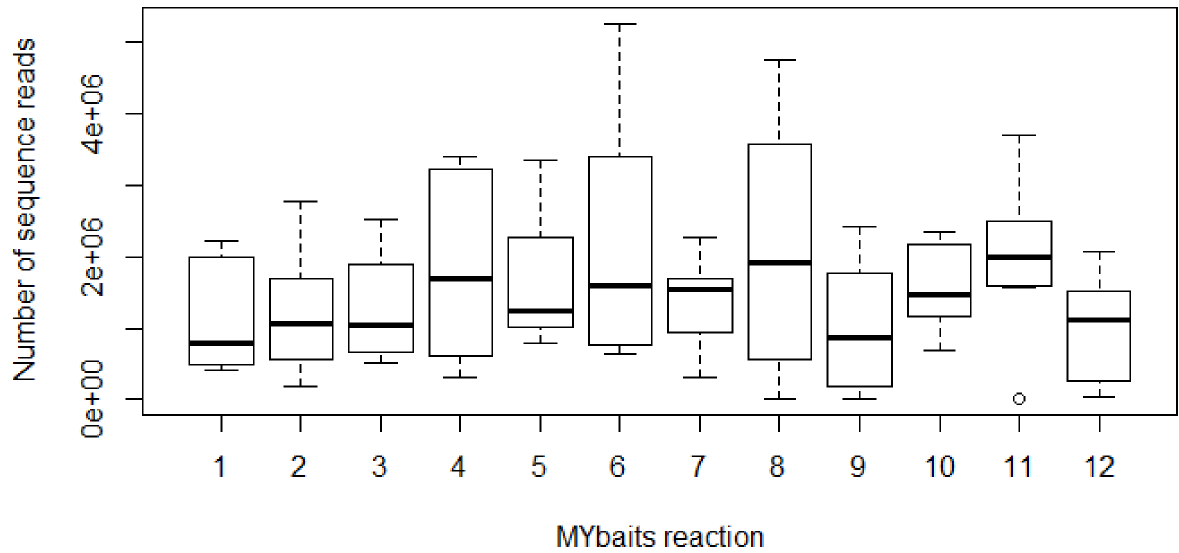



**Figure 3.** Boxplots showing the variation in individual capture (MYbaits) reactions on the number of sequence reads per individual. The median is indicated by the solid dark line in the center of the interquartile range (IQR) encompassing the 25th to 75th percentile of the data in the box. The whiskers extend from the IQR to represent extreme values in the dataset within ±1.5IQR , and outliers are shown as open circles. There was no significant impact of the capture reaction on the number of reads per individual ($F_{(11,84)}$ = 1.23, $P$ = 0.28).

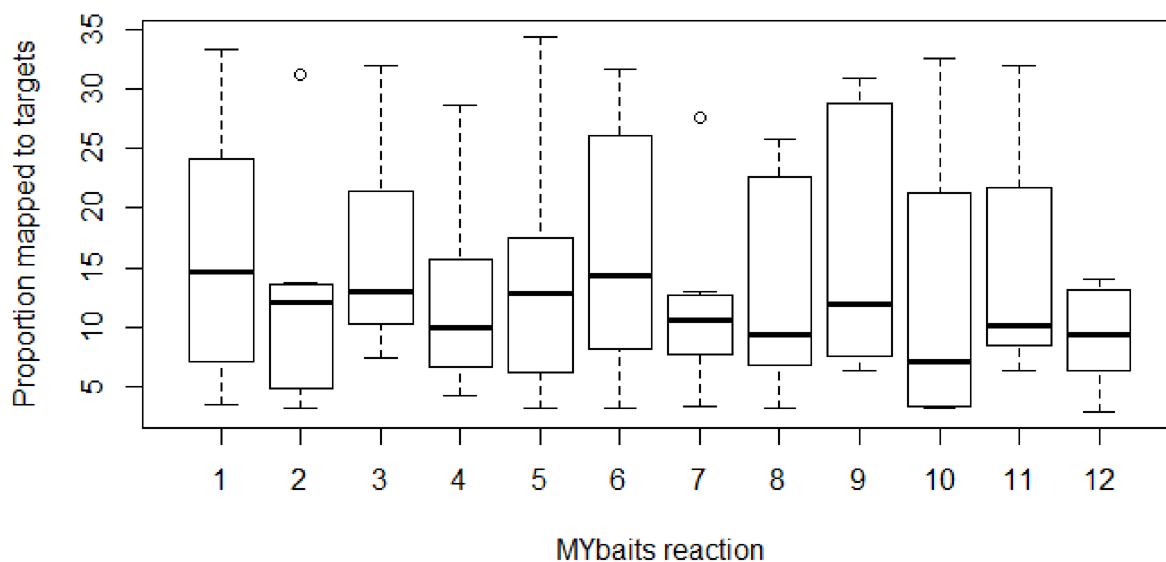



**Figure 4.** Boxplots showing the variation in individual capture (MYbaits) reactions on the proportion of reads mapped to target regions. The median is indicated by the solid dark line in the center of the IQR encompassing the 25$^{th}$ to 75$^{th}$ percentile of the data in the box. The whiskers extend from the IQR to represent extreme values in the dataset within ±1.5IQR , and outliers are shown as open circles. There was no significant difference in the mapping rate of sequences based upon the capture reaction ($F_{(11,84)}$ = 0.49, $P$ = 0.91).

