Variation among metabolic enzymes along a thermal gradient in a montane ectotherm

By

Kevin T. Roberts

A thesis submitted to

Sonoma State University

In partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

In

Biology

Committee Members

Dr. Nathan Rank, Chair

Dr. Elizabeth Dahlhoff

Dr. Sean Place

4/10/2016

Copyright 2016

By Kevin T. Roberts

Authorization for Reproduction of Master's Thesis

I grant permission for the print or digital reproduction of this thesis in its entirety, without further authorization from me, on the condition that the person or agency requesting reproduction absorb the cost and provide proper acknowledgement of authorship.

Kevin Roberts 6/1/2017

VARIATION AMONG METABOLIC ENZYMES ALONG A THERMAL GRADIENT IN A MONTANE ECTOTHERM

Thesis by
Kevin T. Roberts

Abstract
Many montane organisms live in fragmented populations that are especially vulnerable to climate change. The ability of populations to persist in the face of immanent environmental change depends partly on whether they possess genetic variation in their capacity to respond and adapt physiologically to altered environments. In the Sierra Nevada Mountains of California, the willow leaf beetle *Chrysomela aeneicollis* occurs at high elevations just below tree line (2400-3600 m). Individuals living in these high-elevation habitats experience highly variable environmental temperatures, ranging from below 0°C to above 30°C in a 24 hour period. Genetic markers reveal significant differentiation among montane drainages along a 40 km transect from Big Pine Creek to Rock Creek in this region. With the increasing accessibility of next-generation sequencing, whole-genome approaches to novel species are more readily available. I used next-generation sequencing to examine how genetic differentiation occurs in metabolic genes for three populations of willow beetles living along a latitudinal thermal gradient and three populations along an elevational (2800-3200m) gradient in a geographically-central drainage, Bishop Creek. My results revealed that single-nucleotide polymorphisms (SNPs) in exons of metabolic genes showed higher differentiation than SNPs in introns. This was not surprising, because differentiation at introns reflects population processes such as genetic drift and migration, whereas differentiation at exons reflects the same processes but may also involve natural selection. I also found that the glycolytic enzyme locus phosphoglucose isomerase (*Pgi*) and the FeS subunit of succinate dehydrogenase (*SdhB*) showed significantly higher genetic differentiation at non-synonymous SNPs than thirteen other examined loci, which show little or no variation. *Pgi* shows unusually high genetic differentiation across latitude, while *SdhB* shows high genetic differentiation among elevations in Bishop Creek. Lactate dehydrogenase and NADP dependent mitochondrial isocitrate dehydrogenase show similar levels of genetic differentiation as those observed for *Pgi* and *SdhB*, but data analyzed thus far suggest that it is driven by synonymous SNPs. Genetic variation found across geographic and elevational gradients suggests that both temperature and hypoxia may act as important selective pressures on insect populations inhabiting montane habitats. Combined, these pressures may challenge the ability of locally adapted populations to persist in the face of climate change.

Chair: Nathan Rank

MS Program: Biology Date: 6/1/2017

ACKNOWLEDGEMENTS

I would like to first thank Nathan Rank and Elizabeth Dahlhoff for years of guidance and for sparking my interest in research. I would also like to thank Chris Wheat for his many hours of guidance and for putting up with many lengthy, panicked emails. I would like to thank Sean Place for his advice along the way and patience with the creation of this document. I would also like to thank John Smiley for all of his advice, and all of the Sonoma State and Santa Clara students that helped over the years, especially Jordan Sayre and Jared Deyarmin. I would also like to thank all of the Sonoma State Biology staff, faculty, and students for the help over the past many years, especially Derek Girman, Dan Crocker, Jane Khudyakov, Mike Cohen, Rosemary Galten, Jen Aseth, Kandis Gilmore, Stephanie Thibault, Sarah Heidl, Kerry Winninger, Christine Windsor, Sage Wikman, Olivia Pipitone, Richard Regello, and Natalie Graham.

This research was funded by grants from the California State University Program for Education and Research in Biotechnology Entrepreneurial Joint Venture Grant to N. E. Rank and C. W. Wheat, and by a grant from the National Science Foundation (Divison of Environmental Biology 0844404 and 844406 to N. E. Rank and E. P. Dahlhoff.

Table of Contents

Chapter		Page
I.	Introduction	1
II.	Methods	8
III.	Results	14
IV.	Discussion	16
V.	References	24
VI.	Tables and Figures	31

Tables and Figures

Tables: Page

Table 1. Localities and numbers of samples used to construct sequence libraries. Drainages are listed from north to south. 31

Table 2. Genes of central metabolism showing genetic variation among eastern Sierra Nevada populations of the willow beetle *Chrysomela aeneicollis.* Data show number of exons, number of non-synonymous (NonSyn) and synonymous (Syn) single nucleotide polymorphisms (SNPs), and scaffold length of each contig. (number of base pairs). Enzyme name abbreviations used in Figures 3 and 4. 32

Table 3. Two-way factorial analysis of variance of genetic differentiation (F_{st}) between northern (Rock Creek) and southern (Big Pine Creek) populations of *Chrysomela aeneicollis*. F_{st} values were Box-Cox transformed before analysis using SAS PROC GLM. Data shown in Figure 1. 33

Table 4. Two-way factorial analysis of variance of genetic differentiation (F_{st}) between low and high elevation populations of *Chrysomela aeneicollis*. F_{st} values were Box-Cox transformed before analysis using SAS PROC GLM. Data shown in Figure 2. 34

Table 5. Two-way analysis of covariance for the relationship between distance and Fst value for pairs of populations along latitudinal and elevation gradients. Results based on a repeated measures analysis of exons using SAS PROC MIXED, with individual SNPs specified as subjects and a compound symmetry covariance structure. F_{st} and distance values were Box-Cox transformed before analysis. Data shown in Figure 3. 35

Table 6. Synonymous and Non Synonymous mutations in the most variable genes. 36

Figures:

Figure 1. Localities of beetles collected. Rock Creek (RC), Bishop Creek (BC), and Big Pine Creek (BP). 37

Figure 2. F_{st} before and after best fit Box Cox transformation. Distribution of F_{st} values before (top) and after (bottom) Box Cox transformation. 38

Figure 3. F_{st} values of SNPs found in variable introns and exons of metabolic genes for beetles from Big Pine Creek and Rock Creek. Genes coding for glycolytic enzymes shown in blue, TCA cycle in orange, and other dehydrogenases in red. Values shown are back-transformed least-square means of F_{st} values. Statistical analysis shown in Table 3. 39

Figure 4. F_{st} of SNPs found in variable introns and exons of metabolic genes for beetles from high and low elevation sites in Bishop Creek. 40

Legend code and values generated as described for Figure 1. Statistical analysis shown in Table 4.

Figure 5. Isolation by distance plots for metabolic enzyme loci that are variable among all populations. Data shown are based on Box-Cox transformed values for both axes. Lines represent predicted values based on pair-wise comparisons between five study populations. Genes are categorized by association with glycolysis (left), TCA cycle (center) or other pathways (right). Gene abbreviations are defined in Table 2. Statistical analysis is shown in Table 5. 41

Figure 6. Differences in allele frequency between low-, mid- and high-elevation populations for beetles in the Bishop Creek drainage. Data shown are least squares means for each gene by type of SNP (Non-synonymous/synonymous). Within the legend, genes are categorized by association with glycolysis (left), TCA cycle (center) or other pathways (right). Gene abbreviations are defined in Table 2. 42

INTRODUCTION

Organisms with wide geographic distributions often encounter variable environmental conditions across their ranges (Gaston, 2000). Local adaptation along environmental gradients can lead to genetic and phenotypic diversity over a large geographic range (Huxley, 1932; Endler, 1977). Elevation gradients experience thermal variation over shorter spatial scales than latitudinal gradients (Diaz and Bradley, 1997). Local adaptation along altitudinal gradients can thus promote high genetic differentiation over a small geographic range (Bonin et al., 2005). High elevation environments are highly prone to highly variable diurnal and seasonal temperatures, which require species inhabiting these environments to possess the capacity to cope with thermal stress across their lifecycle (Diaz and Bradley, 1997). Cellular respiration powers the stress response, making thermal tolerance partially dependent on cellular metabolism. An organism's physiological response to this stress is energetically costly, and these energetic requirements must be met for survival and reproduction to occur (Somero, 2002; Silbermann and Tater, 2007). Since insects are ectotherms, many physiological processes are reliant on thermal conditions (Wilmer, 1991). Temperature profoundly affects cellular metabolism, with colder temperatures slowing metabolic rate by reducing reaction rates and metabolic flux (Huey and Stevenson, 1979; Clarke, 1993). Thus, examining local adaption of ectotherm populations along elevation gradients provides a powerful tool to study how organisms adapt to differing thermal environments.

Populations can respond to changing environments by shifting their range

to habitats with more moderate climates. This has been observed for birds, butterflies, and trees, which all have been documented making pole-ward shifts in response to warming climate (Parmesan and Yohe, 2003; Chen et al., 2011). Many range-restricted species are faced with loss of suitable habitat as they shift, leaving them with little opportunity to disperse to environments where they can survive. In high elevation environments, species may shift to higher elevation habitats to cope with new thermal regimes (Wilson et. al, 2005; Parmesan, 2006). High elevation environments have longer cold seasons, which can cause organisms inhabiting them to have shorter development windows or lower fecundity (Parmesan, 2006). They could also be exposed to more drastic levels of thermal variation than to which they can physiologically cope (Thomas et al., 2006). Along with thermal stress, high elevation habitats are associated with hypoxic conditions, higher evaporation rates, and high solar radiance, which can pose additional challenges to survival (Buckley et al., 2013; Tracy, 1976). Hypoxic conditions inhibit development and lead to decreased adult body mass (Callier et al., 2013). Increased solar radiance can cause body temperature to increase causing additional thermal stress, and altered behavioral thermal regulation (Sunday et al., 2014). Given these additional environmental stresses, range shifts to higher elevation are not always desirable or even possible.

Metabolic enzymes have long been the focus of evolutionary studies, partly owing to the ease of scoring electromorphs using allozyme electrophoresis (Marden, 2013a). Proper functioning of proteins of basic cellular metabolism is critical for growth, movement, and recovery from stress; however, these

processes are very sensitive to changes in the environment (Tolmanek and Zuzow, 2010). For example, most enzymes have a relatively narrow window of optimal function; a small change in temperature can cause the binding constant (K_m) or maximal rate (V_{max}) to alter substantially (Somero, 1995), resulting in reduced performance and fitness (Watt 1977; Watt et al. 1983; Dahlhoff and Rank 2007). Enzymes of many ectotherms have thermal optima that correspond to native environmental temperature ranges (Hochachka and Somero, 2002). Decreased function of metabolic enzymes can lead to decreased organismal performance and reproductive success. There are gradients in frequency of alleles at polymorphic metabolic loci along environmental gradients in a large number of animals, including deer mice (Cheviron et al., 2012), mummichog fishes (Whitehead and Crawford, 2006), barnacles (Veliz et al., 2006), crickets (Huestis et al., 2009), and mussels (Lockwood and Somero, 2012). Metabolic enzyme performance is also important for the cellular stress response, which is energetically costly from production of chaperone proteins (Krebs and Loeschcke, 1994; Svensson et al., 1998). In addition, many metabolic enzymes or their products function as cellular signals or play secondary roles in other enzymatic pathways (Marden, 2013a). Moonlighting functions can range from enolase's involvement in cold-acclimation to aconitase's role in iron homeostasis (Lee et al., 2002; Volz, 2008). It is for these reasons that metabolic genes are ideal target genes to examine thermal adaptation.

The willow beetle *Chrysomela aenecollis* (family Chrysomelidae) is well suited to study natural selection based on environmental gradients in populations

of ectotherms. This beetle's range extends across western North America (Brown, 1956; Dellicour et al., 2014). In California, *C. aeneicollis* reaches the southern-most extent of its range in the Eastern Sierra Nevada Mountains, where it is restricted to high elevation environments (Smiley and Rank, 1986). Populations in Rock Creek (RC), Bishop Creek (BC), and Big Pine Creek (BP) have been the focus of long-term studies of local temperature adaptation (Dahlhoff et al., 2008; Dahlhoff and Rank, 2000; Rank, 1992). Adult beetles overwinter under host plants and emerge in early June to mate and lay eggs. Larvae hatch and develop on the same host plant until the beginning of winter, when new adults enter diapause. Populations are fragmented and microclimate varies significantly among localities (Dahlhoff and Rank , 2007). Not surprisingly, air temperature declines with increasing elevation (Boychuck et al., 2015).

Previous work suggests that the glycolytic enzyme phosphoglucose isomerase (*Pgi*) experiences temperature selection in these willow beetle populations (Rank, 1992). This enzyme was found to have much higher genetic differentiation from BP to RC than other allozyme loci (Rank, 1992). Work conducted for the past two decades has demonstrated that *Pgi* genotype has been linked to physiological responses to stress, thermal tolerance, and performance, and fitness traits (Neargarder et al. 2003; Dahlhoff et al., 2007; Rank et al., 2007; Dick et al., 2013), and PGI allozymes have different functional properties (Dahlhoff and Rank 2000). Allele frequency variation at *Pgi* corresponds with a latitudinal air temperature gradient among drainages (Rank and Dahlhoff, 2002). Within Bishop Creek, an allele frequency cline in the

opposite direction exists. Warm adapted southern *Pgi* alleles are found in greater abundance at high elevations, while cold adapted northern *Pgi* alleles are more predominant at low elevations (Rank, 1992).

The mitochondrial gene cytochrome oxidase II (*COII*) follows a similar pattern as *Pgi* along the latitudinal gradient and within Bishop Creek. Recent studies have indicated that mt*COII* and *Pgi* genotypes jointly influence larval development rate, adult running speed and larval cold tolerance in *C. aenecollis* (Heidl, 2013; Sayre, 2014). There is an increasing body of research that suggests that co-adaptation of mitochondrial and nuclear genes is important, and that mismatches between the two can lead to reduced fitness (Rand, 2001; Ellison and Burton, 2006; Gompart et al., 2008; Dowling et al., 2008; Montooth et al., 2010; Hoekstra et al., 2013; Burton et al., 2013). The *C. aeneicollis* populations in eastern California offer an excellent case to gain greater understanding of how the two genomes interact to mold thermal adaptation.

To date little is known about sequence variation that underlies the genetic markers used in earlier studies of *C. aeneicollis*. Most work has focused on five polymorphic proteins screened using allozyme electrophoresis. This method allows for electromorph genotyping of many individuals at once and allows for multiple alleles to be detected in a population. However, it misses underlying genetic variation that does not translate into electromorph differences. This includes variation at introns within each gene, and synonymous variation in the gene that does not affect amino acid sequence but might affect gene expression in other ways. Prior research on genetic differentiation in *C. aeneicollis* using five

microsatellite markers and one mitochondrial locus genotyped by the restriction fragment length polymorphism (RFLP) method has been done (unpublished data), but as microsatellites are neutral markers, they are limited in their functionality in studies of adaptation. Furthermore, RFLP is time consuming and must be run on each individual for each gene, limiting sample size and size of the gene set used. No work in this system has been done on how entire metabolic pathways vary geographically.

Often sequence level studies reveal the underlying basis for phenotypic variation. The use of Next-generation sequencing allows for generation of massive amounts of DNA sequence data (Ellegren and Sheldon, 2009). Sequencing pooled individuals from a population (Pool-Seq) provides an accurate representation of population allele frequencies (Kofler et al., 2011b). Pool-Seq is a useful approach when dealing with large samples sizes or dealing with population differentiation without reducing the genome coverage (Schlotterer et al., 2014). Pool-seq specific programs are making it increasingly accessible to perform using non-model organisms without a loss of accuracy (Ramprasad et al., 2016).

With my thesis work, I used Pool-Seq to answer the following questions:

1) Do SNPs in exons of metabolic genes show different patterns of genetic differentiation than SNPs in introns, and does this pattern differ along environmental gradients? Comparing SNPs in exons of genes to SNPs in introns will allow me to compare loci potentially under selection to SNPs that are in the same gene, but have weaker selective forces acting on them, because

they are spliced out. **2) Do exons of metabolic genes show evidence of environmental effects on geographical differentiation? For example, does isolation by distance vary among genes, and do non-synonymous SNPs differ from synonymous SNPs with respect to genetic differentiation?** Isolation by distance shows how genetic differentiation between populations changes over distance. It is a powerful tool in the comparison of multiple populations, and allows for comparison of patterns of differentiation over distance among loci. **3) How does elevation influence allele frequency within populations in Bishop Creek? Does elevation influence allele frequencies of non-synonymous SNPs differently than synonymous SNPs?** *Pgi* and COII exhibit a cline in Bishop Creek, looking at allele frequencies across elevation will show if other genes follow the same pattern. Since synonymous SNPs do not change the amino acid they code for, they are assumed to have little function on enzyme function; so by comparing differentiation of non-synonymous SNPs to synonymous SNPs we can examine where there appear to be selective pressures on the non-synonymous sites.

METHODS

Sample Collection

Free-living adult beetles were collected in the Eastern Sierra Nevada during the summer of 2013. Individuals of both sexes were sampled from within three drainages across the latitudinal gradient, Big Pine Creek, Bishop Creek, and Rock Creek (Table 1; Figure 1). Individuals within Bishop Creek were collected from three different localities at varying elevations(Table 1; Figure 1). After collection, individuals were taken to the White Mountain Research Center (Bishop, CA) and frozen at -80°C until DNA extraction.

DNA preparation

Samples were flash frozen using liquid nitrogen and homogenized using a mortar and pestle. Once powderized, 400ug RNAase was added to each sample and samples were incubated at 56°C for 12 hours. DNA was extracted using Qiagen Dneasy Plant Mini Kits (Qiagen, Valencia, CA) using a standard plant tissue protocol. Extracted DNA was quantified using a Qubit Fluorimeter (Life Technologies, Grand Island, NY) with hsDNA kits. Samples were then combined at equimolar concentration (> 25 ng/µL) into five libraries, Big Pine Creek, Bishop Creek low elevation (La Hupp), BC mid elevation (Pipeline), BC high elevation (High Stream Crossing), and Rock Creek. Each library contained greater than 3µg total low fragmented DNA. Library preparation (200-300bp insertion size), barcoding, and full genome sequencing (Illumina HiSeq2000, 100PE) was performed by The Beijing Genomics Institute (Sacramento, CA). The genome of *C. aeneicollis* is estimated to be around 600Mb, using 100PE HiSeq2000

sequencing we would expect roughly 50 times coverage for the entire genome per lane. All sequencing was done in six lanes, resulting in 180 billion base pairs of data from 240 beetle genomes.

Data Processing

After sequencing, several stages of file preparation were implemented before data analysis. PCR duplication removal was completed using clonefilter (Catchen et al., 2013). Adapters were removed in the forward and reverse direction using BBduk2, which was also used to remove low confidence base pairs with a phred quality score below 10. A de novo assembly was created (kmer = 64, bubble size = 2000) using Pool-Seq data from 50 individuals collected in 2011 from High Stream Crossing and assembled using CLC genomics workbench (CLC Bio-Qiagen, Aarhus, Denmark). A gene set was created using orthologs from *Tribolium castaneum,* using genes involved in glycolysis/gluconeogenesis, the citric acid cycle, oxidative phosphorylation, pentose phosphate pathway, fructose and mannose metabolism, heat shock response, and hypoxia-inducible factor pathway obtained from Kyoto Encyclopedia of Genes and Genomes (www.kegg.jp). Contigs containing a sequence that matched a Tribolium ortholog 90% or greater were isolated using MESPA (Ramprasad et al., 2016). Once contigs containing our genes of interest were isolated, clean reads from our libraries were mapped using Next-Gen Mapper (Sedlazeck et al., 2013). Once the five pools were mapped to contigs containing their orthologs, sequences of genes involved in glycolysis and TCA cycle were identified and manually examined using Integrative Genomics Viewer

(Robinson et al., 2011; Thorvaldsdottir et al., 2012). Annotated SNPs and allele frequencies were generated for all SNPs in introns and exons with coverage between 35X and 150X using Popoolation (Kofler et al., 2011a). Pair-wise comparisons among populations and F_{st}'s were generated using Popoolation2 for all SNPs previously identified (Kofler et al., 2011b).

Analysis

Preparation of SNP and F_{st} files and data transformations

Files containing all SNPs and their frequencies were exported from Popoolation. These files were then imported into JMP 11 (SAS Institute, Cary, NC) and all SNPs in 30 contigs containing genes of interest were extracted. All variable sites that appeared between start and end of the gene within the three Bishop Creek populations were then isolated. These sites were used as reference SNPs for all further analysis. Raw pairwise F_{st} files were exported from Popoolation2 for each of ten comparisons among 5 study populations. These F_{st} output files were then imported into JMP 11 and merged one by one with the reference SNP file, then concatenated together. Once in JMP, F_{st} values were transformed using the Box Cox method to meet assumptions of parametric analysis (Figure 2). To calculate the best transformation, a small number (0.0001) was added to all values in order to remove zero values, then a statistical model was fit in the 'Fit Model' platform in JMP (using the expected mean-squares option) and selecting “Factor Profiling” option in the output window. The transformed values were negative, so 0.1 was added to make them positive. The final formula for transformed F_{st} was $\{[\ln(F_{st} + 0.0001) * 0.008846] + 0.1\}$.

Pairwise distances among populations were calculated from geographic position using the online application GDMG (http://biodiversityinformatics.amnh.org/open_source/gdmg/), and also transformed using a best-fit Box Cox transformation [(Distm $^{0.2}$ - 1) / 0.0001927)].

How do genes vary between coding and non-coding regions?

In JMP, the file containing all F_{st} values and the file with reference SNPs were merged, returning all F_{st} values for each SNP in each intron and exon of targeted genes. The number of intron SNPs was too large for analysis, thus a subset of 10% of SNPs in introns was randomly selected to yield similar sample sizes as exon SNPs before merging the exon and intron files. Genes with missing F_{st} values for comparisons between Big Pine Creek and Rock Creek were removed for further analysis. Transformed F_{st} values were then analyzed in SAS 9.4 (SAS Institute, Cary, NC) using a two way factorial ANOVA with gene, gene region (intron/exon) and the interaction term. To produce graphs, least-square mean F_{st} values were back transformed (Figures 2 and 3). A similar approach was used to compare transformed F_{st} values from low and high elevation Bishop Creek populations.

Does isolation by distance vary by gene?