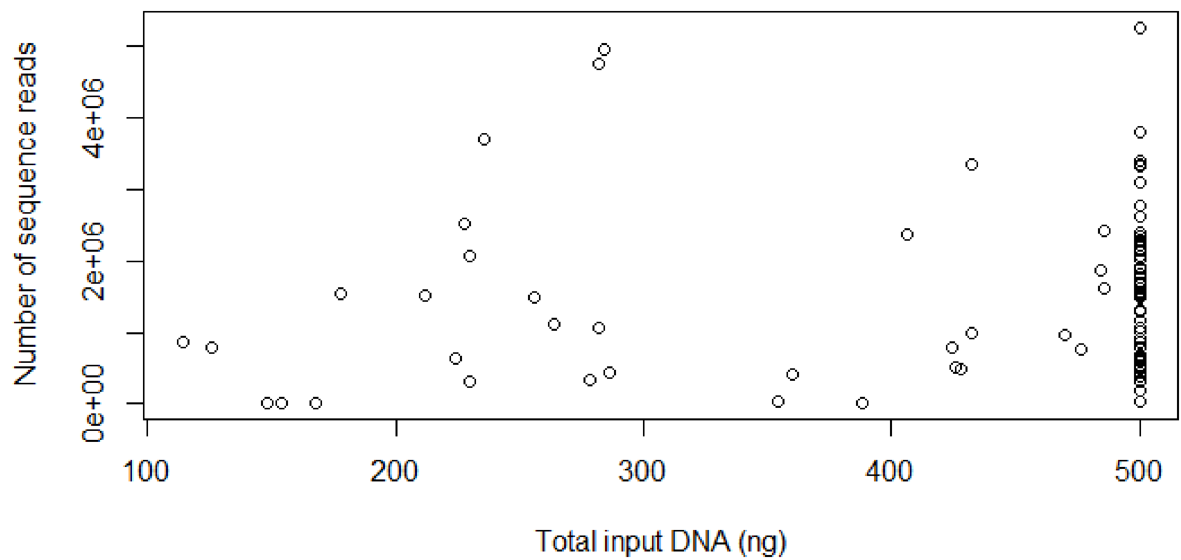



**Figure 5.** Scatterplot of total input DNA (ng) to library preparations to the number of sequence reads per individual. There was no significant relationship between total input DNA and the number of reads ($F_{(1,94)} = 1.60$, $P = 0.21$).

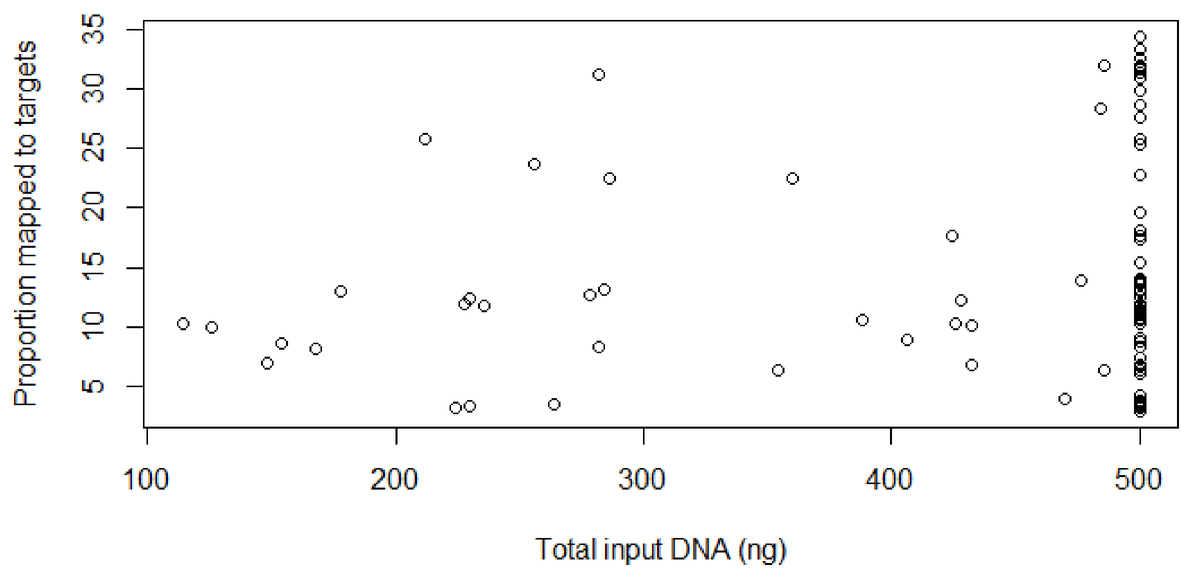



**Figure 6.** Scatterplot of total input DNA (ng) to the proportion of sequences mapped to target regions. There was no significant relationship between the two ($F_{(1,94)} = 0.95$, $P = 0.33$).

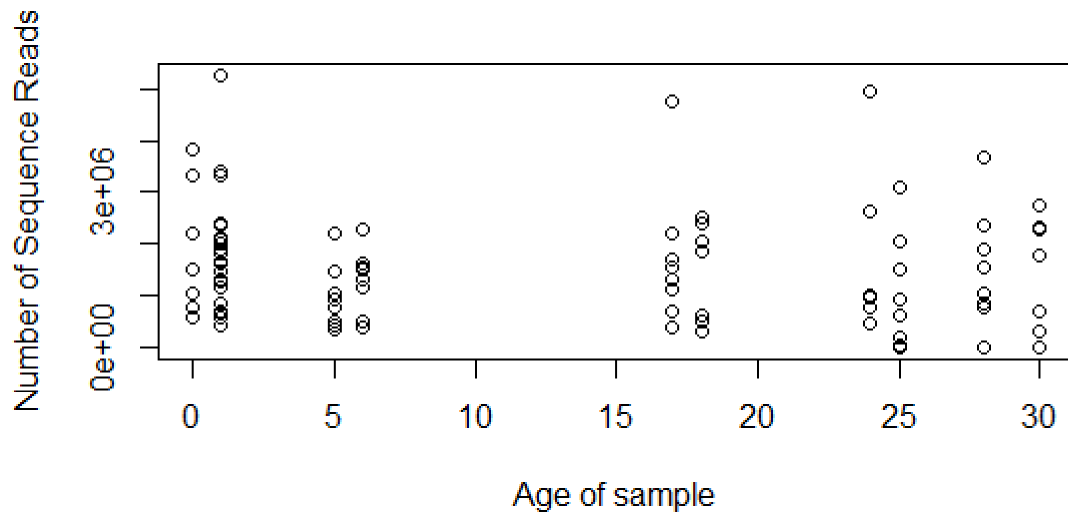



**Figure 7.** Scatterplot of the age of the sample in years at the time of DNA extraction to the number of sequence reads for each individual. There was no significant relationship between the two ($F_{(1,94)} = 1.85$, $P = 0.18$).

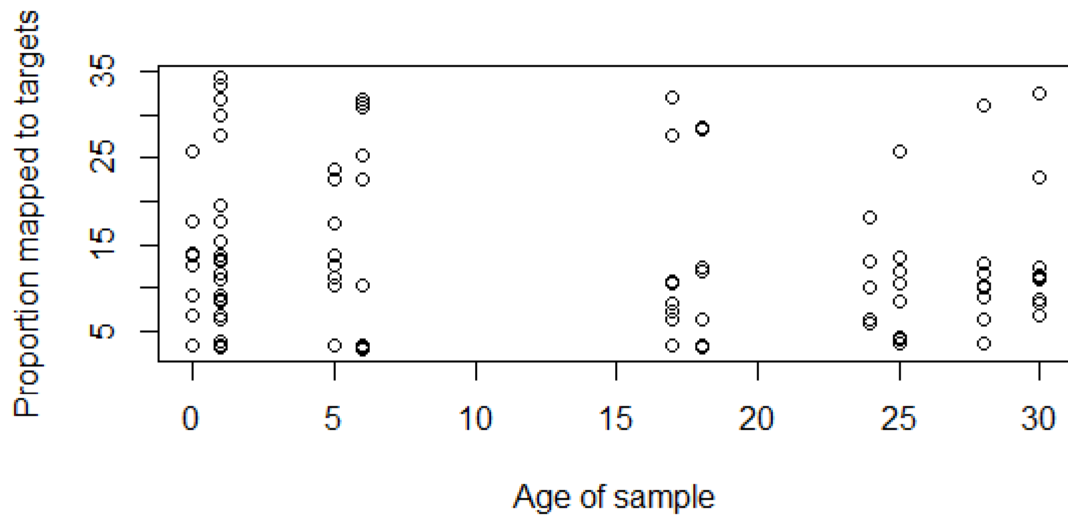



**Figure 8.** Scatterplot of the age of the sample at the time of DNA extraction to the proportion of sequence reads mapped to target areas of the genome. There was no significant relationship between the two ($F_{(1,94)} = 1.96$, $P = 0.16$).

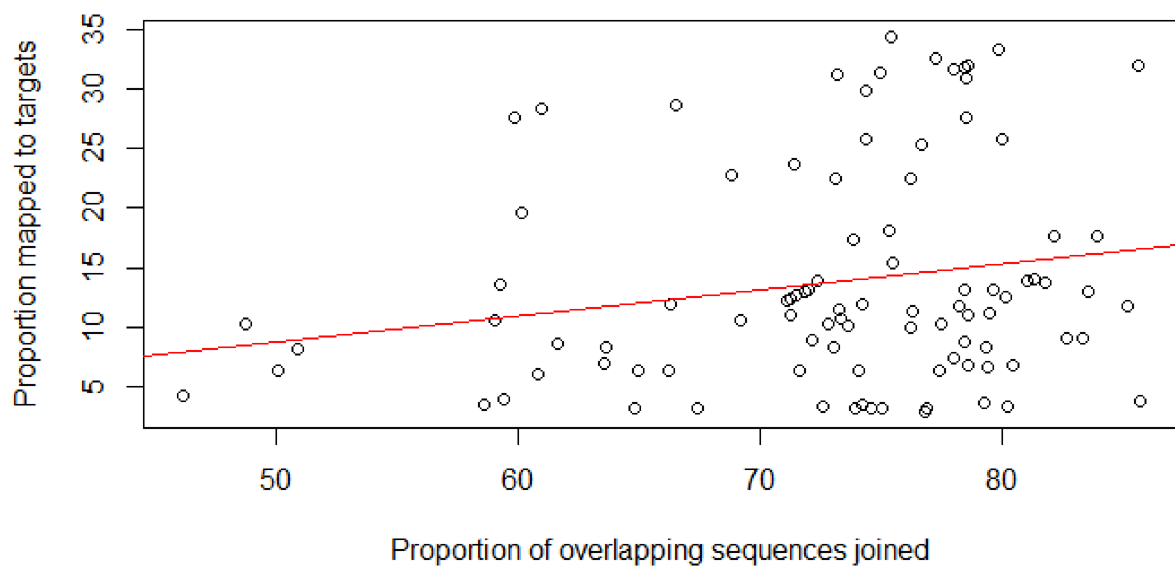



**Figure 9.** Scatterplot of the proportion of overlapping sequence reads joined using fastq-join to the proportion of sequences mapped to target regions with a best fit line shown in red. Greater proportion of joined sequences marginally affected the proportion of sequences mapped, though the predictability of this was low ($F_{(1,94)} = 3.73$, $P = 0.06$, $r^2 = 0.04$).