For this analysis, we omitted SNPs in introns and relied on SNPs that included at least one comparison between Big Pine or Rock Creek and any Bishop Creek population. This resulted in 22 genes with ten F_{st} comparisons among populations and one with only six comparisons. Because of these multiple F_{st} values for each SNP, a repeated-measures ANCOVA using the restricted

maximum likelihood option in SAS PROC MIXED was employed. A repeated measures model accounts for lack of independence of F_{st} values calculated from the same pairs of populations. An ANCOVA was used because distance was fundamental to the research question. Multiple covariance structures were tested using the maximum-likelihood option in PROC MIXED, and Akaike Information Criterion (AIC) values revealed that compound symmetry was the best one. Multiple models were run, with full factorial design, subsets with each grouping factor alone, different combinations of interactions, and all combinations of covariate by grouping factor; the best model was picked using AIC. The final model included gene and SNP type (synonymous or non-synonymous), as grouping factors and distance as a covariate. This approach to analyzing isolation by distance helps avoid Type I error associated with traditional isolation by distance analysis (Meirmans, 2012). In order to plot data, the model was rerun with distance as a categorical variable to calculate least squares means. These means were then plotted against distance and slope values calculated to represent the relationship between F_{st} and distance for each gene (see figures).

How does elevation affect metabolic genes within a population?

Raw files containing allele frequencies were taken from Popoolation for five study populations. All files were merged together in JMP, and then merged with the file containing all reference SNPs. SNPs from exons were then extracted and allele frequencies were calculated for each locus. Allele frequencies were calculated for the most common allele in Big Pine Creek, if Big Pine Creek data was missing, then the high elevation population from Bishop Creek was used.

Allele frequencies were plotted using least squares means calculated using a PROC MIXED model in SAS in which allele frequency was analyzed with gene, population, and SNP type as grouping factors and no interaction terms.

RESULTS

Genes recovered for analysis

Of all genes involved in glycolysis and TCA cycle, 23 genes were found in contigs of our assembly (Table 2). Of those 23 genes, 10 were in glycolysis and 13 were in the TCA cycle. Seven other biologically relevant genes from various metabolic pathways were added to the gene list, bringing our total to 30 genes. Of these genes, 14 were subunits of larger complexes, and all genes had introns that varied in length. Only 19 of 30 genes had non-synonymous SNPs within exons.

How do genes vary between coding and non-coding regions?

Levels of genetic differentiation (F_{st}) between Big Pine Creek and Rock Creek depend on the gene (Figure 3, Table 3). Introns exhibited much less genetic differentiation between localities than exons (Figure 3, Table 3). *Pgi* showed the highest differentiation in introns, while *Pgi* and NADP-dependent mitochondrial isocitrate dehydrogenase (*Idh2*) showed the highest differentiation at exons (Figure 3, Table 3).

As expected, genetic differentiation was much lower within Bishop Creek than between Big Pine and Rock Creek. SNPs in exons tended to show more genetic differentiation than SNPs in introns. Levels of genetic divergence among populations varied among genes. Succinate dehydrogenase FeS (*SdhB*) and synonymous SNPs in lactate dehydrogenase (*Ldh*) exhibited the highest amount of genetic differentiation within exons (Figure 4, Table 4).

Does isolation by distance vary by gene?

Every gene had increasing differentiation with increasing distance, but the amount depended on the gene (Figure 5, Table 5). Genetic differentiation of non-synonymous SNPs increased more with distance than synonymous SNPs (Table 5). Genetic differentiation depends on the distance between two populations, with *Pgi* having the highest genetic differentiation over distance of glycolytic genes. Of TCA cycle genes, *Idh2* had the largest genetic differentiation between distant populations, while *SdhB* had the highest genetic differentiation between close populations (Figure 5, Table 5).

How does elevation affect metabolic genes within a population?

Allele frequencies of SNPs within genes in Bishop Creek showed that most non-synonymous allele frequencies were high throughout the elevation gradient. Non-synonymous SNPs in *Pgi* and *SdhB* stand out for having much lower allele frequencies than other genes, and allele frequency of each gene varies monotonically with elevation (Figure 5). Synonymous SNPs vary more in allele frequency along the elevation gradient than non-synonymous SNPs (Figure 5). Within Bishop Creek, allele frequency of *Pgi* looks most like Big Pine Creek at high elevation. *SdhB* follows an inverse pattern, the allele frequency at low elevation looks most like Big Pine Creek (Figure 5). Synonymous variation at lactate dehydrogenase was similar to that observed at non-synonymous sites at *SdhB* (Table 6).

DISCUSSION

In this study I found high levels of genetic variation in both introns and exons of metabolic loci in eastern Sierra Nevada populations of *C. aeneicollis*. My findings support prior work that found populations show moderate levels of genetic divergence and unusually high divergence at several loci (Rank, 1992; Rank and Dahlhoff, 2002). The genetic differentiation between populations in *Pgi* matches prior studies, but this is the first that employs DNA sequence level analysis at this locus and quantifies the genetic basis of the *Pgi* polymorphism in *C. aeneicollis* (Rank 1992; Rank and Dahlhoff, 2002). Two genes exhibiting high genetic differentiation at non-synonymous sites, *Pgi* and *SdhB*, showed a strong cline across elevation in Bishop Creek. One of my most exciting findings is that *SdhB* displays a strong cline along the Bishop Creek elevation gradient. Any gene that shows such a strong cline among elevations has significant potential to be important for coping with variation in oxygen supply. This is particularly exciting because Sdh is involved in hypoxia signaling, discussed below. The pattern found in *SdhB* shows that there can be strong allele frequency shifts at small spatial scales and has large implications into the amount of local adaptation that these insects may exhibit. Two other genes exhibited high genetic differentiation in SNPs that are synonymous, *Idh2* and *Ldh*. These two genes showed interesting patterns of differentiation, but more work is needed to better understand why they would do so.

Surprisingly few population genetic studies have examined genetic differentiation of SNPs in introns versus exons. Introns are frequently used as

markers for phylogenetic studies due to the commonly held assumption that they are selectively neutral and accrue mutations at a constant pace (Palumbi and Baker, 1994, Johansson et al., 2013). Comparing the variation in introns to the variation in exons will be informative on the amount of selection acting on coding regions of a gene. In a study of *Drosophila melanogaster,* it was found that the amount of selection acting on SNPs in introns is similar to synonymous SNPs, and the best markers to use were SNPs in introns 8-30bp from an exon (Parsch et al., 2010). Introns can have important functions in regulatory sites involved in alternative splicing, causing them to be positively selected for (Lynch, 2002). Alternately, introns may show signatures of selection due to linkage with exons, known as hitchhiking (Nielsen et al., 1997). High genetic differentiation in introns could therefore be used as an indicator that there is strong selection acting upon a gene. This is seen in introns of *Pgi* and *SdhB*, but not in *Idh2* or *Ldh.*

As expected, in *C. aeneicollis* non-synonymous SNPs are rare compared to synonymous SNPs and SNPs in introns. The allele frequencies of non-synonymous SNPs in Bishop Creek show little variation, with the exception of those found in two genes, *Pgi* and *SdhB*. The pattern of genetic differentiation seen between synonymous and non-synonymous SNPs suggests that there is a selective benefit to maintain that non-synonymous variation. Since there is little variation in non-synonymous SNPs in genes, it is likely that the SNPs found in *Pgi* and *SdhB* are maintained due to performance variation. There is evidence that *Pgi* alleles perform differently under different thermal conditions (Dahlhoff and Rank, 2000; Rank et al., 2007; Dahlhoff and Rank, 2007). It is possible that

the same is true in *SdhB,* but further work is needed to test functional variation between *SdhB* genotypes.

Pgi stands out as the enzyme that varies most between Big Pine Creek and Rock Creek (Figure 3). The allele that is most common in Big Pine Creek declines in frequency until it becomes relatively rare in Rock Creek. Ambient temperatures decline with increasing latitude within this eastern Sierra Nevada transect. Thus, the latitudinal gradient in *Pgi* frequencies suggests that the Big Pine Creek allele is associated with greater tolerance to warmer temperatures. This matches prior work using allozyme analysis of *Pgi,* which showed that individuals with Big Pine Creek alleles for *Pgi* have higher maximum critical thermal limit, higher onset of expression of a 70kD heat shock protein, and higher running speed after repeated heat stress than those with Rock Creek alleles (Neargardner et al., 2003; Rank and Dahlhoff, 2002; Rank et al., 2007). Paradoxically, the variation at Pgi across elevations within Bishop Creek is somewhat inconsistent with this pattern. At high elevation, *Pgi* allele frequencies are more similar to Big Pine Creek, while frequencies at low elevation are more similar to Rock Creek (Figure 5). Due to the geographic proximity of the High Stream Crossing sampling locality to Big Pine Creek, this unusual pattern could develop through gene flow over a pass in the Coyote Flat region and reduced gene flow between the southern and northern basin of Bishop Creek (Figure 7). Alternatively it could be that thermal variation between high and low elevations in Bishop Creek is not great enough to cause selective differences among alleles. Finally, effects of elevation on environmental oxygen, and its physiological

consequences, may play a role.

In *C. aeneicollis*, the two different alleles of *Pgi* have been found to be associated with differing organismal thermal performance, which could result in the maintenance of both alleles in populations experiencing varying thermal conditions (Dahlhoff and Rank, 2000; Dahlhoff and Rank, 2002; Neargarder et al., 2003; Dahlhoff et al., 2008; Dick et al., 2013; Heidl 2013; Sayre 2014). In many other insect species *Pgi* has been found to be under balancing selection due to higher performance of heterozygote individuals (Watt, 1977; Haag et al., 2005). There have been several hypothesized mechanisms responsible for the maintenance of diversity, the first is that it is caused by selective differences on the demands of glycolytic flux (Watt, 1983; Watt, 2003), *Pgi* 's moonlighting functions (Marden, 2013), or a mixture of both (Wheat and Hill, 2014). Glycolytic flux is the rate that metabolites are able to move through the glycolytic pathway. An enzyme's ability to catalyze metabolites quickly in order to meet the demand of rate limiting enzymes and ultimately provide more ATP when needed can result in higher physiological performance. *Pgi* has several moonlighting functions in the regulation of the stress response in the endoplasmic reticulum, cell motility and metastasis (Fairbank et al., 2009; Fu et al., 2011).

Within Bishop Creek, *SdhB* shows greater variation in F_{st} and allele frequency across elevation than other loci. *SdhB* has a striking allele frequency shift with elevation over very short distance. In the less than one kilometer distance from La Hupp to High Stream crossing there is a shift of 0.4 for one of the two non-synonymous SNPs in *SdhB*. Selection could be acting on the role

succinate dehydrogenase plays in the hypoxia-inducible factor (HIF) pathway, which monitors cellular oxygen and unmet metabolic demand for oxygen (Marden et al., 2013). HIF-1 alpha is present in the cytosol, and is constantly tagged for degradation by Hif-prolyl hydrolases (Semenza, 2004). When Sdh is inhibited, buildup of succinate can leak into the cytosol, inhibiting Hif-prolyl hydrolase activity. The increased presence of HIF-1 alpha then forms a dimer with Hypoxia-Inducible Factor beta, which causes several dozen hypoxia responsive genes to be upregulated (Semanza, 1999). HIF accumulation has also been linked to tracheal remodeling as a means of delivering oxygen to hypoxic tissues (Centanin et al., 2008, Marden et al., 2013). Along with its role in hypoxia signaling, Sdh is also the branch-point enzyme connecting the TCA cycle and oxidative phosphorylation. *SdhB* within the Sdh complex provides a place for electrons generated from conversion of succinate to fumarate to attach as they enter the oxidative phosphorylation pathway. Sdh plays a large role in cellular metabolism and in hypoxia signaling, which are both very important pathways for high elevation organisms (Harrison et al., 2006). In the Glanville fritillary, both *Pgi and SdhD* genotype are associated with flight endurance and population growth (Wheat et al., 2011). Further work is needed to distinguish among alternative hypotheses for the adaptive significance of the *SdhB* polymorphism in *C. aeneicollis*, but it an exciting candidate gene nonetheless.