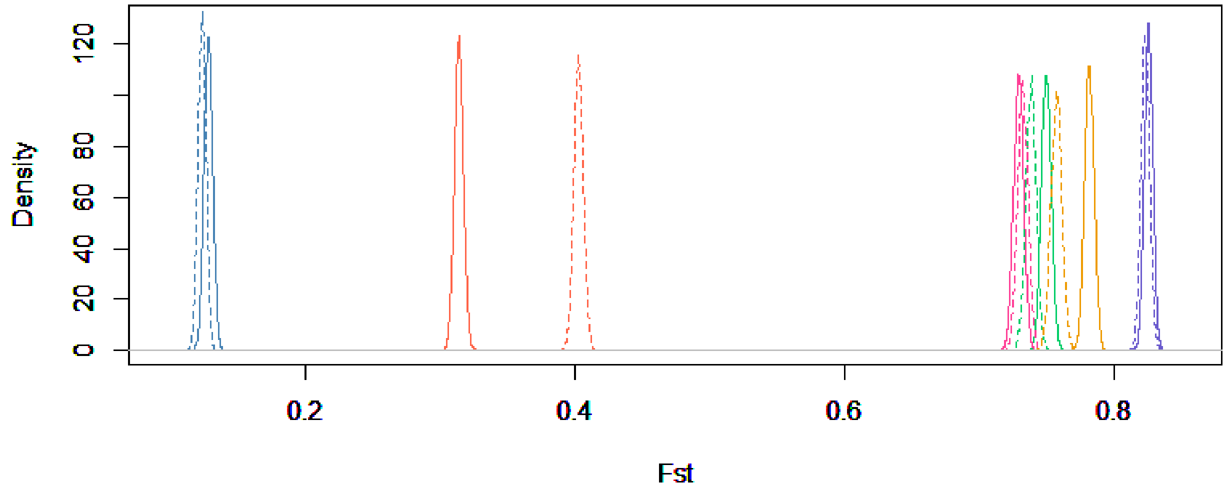



**Figure 10.** Posterior distributions of population-level $F_{ST}$ values for each sample site and time period calculated by BayeScan. Contemporary $F_{ST}$ values are shown as solid lines, historic $F_{ST}$ values are indicated by dashed lines, and populations are designated by color from left to right: Bluestone Quarry (blue), JCL-Pond H (red), Camino Diablo 2 (pink), Jepson-Olcott (green), Grant 1 (yellow), Dunnigan (purple).

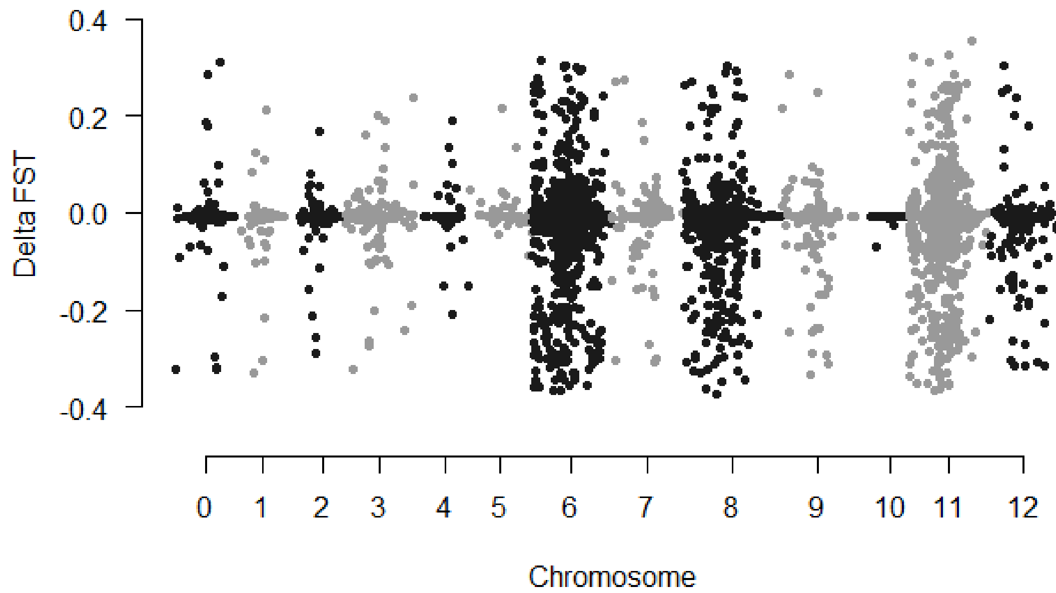



**Figure 11.** Manhattan plot of $\Delta F_{ST}$ values for all 13,647 loci arranged by chromosome. Arrangement of loci within chromosomes are arbitrary as exact locations in the genome were not determined. Loci found within chromosome zero are unmapped. The average $\Delta F_{ST}$ across all loci was -0.012.

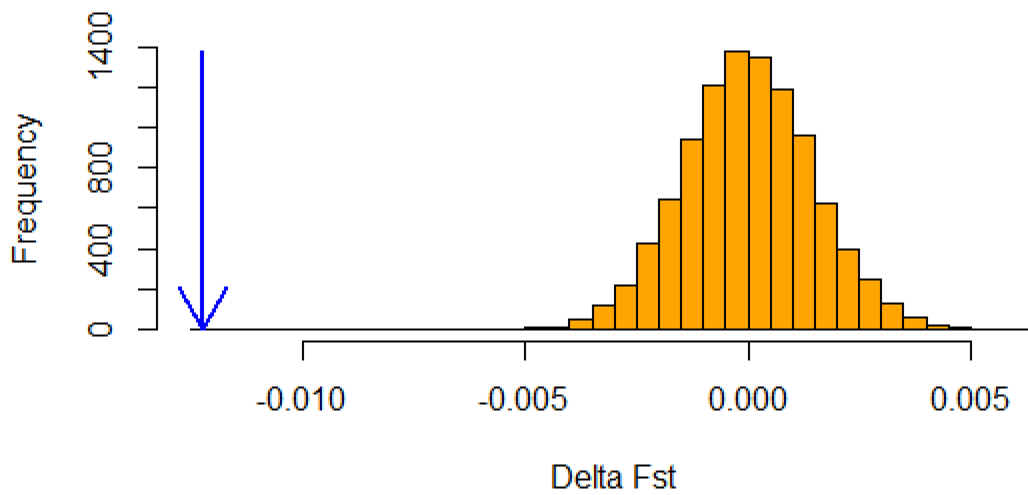



**Figure 12.** Histogram showing the differences in mean $F_{ST}$ of 10,000 random permutations of historic and contemporary $F_{ST}$ data. The mean of the randomized data was $5.70\times10^{-6}$. The blue arrow depicts the location of the average $\Delta F_{ST}$ for this dataset. A one-sample t-test reveals that the observed average $\Delta F_{ST}$ was significantly different than what would be expected at random ($P < 0.001$).

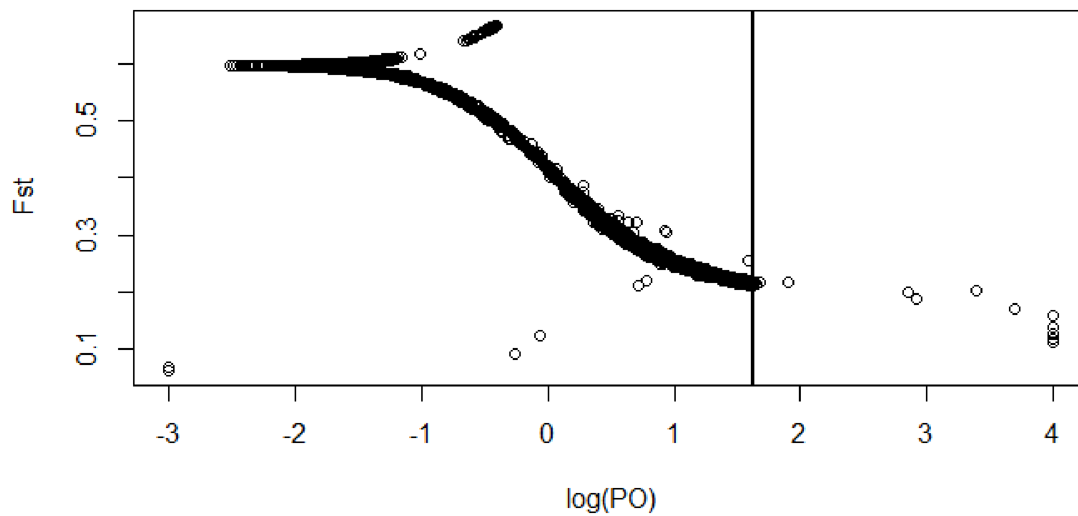



**Figure 13.** BayeScan plot for 13,647 loci of historic samples including all six study populations. $F_{ST}$ is plotted against log10 of the posterior odds (PO) that a locus is under selection. The vertical line represents the threshold false discovery rate (FDR) of 0.01. The 17 points to the right of the line were loci considered to be under selection.