Idh2 showed a strong positive relationship between geographic distance and genetic differentiation among populations. This pattern suggests local adaptation among drainages, as variation was much higher than expected by

drift, and it occurred in a clinal pattern. The SNPs with highest Fst values within *Idh2* are synonymous, leading to no structural variation from coding and no variation in amino acid sequence. The variation in SNPs is likely due to selection acting on a nearby gene or regulatory region outside with hitchhiking of synonymous SNPs in the *Idh2* coding region. Levels of differentiation at *Idh2* allozyme electromorphs observed in prior studies of *C. aeneicollis* do not match the SNP variation documented here (Rank, 1992, Rank and Dahlhoff, 2002; Dahlhoff et al., 2008). Those studies revealed relatively low differentiation along the latitudinal gradient and changes in allele frequency at *Idh2* between 1988 and 1996 were not consistent with temperature changes, as they were with *Pgi* (Rank and Dahlhoff 2002, Dahlhoff et al. 2008). While it is possible that there is no functional variation between *Idh2* genotypes, several studies have shown that *Idh* orthologs from warm vs cold adapted mussel species differ in thermal performance (Lockwood and Somero, 2012; Cashon, 1981). It is possible that functional variants in regulatory regions for *Idh2* cause the high differentiation observed in this study, but this does not explain the lack of variation in prior allozyme studies. Two SNPs in *Idh2* are cytosine/thymine nucleotides; methylated cytosine residues spontaneously deaminate and form thymine residues that cannot be picked up by DNA repair machinery (Bird, 1980). Through methylation, there could be potentially variation in expression that shows up as synonymous C/T SNPs in sequencing (Mandriolo and Manicardi, 2015). Further work is needed to examine the cause of the differentiation of the polymorphisms found in *Idh2* and if there are performance differences between

genotypes in *C. aeneicollis*.

Ldh was found to have as high differentiation across elevation in Bishop Creek as *SdhB*. *Ldh* in insects has a much smaller body of research than *Ldh* in vertebrates, but *Ldh* has been hypothesized to vary if the species uses primarily carbohydrate or lipid metabolism to fuel energetically demanding activity (Kitto and Briggs, 1962). In humans, the product of the reaction, lactate, is used as the primary substrate in gluconeogenesis during exercise (Brooks, 1986). Recent research suggests that lactate is a valuable source of energy during activity where oxygen is limited. While *C. aeneicollis* is rarely seen in energetically demanding activities such as flight, they may rely on anaerobic metabolism while over wintering. These beetles overwinter in leaf litter that is buried in snow for months at a time. A European chrysomelid beetle *Melasoma collaris,* which overwinters in a similar way, was found to accumulate high concentrations of lactate after weeks under anoxic conditions (Meidell, 1983). The *Ldh* SNP detected in *C. aeneicollis* was synonymous, but the amount of differentiation across latitude and elevation was substantial. It, like *Idh2* is most likely due to selection acting on a regulatory region or nearby gene, and the SNP in *Ldh* may reflect genetic differentiation by hitchhiking. Additionally, the SNP in *Ldh* is also a cytosine/thymine nucleotide substitution, this means that there is also the possibility that cytosine methylation deamination is occurs in some allelic variants. It is also possible that since one SNP is responsible for the variation in *Ldh,* codon usage bias if one SNP is favored by the tRNA that synthesizes the protein (Behura and Severson, 2013). Synonymous codon usage has been found

to affect the ability to grow at high temperature in thermophilic bacteria (Zeldovich et al., 2007). More work is needed to examine regulatory regions and enzyme function in order to find the functional driver of the high differentiation across both latitude and elevation.

In this study, I have examined how genes over entire metabolic pathways differentiate in response to environmental variation. I have also been able to fill gaps in genetic knowledge of *C. aeneicollis* by uncovering the genetic basis behind prior protein based population genetic studies, and set a foundation for SNP based validation of allozyme work. I have demonstrated that *C. aeneicollis* exhibits very high levels of genetic differentiation across very short spatial scales, and this differentiation mirrors microclimate variation across latitude and elevation. I have also found novel candidate genes for adaptation to high elevation environments, and demonstrated that by using a Pool-Seq approach and assembling the sequence data using the MESPA pipeline, I am able to accurately mine target genes and large portions of the genome of non-model organisms, and achieve insights into the genetic architecture of adaptation to environmental stress.

LITERATURE CITED

Behura, S.K. and D.W. Severson. 2013. Codon usage bias: causative factors, quantification meathods and genome-wide patterns: with emphasis on insect genomes. *Biological Reviews, 88(1): 49-61*

Bird, A.P. 1980. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research. 8(7): 1499-1504*

Bonin, A., P. Taberlet, C. Miaud, and F. Pompanon. 2005. 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(4): 773-783

Boychuck, E.C., J.T. Smiley, E.P. Dahlhoff, M.A. Bernards, N.E. Rank, and B.J. Sinclair. 2015. Cold tolerance of the montane Sierra leaf beetle *Chrysomela aeneicollis. Journal of Insect Physiology* 81: 157-166

Brooks, G.A. 1986. The lactate shuttle during exercise and recovery. *Medicine and science in sports and exercise.* 18(3):360-368

Brown, W.J. 1956. The New World Species of *Chrysomela* L. (Coleoptera: Chrysomelidae). *Memoirs of the Entomological Society of Canada*. 88: 5-56

Buckley L.B., E.F. Miller, and J.G. Kingsolver (2012). The demographic impacts of shifts in climate means and extremes on alpine butterflies. *Integrative and Comparative Biology*. 26:969–977.

Burton, R.S., R.J. Pereira, F.S. Barreto. 2013. Cytonuclear Genomic Interactions and Hybrid Breakdown. *Annual Revie of Ecology, Evolution, and Systamatics* 44: 281-302.

Callier, V., A.W. Shingleton, C.S. Brent, S.M. Ghosh, J. Kim, J.F. Harrison. 2013. The role of reduced oxygen in developmental physiology of growth and metamorphosis in Drosophila melanogaster. *Journal of Experimental Biology*. 216(23): 4334-4340.

Cashon, E.J., R.J. Van Beneden, and D.A. Powers. 1981. Biochemical genetics of *Fundulus heteroclitus* (L.). IV. Spatial variation in gene frequencies of *IDH-A, IDH-B, 6-Pgdh,* and *Est-S. Biochemical Genetics.* 19: 715-728

Catchen, J., P. Hohenlohe, S. Bassham, A. Amores, and W. Cresko. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology,* 22(11): 3124-3140.

Centanin, L., A. Dekanty, N. Romero, M. Irisarri, T.A. Gorr, and P. Wappner. 2008. Cell Autonomy of HIF Effects in Drosophila: Tracheal Cells Sense Hypoxia and Induce Terminal Branch Sprouting. *Developmental Cell.* 14: 547-558.

Chen I., J.K. Hill, R. Ohlemüller, D.B. Roy, and C.D. Thomas. 2011. Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *Science.*333 (6045): 1024-1026.

Cheviron Z.A., G.C. Bachman, A.D. Connaty, G.B. McClelland, and J.F. Storz. 2012. Regulatory changes contribute to the adaptive enhancement of thermogenic capacity in high-elevation deer mice. *Proceedings of the National Academy of Sciences, USA.* 109: 8635-8640.

Clarke, A. 1993. Seasonal acclimatization and latitudinal compensation in metabolism: do they exist? *Functional Ecology,* 7 (2): 139-149.

Dahlhoff, E.P. , S.L. Fearnley, D.A. Bruce, A.G. Gibbs, R. Stoneking, D.M. McMillan, K. Diener, J.T. Smiley and N.E. Rank. 2008. Effects of temperature on physiology and reproductive success of a montane leaf beetle: implications for persistence of native populations enduring climate change. *Physiological and Biochemical Zoology*, 81(6): 718-732.

Dahlhoff, E.P. and N.E. Rank. 2007. The role of stress proteins in responses of a montane willow leaf beetle to environmental temperature variation. *Journal of Biosciences*, 32: 477-488.

Dellicour, S., Fearnley, A. Lombal, S. Heidl, E.P. Dahlhoff, N.E. 2014. Inferring the past and present connectivity across the range of a North American leaf beetle: Combining ecological niche modeling and a geographically explicit model of coalescence. *Evolution,* 68(8): 2371-2385.

Diaz, H.F. and R.S. Bradley. 1997. Temperature Variations During the Last Century at High Elevation Sites. *Climatic Change at High Elevation Sites.* Springer Netherlands. Houten, NL.

Dick, C.A., N. E. Rank, M. McCarthy, S. McWeeney, D. Hollis, and E.P. Dahlhoff. 2013. Effects of Temperature Variation on Male Behavior and Mating Success in a Montane Beetle. *Physiological and Biochemical Zoology*. 86: 432-440.

Dowling, D.K, U. Friberg, and J. Lindell. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends in* Ecology and Evolution. 23:10 546-554

Dunning L.T., A.B. Dennis, G. Thomson, B.J. Sinclair, R.D. Newcomb, T.R. Buckley. 2013. Positive selection in glycolysis among Australasian stick insects. *BMC Evololutionary Biolology,* 13: 215.

Ellegren, H. and B.C. Sheldon. 2008. Genetic basis of fitness differences in natural populations. *Nature,* 452: 169-175

Ellison, C.K. and R.S. Burton. 2006. Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus. Evolution 60(7):* 1382-1391

Fairbank M, P. St-Pierre, I.R. Nabi. 2009. The complex biology of autocrine motility factor/phosphoglucose isomerase (AMF/PGI) and its receptor, the gp78/AMFR E3 ubiquitin ligase. *Molecular BioSystems,* 5, 793–801.

Fu, M., L. Li, T. Albrecht, J.D. Johnson, L.D. Kojic, I.R. Nabi. 2011. Autocrine motility factor/phosphoglucose isomerase regulates ER stress and cell death through control of ER calcium release. *Cell Death and Differentiation*, 18, 1057–1070.

Gaston, K.J. 2000. Global patterns in biodiversity. *Nature*. 405(6783):220-227.

Gompert, Z., M.L. Forister, J.A. Fordyce, and C.C. Nice. 2008. Widespread mito-nuclear discordance with evedence for introgressive hybridization and selective sweeps in *Lycaeides. Molecular Ecology.* 17:24 5231-5244.

Haag C.R., M. Saastamoinen, J.H. Marden, and I.Hanski. 2005. A candidate locus for variation in dispersal rate in a butterfly metapopulation. *Proc Biol Sci*, 272: 2449-2456.

Harrison, J., M.R. Frazier, J.R. Henry, A. Kaiser, C.J. Klok, and B. Rascon. 2006. Responses of terrestrial insects to hypoxia or hyperoxia. *Respiratory Physiology and Neurobiology, 154(1-2): 4-17*

Heidl, S. 2013. Cytonuclear Interactions among Metabolic Enzyme Loci Mediate Reproductive Success and Larval Performance along Environmental Gradients In Natural Populations of Montane Leaf Beetle. Master's Thesis, Department of Biology, Sonoma State University, Rohnert Park, CA. USA.