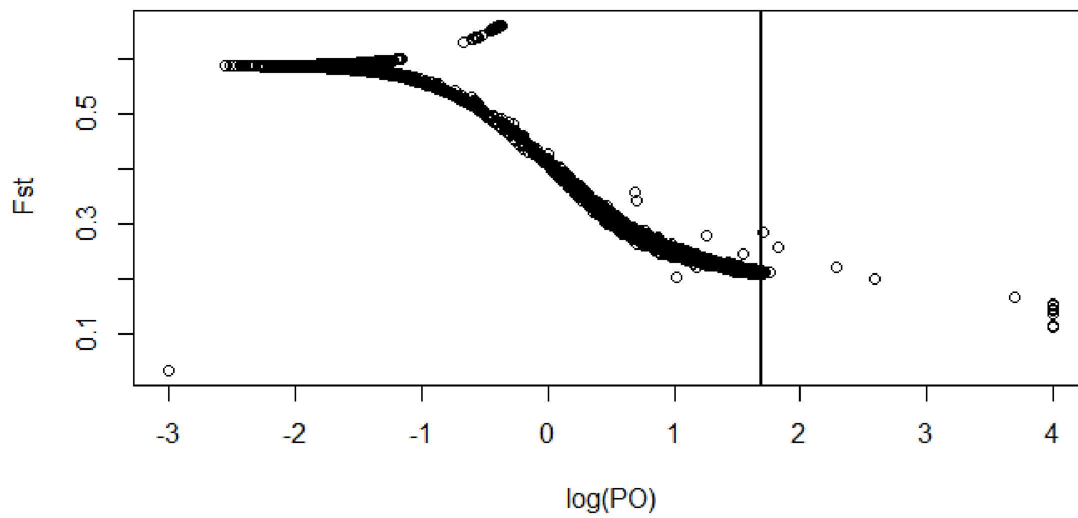



**Figure 14.** BayeScan plot for 13,647 loci of contemporary samples including all six study populations. $F_{ST}$ is plotted against log10 of PO that a locus is under selection. The vertical line represents the threshold FDR of 0.01. The 26 points to the right of the line were loci considered to be under selection.

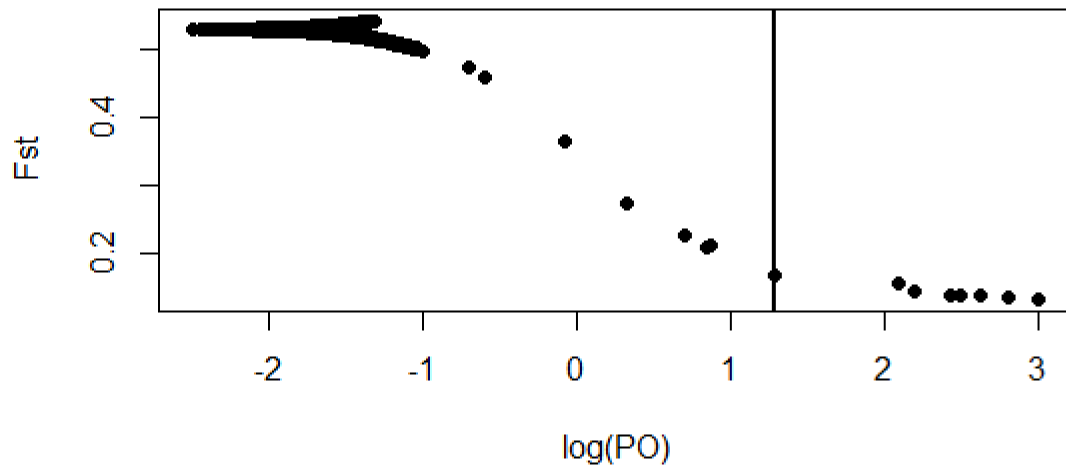



**Figure 15.** BayeScan plot for 13,647 loci of historic samples of populations excluding the hybrid swarm. $F_{ST}$ is plotted against log10 of PO that a locus is under selection. The vertical line represents the threshold FDR of 0.01. The 8 points to the right of the line were loci considered to be under selection.

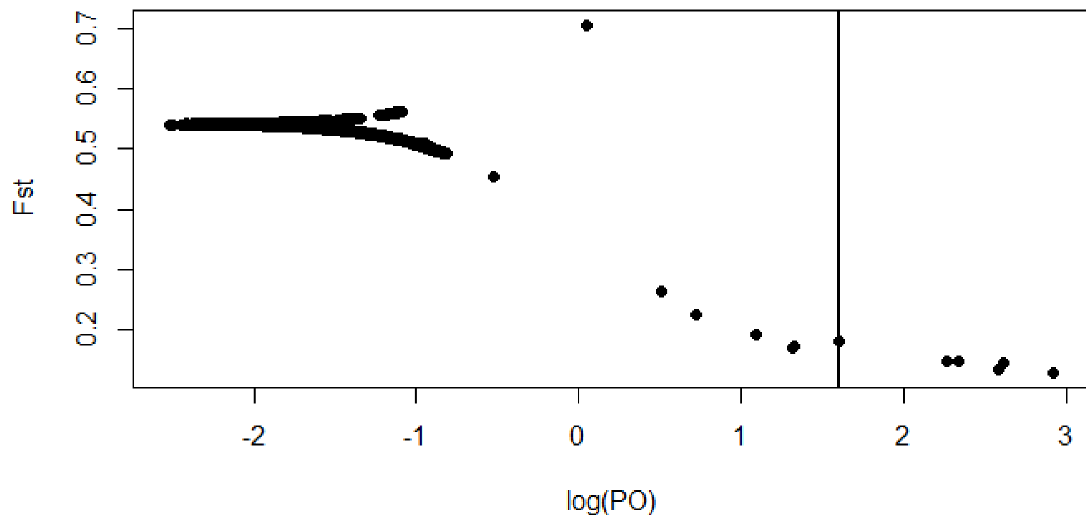



**Figure 16.** BayeScan plot for 13,647 loci of contemporary samples of populations excluding the hybrid swarm. $F_{ST}$ is plotted against log10 of PO that a locus is under selection. The vertical line represents the threshold FDR of 0.01. The 6 points to the right of the line were loci considered to be under selection.

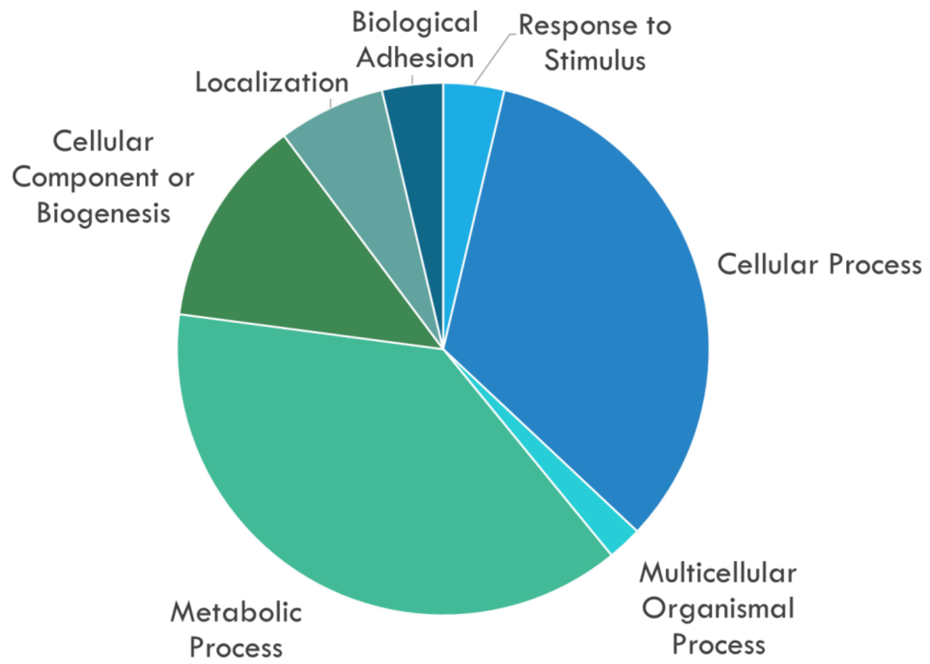



**Figure 17.** Biological processes of all outlier loci detected in both the historic and contemporary time periods categorized by GO annotations.

APPENDIX C: BIOINFORMATICS PROGRAMS USED

**Table S1.** List of programs and their functions in the bioinformatics pipeline used for management of sequence data.

| Program | Function in Sequence Management |
|---|---|
| fastQC | Visualized a general check on quality of raw sequence reads |
| Trimmomatic 0.32 | Removed Illumina and indexing adapter sections of sequence reads and trimmed sequences based on quality of base calls from sequencing |
| fastq-join | Merged overlapping paired-end reads |
| Burrows-Wheeler Aligner (BWA) | Mapped sequences for each individual to a combined assembly of targeted exons |
| SAMtools v. 1.3.1 | Sorted and merged SAM files into BAM files from the outputs of BWA |
| Picard Tools MarkDuplicates | Marked duplicate reads of a single DNA strand from sequencing |
| SAMtools flagstat | Counted flags in BAM files to obtain data on mapping rates and number of duplicate sequences |
| Picard Tools AddOrReplaceReadGroups | Added individual identifiers to BAM files prior to merging |
| BamTools | Merged BAM files of individuals from the same population and time period for variant calling |
| FreeBayes v. 1.1.0 | Called variants (SNPs) |
| VCFtools | Filtered and removed called SNPs based on missing data, mean depth, etc. |

APPENDIX D: ENVIRONMENTAL DATA

**Table S2.** Summary of environmental variables used in this study. Units are as follows: temperature ($^oC$), precipitation (cm), drought indices (Palmer Z-scores, and PDSI scores).

| Environmental Variable | Definition |
| --- | --- |
| ENV1 | April extreme maximum temperature |
| ENV2 | May extreme maximum temperature |
| ENV3 | June extreme maximum temperature |
| ENV4 | April extreme minimum temperature |
| ENV5 | May extreme minimum temperature |
| ENV6 | June extreme minimum temperature |
| ENV7 | April average maximum temperature |
| ENV8 | May average maximum temperature |
| ENV9 | June average maximum temperature |
| ENV10 | April average minimum temperature |
| ENV11 | May average minimum temperature |
| ENV12 | June average minimum temperature |
| ENV13 | April rain days above trace amounts |
| ENV14 | May rain days above trace amounts |
| ENV15 | June rain days above trace amounts |
| ENV16 | Total rain days above trace amounts April-June |
| ENV17 | April total precipitation |
| ENV18 | May total precipitation |
| ENV19 | June total precipitation |
| ENV20 | Total precipitation April-June |
| ENV21 | April Palmer Z-score |
| ENV22 | May Palmer Z-score |
| ENV23 | June Palmer Z-score |
| ENV24 | Self-calibrated Palmer Drought Severity Index |

**Table S3.** Environmental variable measures for each population and time period used for environmental correlations. Populations listed in north-south order, see Fig. 1.