Hochachka, P. W. and G.L. Somero. 2002. Biochemical adaptation: mechanism and process in physiological evolution. Oxford University Press, New York.

Hochachka, P.W. and G.L. Somero. 1968. The Adaptation of Enzymes to Temperature. *Comparative Biochemistry and Physiology,* 27: 659-668.

Hoekstra, L.A., M.A. Siddiq, K.L. Montooth,. 2013. Pleiotropic Effects of a Mitochondrial-Nuclear Incompatibility Depend Upon the Accelerating Effect of Temperature in Drosophila. *Genetics.*

Huestis, D.L., B. Oppert, J.L. Marshall. 2009. Geographic distribution of IDH-1 alleles in a cricket are linked to differential enzyme keenetic performance across thermal environments. *BMC Evolutionary Biology,* 9: 113.

Huey, R., and Stevenson, R. 1979. Integrating thermal physiology and ecology of ectotherms. *American Zoology.* 19: 357-366.

Huxley, J. 1932. Clines: an auxiliary taxonomic principle. *Nature,* 142: 219-220

Johansson, U.S., J. Ekman, R.C.K. Bowie, P. Halvarsson, J.I. Ohlson, T.D. Price, P.G.P. Ericson. 2013. A complete multilocus species phylogeny of the tits and chickadees (Aves: Paridae). *Molecular Phylogenetics and Evolution, 69(3):852-860.*

Kitto, G.B. and M.H. Briggs. 1962. Lactate Dehydrogenase in Some Insect Muscles. *Nature,* 193: 1003-1004.

Kofler R., P. Orozco-terWengel, N. De Maio, R.V. Pandey, and V. Nolte. 2011a PoPoolation: A Toolbox for Population Genetic Analysis of Next Generation Sequencing Data from Pooled Individuals. *PLoS ONE,* 6(1): e15925.

Kofler, R., R. Vinay Pandey, and C. Schloetterer. 2011b PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq); *Bioinformatics*, 27 (24): 3435–3436.

Krebs, R.A. and V. Loeschcke. 1994. Cost and benefits of activation of the heat-shock response in *Drosophila melanogaster. Functional Ecology,* 8: 730-737.

Lee, H. Y. Guo, M. Ohta, L. Xiong, B. Stevenson, J.K. Zhu. 2002. LOS2, a genetic locus required for cold-response gene transcription encodes a bi-functional enolase. *EMBO Journal.* 21: 2692-2702.

Lynch, M. 2002. Intron evolution as a population-genetic process. PNAS 99(9): 6118-6123.

Mandrioli, M. and G.C. Manicardi. 2015. Cytosine methylation in insects: new routes for the comprehension of insect complexity. *AIMS Biophysics.* 2: 412–422.

Marden J.H. 2013a. Nature’s inordinate fondness for metabolic enzymes: why metabolic enzyme loci are so frequently targets of selection. *Molecular Ecology,* 22:5743-5764.

Marden, J.H., H.W. Fescemyer, R.J. Schilder, W.R. Doerfler, J.C. Vera, and C.W. Wheat. 2013b. Genetic variation in HIF signaling underlies quantitative variation in physiological and life-history traits within lowland butterfly populations. *Evolution.* 67(4):1105-1115

McMillan, D.M., S.L. Fearnley, N.E. Rank and E.P. Dahlhoff. 2005. Natural temperature variation affects larval survival, development, and Hsp70 expression in a leaf beetle. *Functional Ecology*, 19: 844-852.

Meidell, E.M. 1982. Diapause, aerobic and anaerobic metabolism in alpine, adult *Melasoma collaris* (Coleoptera). *Oikos,* 41: 239-244.

Meirmans, P.E. 2012. The trouble with isolation by distance. *Molecular Ecology* 21(12): 2839-2846.

Montooth, K.L., C.D. Meiklejohn, D.N. Abt, and D.M. Rand. 2010. Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between species of *Drosophila*. *Evolution* 64(12):3364-3379.

Neargarder, G.G., E.P. Dahlhoff and N.E. Rank. 2003. Variation in thermal tolerance and HSP70 expression is linked to phosphoglucose isomerase genotype in a montane leaf beetle. *Functional Ecology* 17: 213-221

Nielsen, D.A., G.L. Jenkens, K.M. Stefanisko, K.K. Jefferrson, and D. Goldman. 1997. Sequence, splice site and population frequency distribution analyses of the polymorphic human tryptophan hydroxylase intron 7. *Molecular Brain Research,* 45(1): 145-148.

Palumbi, S.R. and C.S. Baker. 1994. Contrasting Population Structure from Nuclear Intron Sequences and mtDNA of Humpback Whales. *Mol. Biol. Evol.* 11(3): 426-435.

Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change.*Annual Review of Ecology, Evolution and Systematics*. 37: 637–69.

Parmesan, C. and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421: 37-42.

Parsch, J., S. Novozhilov, S.S. Saminadin-Peter, K.M. Wong, and P. Andolfatto. On the Utility of Short Intron Sequences as a Reference for the Detection of Positive and Negative Selection in *Drosophila. Mol. Biol. Evol.* 27(6): 1226-1234.

Rand, D. 2001. The units of selection on mitochondrial DNA. *Annual Review of Ecology,Evolution, and Systematics.* 37:637-669

Ramprasad, N., E.A. Hornett, J. Hill, C.W. Wheat. 2016. A genome-free genomic scan of Pool-Seq data provides novel insights into the genetic bias of color morph variation in birds and butterflies. *Genome Research.* (in press).

Rank, N.E., Heidl, S., Zavala, N., Roberts, K.T., Smiley, J., Mardulyn, P., Dahlhoff, E. 2013. Interactions between mitochondrial and enzyme loci affect performance and survival in a montane leaf beetle. Evolution Conference 2013. Talk.

Rank, N.E., D.A. Bruce, D.M. McMillan, C. Barclay and E.P. Dahlhoff. 2007. Phosphoglucose isomerase genotype affects runing speed and heat shock protein expression after exposure to extreme temperatures in a montane willow beetle. *Journal of Experimental Biology*, 210: 750-764

Rank, N.E. and E.P. Dahlhoff. 2002. Allele frequency shifts in response to climate change and physiological consequences of allozyme variation in a montane insect. *Evolution*, 56 (11): 2278-2289

Rank, N.E. 1992. A hierarchical analysis of genetic differentiation in a montane leaf beetle (Chrysomela aeneicollis). *Evolution* 46: 1097-1111.

Robinson J.T., H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, and J.P. Mesirov. 2011. Integrative Genomics Viewer. *Nature Biotechnology* 29: 24–26.

Sayre, J. 2014. The effects of mitochondrial and nuclear variation on cold tolerance plasticity in a montane insect. Bachlor's Thesis, Department of Biology, Sonoma State University, Rohnert Park, CA. USA.

Schlotterer, C., R. Tobler, R. Kofler, and V. Nolte. 2014. Sequencing pools of individuals – mining genome-wide polymorphism data without big funding. *Nature Review: Genetics.* 15: 749-763

Sedlazeck, F.J., P. Rescheneder, A. von Haeseler. 2013. NextGenMap: fast and accurate read mapping in highly polymorphic genomes; Bioinformatics, 29 (21): 2790-2791.

Semenza, G.L. 1999. Regulation of mammalian 02 homeostasis by hypoxia-inducible factor 1. *Annual Review of Cellular and Developmental Biology.* 15: 551-578.

Silbermann, R. and M. Tatar. 2000. Reproductive costs of heat shock protein in transgenic *Drosophila melanogaster. Evolution*. 54(6):2038-2045

Smiley, J.T. and Rank, N.E., 1986. Predator protection versus rapid growth in a montane leaf beetle. *Oecologia* 70: 106-112.

Somero, G.N. 1995. Proteins and Temperature. *Annual Review of Physiology*, 57: 43-68

Somero, G.N. 2002. Thermal Physiology and Vertical Zonation of Intertidal Animals: Optima, Limits, and Costs of Living. *Integrative and Comparative Biology.* 42(4):780-789.

Sunday, J.M., A. E. Bates, M.R. Kearney, R.K. Colwell, N. K. Dulvy, J.T Longino, and R. B. Huey. 2014. Thermal-safety margins and the necessity of thermoregulatory behavior across latitude and elevation. *Proceedings of the National Academy of Sciences,* 111: 5610-5615.

Svensson, E., L. Raberg, C. Kooch, and D. HasselQuist. 1998. Energetic stress, immunosuppression and the costs of an antibody response. *Functional Ecology,* 12: 912-919.

Thomas, C.D., A.M.A. Franco, and J.K. Hill. 2006. Range retractions and extinction in the face of climate warming. *Trends in Ecology and Evolution*, 21(8): 415-416.

Thorvaldsdóttir H., Robinson J.T., and Mesirov J.P. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 14: 178-192.

Tolmanek, L. and M.J. Zuzow. 2010. The proteomic response of the mussel congernets *Mytilus galloprovincialis* and *M. trossulus* to acute heat stress: implications for thermal tolerance limits and metabolic costs of thermal stress. *Journal of Experimental Biology 213: 3559-3574.*

Tracy C.R. 1976. Model of dynamic exchanges of water and energy between a terrestrial amphibian and its environment. *Ecological Monographs*. 46: 293–326

Veliz, D., P. Duchesne, E. Bourget, L. Bernatchez. 2006.Stable genetic polymorphism in heterogenous environments: balance between asymmetrical dispersal and selection in the acorn barnacle. *Journal of Evolutionary Biology,* 19: 589-599.

Voltz, K. 2008. The functional duality of iron regulatory protein 1. *Current Opinion in Structural Biology.* 18: 106-111.

Watt WB. 2003. Mechanistic studies of butterfly adaptations. Butterflies: ecology and evolution taking flight. The University of Chicago Press.

Watt, W.B., R.C. Cassin, and M.S. Swan. 1983. Adaptation at specific loci. III. Field behavior and survivalship differences among Colias PGI genotypes are predictable from *in vitro* biochemistry. *Genetics* 103(4): 725-739.

Watt, W.B. Adaptation at specific loci. I. Natural selection on phosphoglucose isomerase of Colias butterflies: biochemical and population aspects. *Genetics* 87(1): 177-194

Wheat, C.W. and J. Hill. 2014. Pgi: the ongoing saga of a candidate gene. *Current Opinion in Insect Science,* 4: 42-47.

Wheat, C.W., H.W. Fescemyer, J. Kvist, E. Tas, J.C. Vera, M.J. Frilander, I. Hanski, and J.H. Marden. 2011. Functional genomics of life history variation in a butterfly metapopulation. *Molecular Ecology, 20: 1813-1828*

Whitehead and Crawford. 2006. Neutral and adaptive variation in gene expression. *Proceedings of the National Academy of Sciences,* 103 (14): 5425-5430.

Willmer, P. 1991. Thermal biology and mate acquisition in ectotherms. *Trends in Ecology and Evolution*, 6(12): 396-399.

Wilson, R.J., D. Gutiérrez, J. Gutiérrez, D. Martínez, R. Agudo, and V.J. Monserrat. 2005. Changes to the elevational limits and extent of species ranges associated with climate change. *Ecology Letters*, 8 (11): 1138–1146.

Zeldovich, K.B., I.N. Berezovosky, and E.I. Shakhnovich. 2007. Protein and DNA sequence determinants of thermophilic adaptation. *PLoS Computational Biology.* 3(1):e5 10.1371/journal.pcbi.0030005

Table 1. Localities and numbers of samples used to construct sequence libraries. Drainages are listed from north to south.

Population	Elevation	Beetles per library
Rock Creek (RC)		
Below Ruby Lake	3299	*8*
Heart Lake	3192	*16*
Mosquito Flat	3067	*16*
Pine Grove	2843	*14*
Bishop Creek (BC)		
High Stream Crossing	3205	*50*
Pipeline Site	3135	*50*
La Hupp	2870	*50*
Big Pine Creek (BP)		
Falls Site	2995	*12*
Black Lake	3197	*12*
Pond Bog	3125	*12*
Sam Mack Lake	3544	*12*
Taboose Pass (TP)		
Pacific Crest Trail Crossing	3279	*16*
Boulder Falls	3316	*16*
Kings River Crossing	3362	*16*

Table 2. Genes of central metabolism showing genetic variation among eastern Sierra Nevada populations of the willow beetle *Chrysomela aeneicollis*. Data show number of exons, number of non-synonymous (NonSyn) and synonymous (Syn) single nucleotide polymorphisms (SNPs), and scaffold length of each contig. (number of base pairs). Enzyme name abbreviations used in Figures 3 and 4.

PATHWAY	Protein	Abbr	Exons	NonSyn SNPs	Syn SNPs	Scaffold length (bp)
Glycolysis	Hexokinase	*Hk*	5	0	7	13,124
	Phosphoglucose Isomerase	*Pgi*	6	1	3	13,868
	Phosphofructokinase	*Pfk*	15	0	10	41,146
	Fructose bP Aldolase	*Fbpa*	3	0	2	7,619
	Triose-phosphate isomerase	*Tpi*	3	1	0	15,812
	Glyceraldehyde 3P dehydrogenase	*G3pdh*	3	0	0	9,389
	Phosphoglycerate kinase	*Pgak*	9	2	7	27,418
	Phosphoglycerate mutase	*Pgam*	4	0	1	15,320
	Enolase	*E*	4	1	4	5,612
	Pyruvate kinase	*Pk*	4	1	8	4,259
TCA cycle	Pyruvate dehydrogenase	*Pdhb*	4	2	5	9,606
	Acetyl CoA synthetase	*Acoas*	6	1	11	14,486
	Citrate synthase	*Cs*	4	0	3	7,570
	Aconitase (m)	*Aco2*	3	3	6	7,523
	Isocitrate dehydrogenase NAD α	*Idh3A*	2	0	2	5,910
	Isocitrate dehydrogenase NAD β	*Idh3B*	7	0	4	26,494
	Isocitrate dehydrogenase NADP(m)	*Idh2*	7	1	6	30,552
	α-ketogutarate dehydrogenase	*αKgdh*	8	2	6	17,370
	Succinyl CoA synthetase α	*ScsA*	5	1	4	13,541
	Succinyl CoA synthetase β	*ScsB*	6	2	1	16,060
	Succinate dehydrogenase FeS	*SdhB*	5	2	3	6,916
	Succinate dehydrogenase b560	*SdhC*	2	0	0	1,894
	Malate dehydrogenase (m)	*Mdh2*	2	1	0	5,099
Other	Malic enzyme I (NADP)	*Me1*	10	3	5	28,302
	Malic enzyme II (NADP)	*Me2*	8	1	11	10,926
	Isocitrate dehydrogenase NADP(c)	*Idh1*	3	5	11	20,631
	Sorbitol dehydrogenase	*Sord*	9	1	7	8,662
	Malate dehydrogenase (c)	*Mdh1*	6	1	6	18,195
	Lactate dehydrogenase	*Ldh*	5	0	1	12,545
	Aconitase (c)	*Aco1*	9	0	0	9,993

Table 3. Two-way factorial analysis of variance of genetic differentiation (F_{st}) between northern (Rock Creek) and southern (Big Pine Creek) populations of *Chrysomela aeneicollis*. F_{st} values were Box-Cox transformed before analysis using SAS PROC GLM. Data shown in Figure 1.

Source of Variation	***df***	**SS**	***F***	***P***
Gene	22	0.01298	3.2	<.0001
Region (Intron/Exon)	1	0.00257	13.8	0.0007
Gene* Region	22	0.00965	2.4	0.0004
Error	2685	0.49117		

Table 4. Two-way factorial analysis of variance of genetic differentiation (F_{st}) between low and high elevation populations of *Chrysomela aeneicollis*. F_{st} values were Box-Cox transformed before analysis using SAS PROC GLM. Data shown in Figure 2.

Source of Variation	***df***	**SS**	***F***	***P***
Gene	22	0. 005260	1.7	0.018
Region (Intron/Exon)	1	0.000152	1.1	0.293
Gene* Region	21	0.000175	1.3	0.186
Error	2765	0.398431		

Table 5. Two-way analysis of covariance for the relationship between distance and Fst value for pairs of populations along latitudinal and elevation gradients. Results based on a repeated measures analysis of exons using SAS PROC MIXED, with individual SNPs specified as subjects and a compound symmetry covariance structure. F_{st} and distance values were Box-Cox transformed before analysis. Data shown in Figure 3.

Source of Variation	***df (num)***	***df (denom)***	***F***	***P***
Gene	22	608	1.2	0.273
SNP (Synonymous/Non-synonymous)	1	608	2.9	0.102
Distance	1	868	49.6	<.0001
Distance*Gene	22	867	3.7	<.0001
Distance*SNP	1	876	7.5	0.006
Error	123	0.336		

Table 6. Synonymous and Non Synonymous mutations in the most variable genes.

Gene	***SNP***	***Amino Acid***
Pgi	A/G	Proline
	A/G	Glutamic Acid
	C/A	Valine
	G/A	**Serine/Phenylalanine**
SdhB	C/T	Isolucene
	T/C	Serine
	C/T	Phenylalanine
	T/C	**Phenylalanine/Serine**
	T/G	**Valine/Glycine**
Idh2	T/C	Lucine
	A/T	Proline
	G/A	Arginine
	G/A	Threonine
	G/A	Glutamic Acid
	C/T	Valine
	C/A	**Proline/Glutamine**
Ldh	C/T	Tyrosine

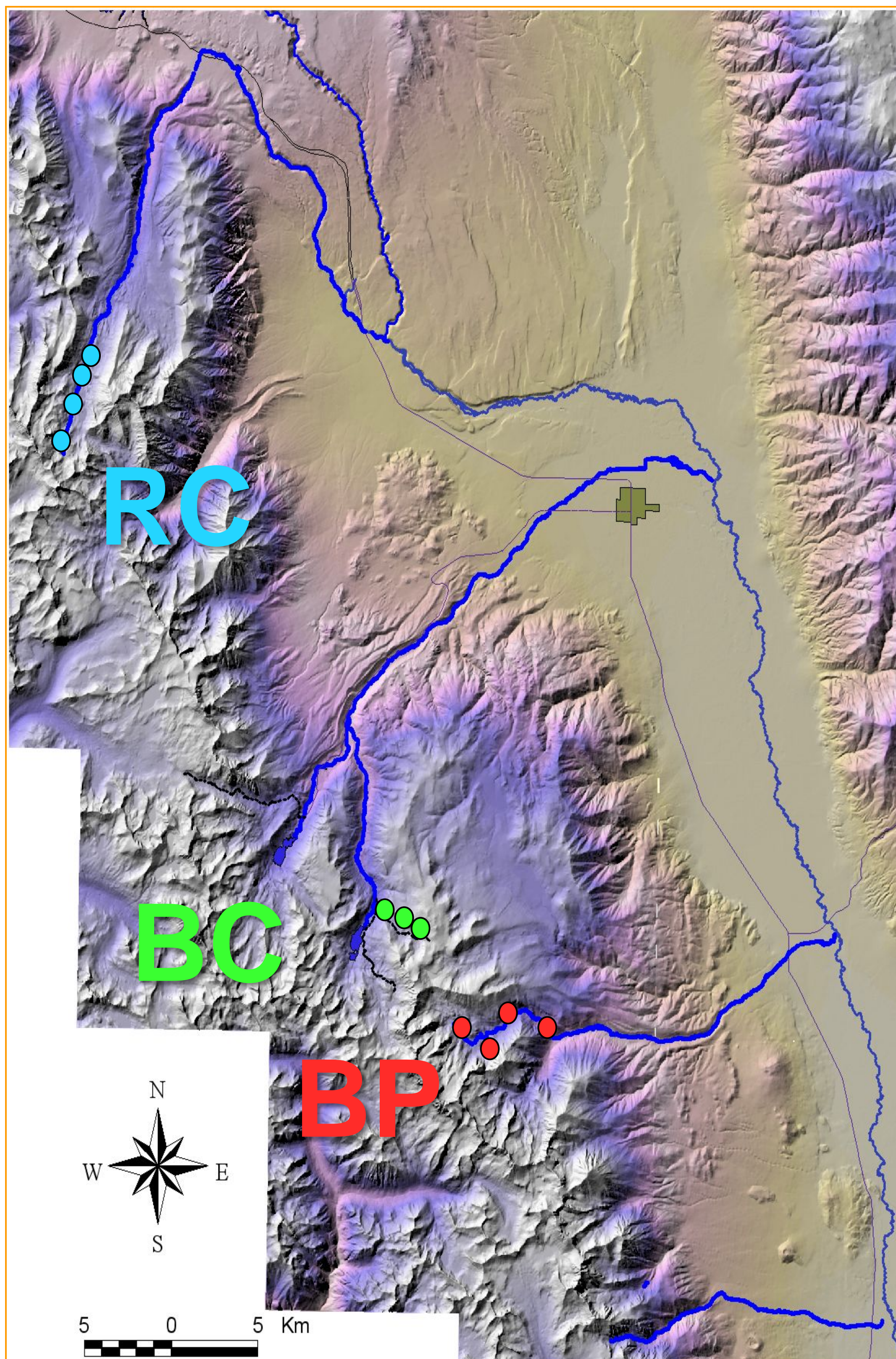

Figure 1. Localities of beetles collected. Rock Creek (RC), Bishop Creek (BC), and Big Pine Creek (BP).