| Site/Year | ENV1 | ENV2 | ENV3 | ENV4 | ENV5 | ENV6 | ENV7 | ENV8 | ENV9 | ENV10 | ENV11 | ENV12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Dunnigan** | | | | | | | | | | | | |
| 1992 | 32.78 | 37.22 | 41.11 | 6.11 | 4.44 | 11.11 | 25.20 | 32.06 | 21.85 | 10.63 | 13.47 | 14.71 |
| 2015 | 32.78 | 33.89 | 41.11 | 4.44 | 6.67 | 12.78 | 26.52 | 27.09 | 34.94 | 10.80 | 10.79 | 15.89 |
| **Jepson-Olcott** | | | | | | | | | | | | |
| 1986 | 33.89 | 36.67 | 41.67 | -1.11 | 2.22 | 7.78 | 23.72 | 29.02 | 32.13 | 4.63 | 6.81 | 9.87 |
| 2015 | 32.22 | 31.11 | 40.56 | 2.78 | 7.22 | 11.67 | 23.43 | 23.33 | 32.71 | 9.02 | 10.57 | 13.65 |
| **Camino Diablo 2** | | | | | | | | | | | | |
| 1988 | 32.78 | 36.67 | 39.44 | 3.33 | 3.33 | 2.78 | 24.31 | 26.05 | 31.08 | 8.20 | 8.75 | 13.49 |
| 2015 | 31.67 | 33.89 | 39.44 | 5.00 | 7.22 | 12.78 | 24.79 | 24.71 | 34.18 | 10.99 | 11.97 | 16.13 |
| **Grant 1** | | | | | | | | | | | | |
| 1991 | 28.33 | 30.56 | 32.78 | 2.78 | 5.00 | 6.67 | 20.18 | 22.98 | 25.52 | 5.94 | 8.12 | 9.68 |
| 2016 | 31.67 | 33.89 | 36.67 | 5.00 | 6.67 | 8.33 | 22.94 | 24.44 | 30.13 | 9.04 | 11.65 | 11.53 |
| **Bluestone Quarry** | | | | | | | | | | | | |
| 1999 | 33.33 | 28.33 | 37.22 | 2.22 | 5.00 | 6.11 | 21.19 | 23.55 | 27.43 | 6.97 | 8.99 | 10.65 |
| 2011 | 32.22 | 35.00 | 37.22 | 1.11 | 6.11 | 6.67 | 20.27 | 23.09 | 26.31 | 7.41 | 8.44 | 11.29 |
| **JCL-Pond H** | | | | | | | | | | | | |
| 1998 | 29.44 | 23.33 | 26.11 | 3.89 | 6.11 | 10.00 | 18.50 | 18.66 | 20.21 | 7.76 | 10.00 | 12.67 |
| 2010 | 25.00 | 25.00 | 30.56 | 3.89 | 4.44 | 9.44 | 18.06 | 18.64 | 20.87 | 7.11 | 8.83 | 11.65 |

**Table S3 Continued.** Environmental variable measures for each population and time period used for environmental correlations.

| Site/Year | ENV13 | ENV14 | ENV15 | ENV16 | ENV17 | ENV18 | ENV19 | ENV20 | ENV21 | ENV22 | ENV23 | ENV24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Dunnigan** | | | | | | | | | | | | |
| 1992 | 2.00 | 0.00 | 3.00 | 5.00 | 1.63 | 0.00 | 0.56 | 2.18 | -0.56 | -1.54 | -1.47 | -2.01 |
| 2015 | 4.00 | 1.00 | 0.00 | 5.00 | 3.07 | 0.10 | 0.00 | 3.18 | -0.57 | -2.20 | -2.20 | -4.00 |
| **Jepson-Olcott** | | | | | | | | | | | | |
| 1986 | 5.00 | 4.00 | 0.00 | 9.00 | 2.51 | 0.48 | 0.00 | 3.00 | 0.08 | 0.17 | 0.36 | 2.12 |
| 2015 | 2.00 | 0.00 | 1.00 | 3.00 | 3.20 | 0.00 | 0.41 | 3.61 | -0.89 | -2.15 | -1.85 | -3.35 |
| **Camino Diablo 2** | | | | | | | | | | | | |
| 1988 | 6.00 | 4.00 | 2.00 | 12.00 | 3.43 | 0.81 | 1.93 | 6.17 | -1.40 | -1.76 | -1.49 | -3.27 |
| 2015 | 3.00 | 4.00 | 1.00 | 8.00 | 1.32 | 1.40 | 0.38 | 3.10 | -1.36 | -1.14 | -0.50 | -3.13 |
| **Grant 1** | | | | | | | | | | | | |
| 1991 | 3.00 | 2.00 | 2.00 | 7.00 | 0.48 | 0.30 | 0.61 | 1.40 | -1.57 | -0.32 | 0.06 | -2.46 |
| 2016 | 4.00 | 3.00 | 0.00 | 7.00 | 3.00 | 0.53 | 0.00 | 3.53 | -0.30 | -0.43 | -0.63 | -1.60 |
| **Bluestone Quarry** | | | | | | | | | | | | |
| 1999 | 6.00 | 1.00 | 0.00 | 7.00 | 3.58 | 0.10 | 0.00 | 3.68 | 0.88 | 0.16 | 1.05 | 2.84 |
| 2011 | 7.00 | 5.00 | 3.00 | 15.00 | 0.66 | 2.24 | 0.71 | 3.61 | -1.45 | 1.93 | 2.81 | 2.89 |
| **JCL-Pond H** | | | | | | | | | | | | |
| 1998 | 10.00 | 10.00 | 4.00 | 24.00 | 5.11 | 4.85 | 0.28 | 10.24 | 2.95 | 7.42 | 4.27 | 6.19 |
| 2010 | 9.00 | 5.00 | 1.00 | 15.00 | 8.53 | 1.50 | 0.05 | 10.08 | 3.72 | 1.49 | 2.33 | 1.53 |