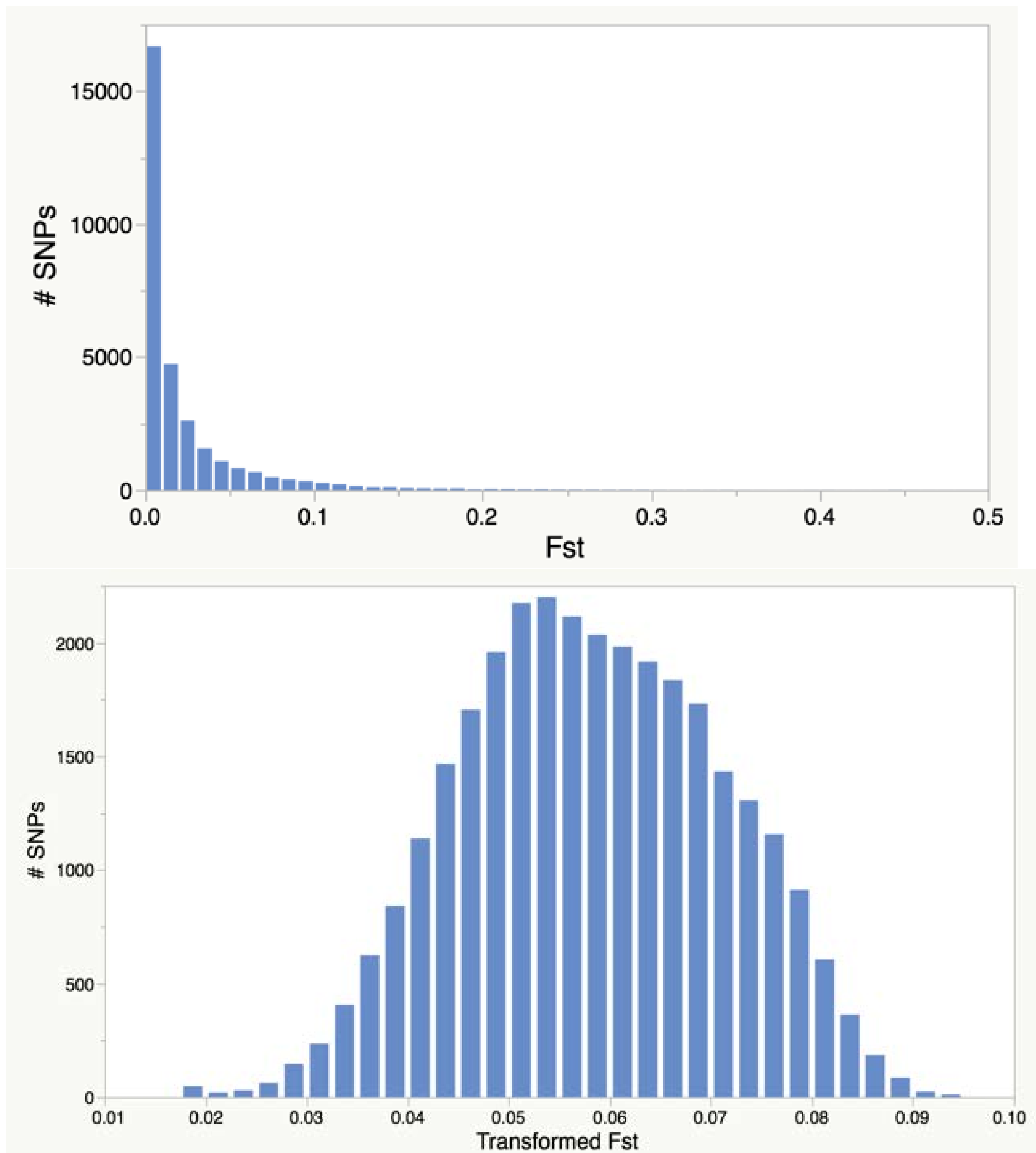

Figure 2. F_{st} before and after best fit Box Cox transformation. Distribution of F_{st} values before (top) and after (bottom) Box Cox transformation.

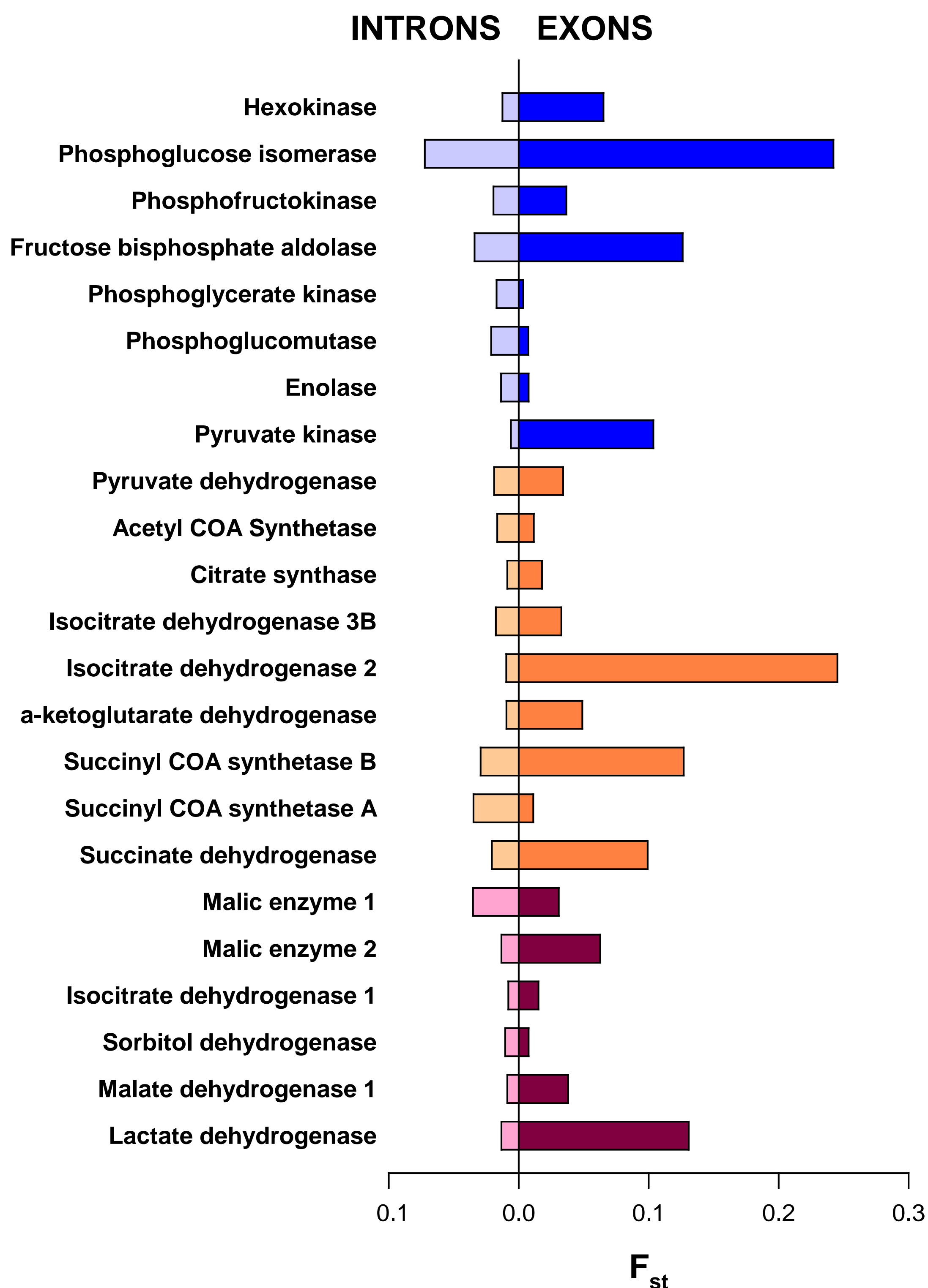

Figure 3. F_{st} values of SNPs found in variable introns and exons of metabolic genes for beetles from Big Pine Creek and Rock Creek. Genes coding for glycolytic enzymes shown in blue, TCA cycle in orange, and other dehydrogenases in red. Values shown are back-transformed least-square means of F_{st} values. Statistical analysis shown in Table 3.

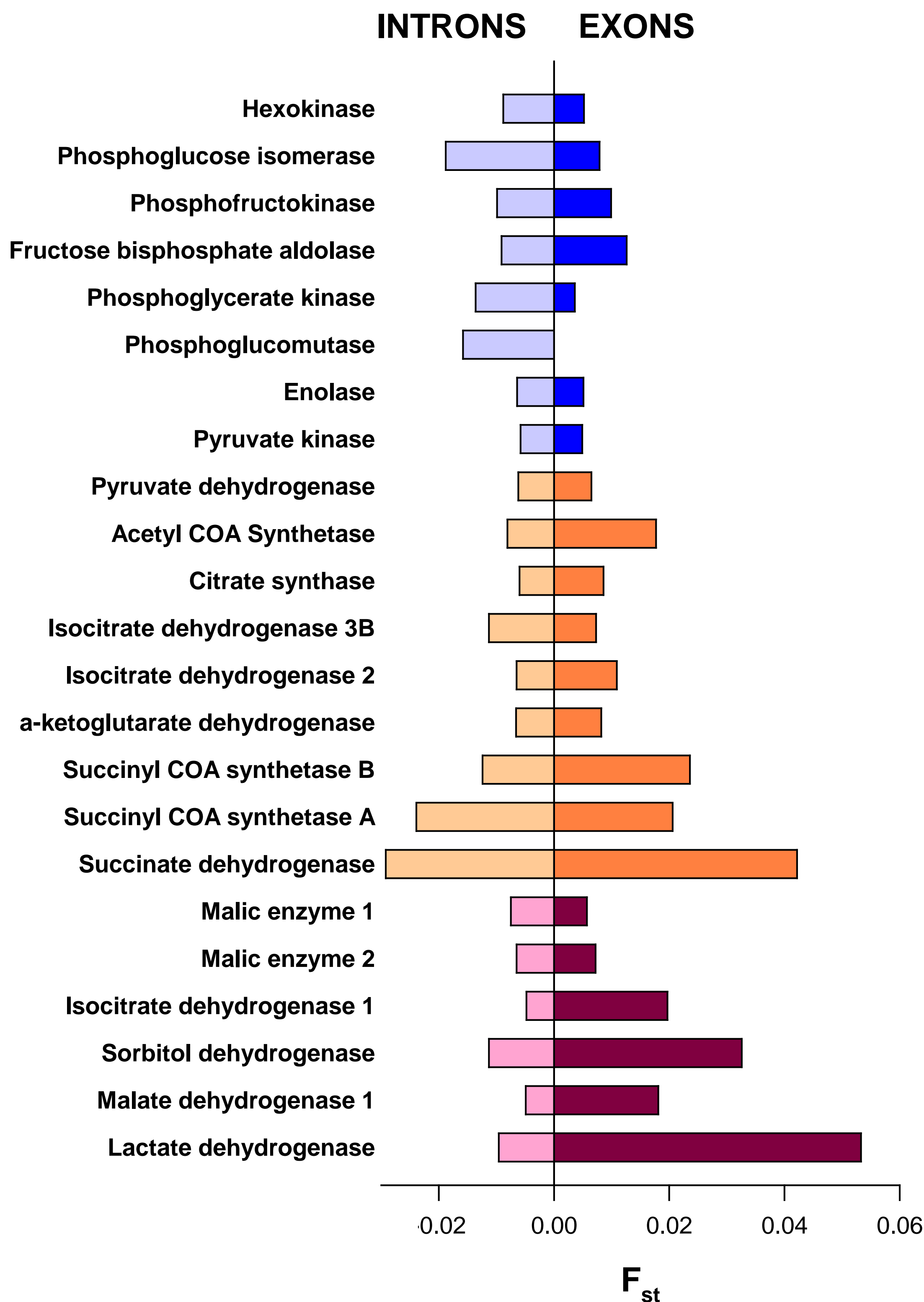

Figure 4. F_{st} of SNPs found in variable introns and exons of metabolic genes for beetles from high and low elevation sites in Bishop Creek. Legend code and values generated as described for Figure 1. Statistical analysis shown in Table 4.

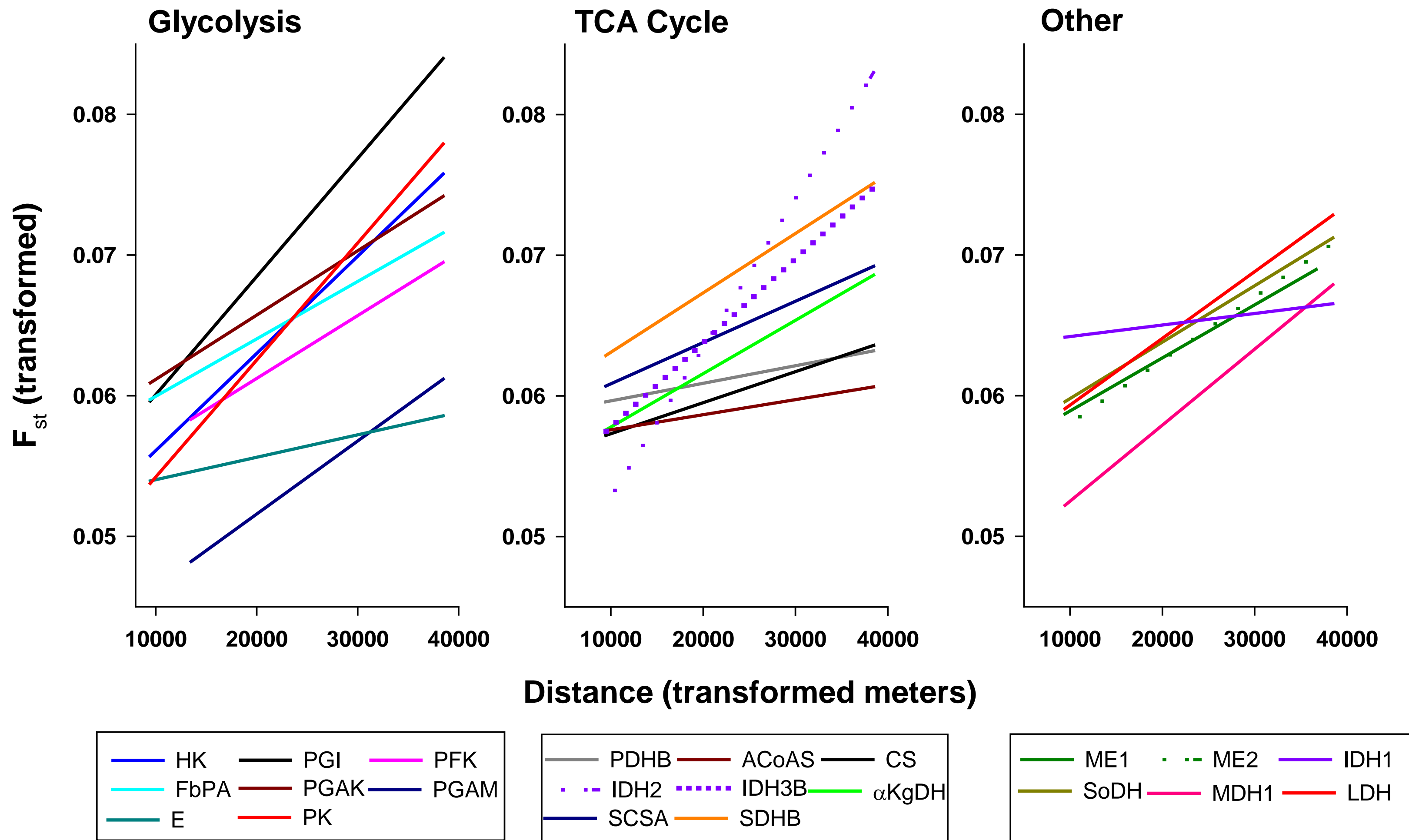

Figure 5. Isolation by distance plots for metabolic enzyme loci that are variable among all populations. Data shown are based on Box-Cox transformed values for both axes. Lines represent predicted values based on pair-wise comparisons between five study populations. Genes are categorized by association with glycolysis (left), TCA cycle (center) or other pathways (right). Gene abbreviations are defined in Table 2. Statistical analysis is shown in Table 5.

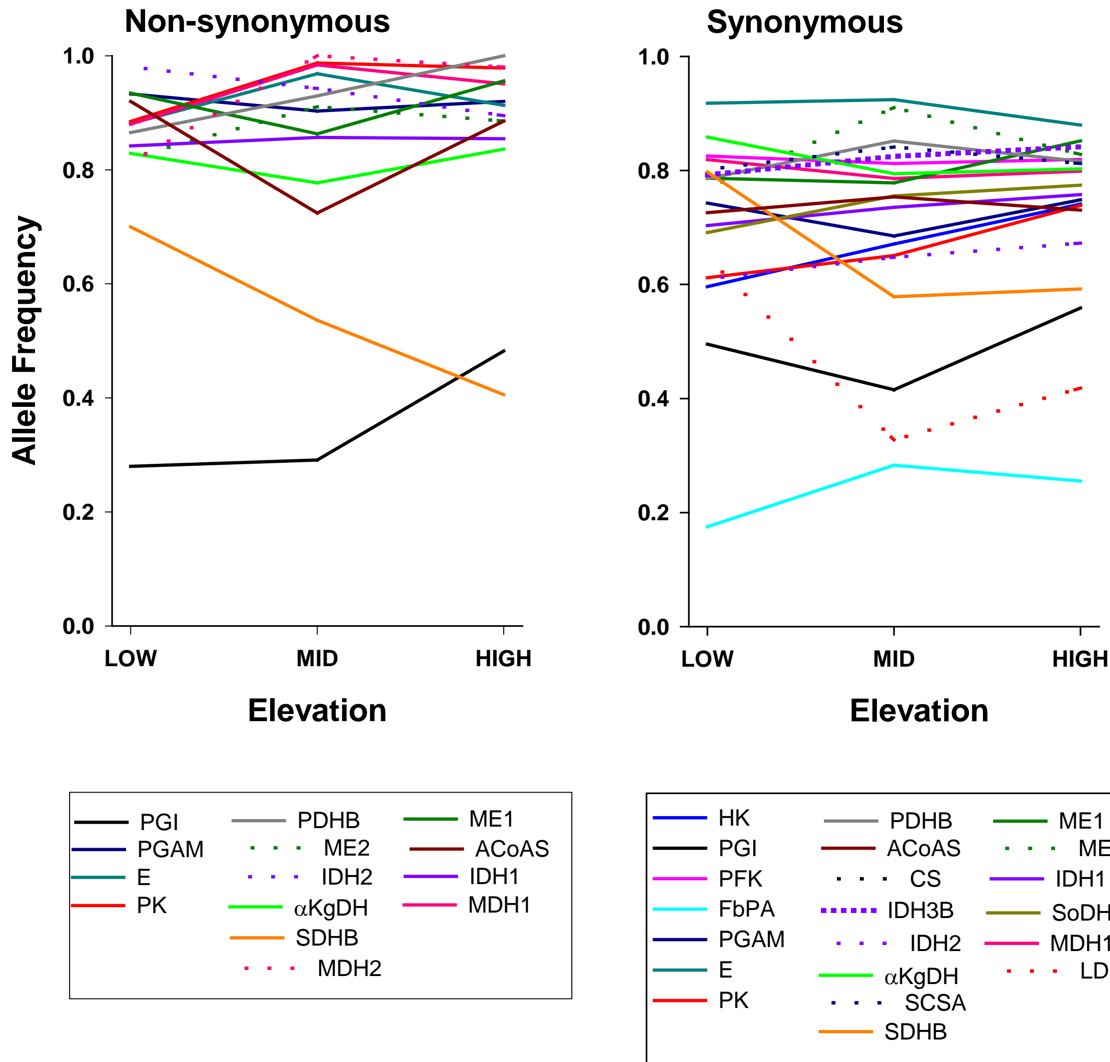

Figure 6. Differences in allele frequency between low-, mid- and high-elevation populations for beetles in the Bishop Creek drainage. Data shown are least squares means for each gene by type of SNP (Non-synonymous/synonymous). Within the legend, genes are categorized by association with glycolysis (left), TCA cycle (center) or other pathways (right). Gene abbreviations are defined in Table 2.

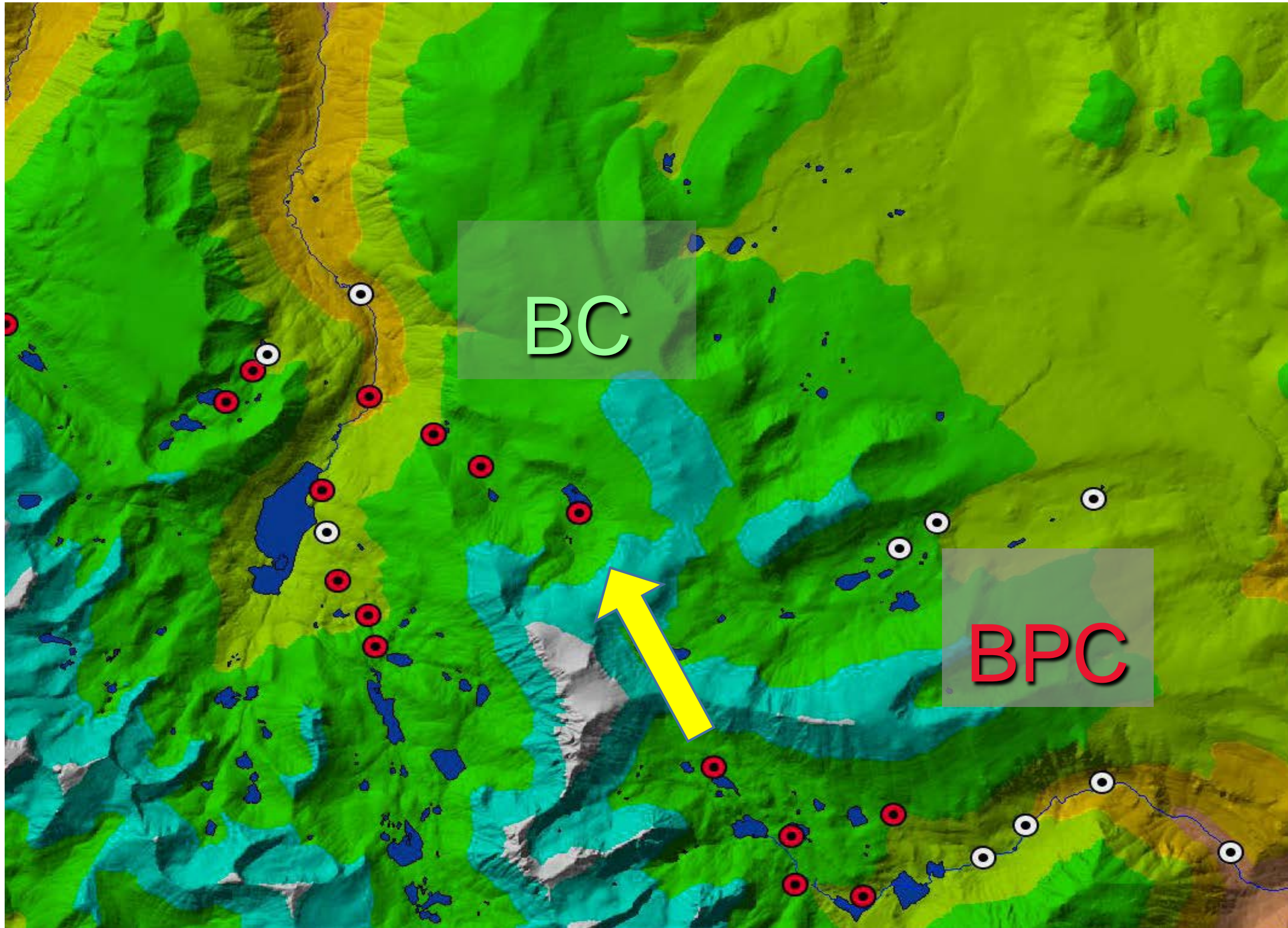

Figure 7. Possible migration corridor between Big Pine Creek and Rock Creek over Coyote Flat. Red circles are extant populations, white circles are extirpated populations.