Fall 2022

Natural Variation of Fructose-1,6-Bisphosphatase in *Colias* Butterflies

Andrea Blair Stokes

Natural variation of fructose-1,6-bisphosphatase in *Colias* butterflies

By

Andrea Blair Stokes

Bachelor of Science
University of South Carolina, 2016

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Science in

Biological Sciences

College of Arts and Sciences

University of South Carolina

2022

Accepted by:

Ward B. Watt, Director of Thesis

Carolyn Wessinger, Reader

Alissa Armstrong, Reader

Cheryl L. Addy, Interim Vice Provost and Dean of the Graduate School

Abstract

In a process known as the evolutionary recursion, it is possible to follow how adaptation affects natural selection at successive stages of a species' life cycle. Through studying allele and genotype frequencies of enzymes in glycolysis, a central metabolic pathway, it is possible to document how polymorphisms can shift the evolving functionality of enzymes within or among species. Fructose-1,6-bisphosphatase (FBPase) is an enzyme in gluconeogenesis that converts fructose-1,6-bisphosphate to fructose 6-phosphate. It is also known to convert the unused amino acid carbon skeletons of glycolysis into glycogen storage. Appearing to play a vital role in the acquisition, storage, and redistribution of carbon skeleton resources derived from larval feeding throughout an individual's lifetime, FBPase could be influential in adaptation leading to diverging life history strategies. FBPase from seven isolated populations of North American Colias butterflies (four C. meadii, two C. eriphyle, and one C. eurytheme) were sequenced. Within C. meadii, we sampled on an elevational gradient both above and below tree line. We first tested the neutral null hypothesis that states any present genetic variation will be neutral. FBPase variation is not neutral, and once the null was violated, we hypothesized that a.) there will be a decrease in FBPase variation as elevation increases within species and b.) there will be a positive phylogentic correlation to variation among species. Strong purifying selection is acting on FBPase nonsynonymous variants. Based

on a linear regression, elevation does not have an effect on FBPase variation within *C. meadii*. Among species, there is no positive phylogenetic correlation to FBPase variation. While initially surprising, the lack of a phylogenetic relationship between *Colias* and FBPase variation now allows for questions to be asked about varying life history strategies that also do not follow phylogenetic relationships. By studying the genetic variation of FBPase, in combination with future field studies, we will be able to directly relate genotypes to performance.

Table of Contents

List of Tables

List of Figures

Chapter 1: Natural variation of fructose-1,6-bisphosphatase of alpine North American *Colias* butterflies

1.1 Introduction

Ever since Charles Darwin introduced the concept of natural selection, scientists have been fascinated with this process and how precisely organisms can adapt to their environments. In 1992 Feder and Watt formalized this process as the "evolutionary recursion": the progression of a genetically variable population through the successive stages of generation *i* in the life cycle. **Figure 1.1** identifies the stages, beginning with a pool of newly fertilized genotypes which build their phenotypes in interaction with their environment. Phenotypes perform biological tasks with environmental variation, leading demography and population genetics to realized fitness values ranging from zero (neutrality) to diverse in response to natural selection. We now have a new array of genotypes for generation *i*+1.

Glycolysis is a highly conserved metabolic pathway that converts glucose into pyruvate, ATP, and reduced electron carriers to be used for various processes. The most important regulating step of glycolysis is the phosphorylation of fructose-6-phosphate to fructose-1,6-bisphosphate by phosphofructokinase (PFK) (Longo 2014). There are also instances when excess energy must be reconverted into carbohydrate storage. This occurs in a process known as gluconeogenesis in which a separate enzyme, fructose-1,6-bisphosphatase (FBPase), dephosphorylates fructose-1,6-bisphosphate to fructose-6-

phosphate (**Figure 1.2**). FBPase facilitates conversion of amino acid or lipid carbon skeletons into glycogen storage. In large mammals, FBPase has been linked to the process of hibernation (Storey, 1997 & Smolinski et al., 2020), but presently there isn't nearly as much known about it in invertebrates.

Colias, or "Sulphurs", are a genus of butterflies composed of 80 or more species that encompass a wide range of biodiversity. The center of its species diversity is in Asian Russia where there are ~50 species present. Here in North America, we have approximately 20 species. North American sulphurs occupy habitats from dry grassland steppes around sea level to alpine tundra at more extreme high elevations. *Colias*, as holometabolous insects, have five "instar" growth stages as larvae. As larvae they feed on the leaves of plants that provide a metabolically complete diet consisting of proteins, lipids, and carbohydrates. In contrast, the same individuals as adults feed almost entirely on nectar containing ~15-25% dissolved solids in the form of simple sugars of angiosperms (Watt et al., 1974). At times, this carbohydrate rich diet can leave the adult lacking metabolically. To overcome this, occasionally adults might access their metabolic stores to make up for what the nectar is lacking. Fortunately, *Colias* have adapted to thrive on these incomplete carbohydrate-rich diets and don't have to rely heavily on their stores for their day-to-day metabolic requirements. Generally, all plants that *Colias* interact with, be it for dietary needs or as host plants for egg laying, are part of the family Fabaceae, consisting of legumes and alfalfas. Within North American *Colias*, there are both univoltine (single broods a year) and multivoltine (multiple broods a year)

species. Here, we focus on three of these *Colias* species: *Colias meadii*, *Colias eurytheme*, and *Colias eriphyle*.

Butterflies are ectotherms requiring the acquisition of heat from the sun for energy to fuel various processes, including all metabolic function. As larvae, *Colias* climb up the stalks of plants to eat the leafy material that is directly in sunlight. While eating, they are also simultaneously warming in the sun. Once they have had their fill, larvae will climb down into the more shaded area of the dense stalks to digest their food (Sherman & Watt, 1973). This is done as a means of escape, due to their camouflaged coloring, from the watchful eye of predators that may be overhead. As adults, *Colias* begin to bask in the sun to produce heat for flight. *Colias* bask with their wings held together over their backs and rely on the heating capabilities of the scales on their underwings. These scale colors are species dependent, but the darker the scale, the faster heat is produced. In colder environments, these scales tend to be either dark green or dark grey. By having darker scales, individuals can bask for shorter periods cutting down on the amount of time they are sitting in the open as potential prey. If a situation arises where an individual isn't allowed enough time to bask before having to avoid potential predation, it is possible that their nutrient stores could be accessed as needed for survival.

Colias meadii are present at higher elevations than *C. eurytheme* or *C. eriphlye*. From previous work of Wheat and Watt (2008), we know that *C. eurytheme* and *C. eriphyle* are sister species and are part of what is known as the "lowland complex", with *C. meadii* being their next closest relative. While slightly misleading, *Colias eriphyle* are

considered a mid-elevation species with some populations found around the same elevation as lower elevation populations of *C. meadii*. Due to the harsh nature of higher elevation winters, *C. meadii* and *C. eriphyle* must diapause as third instars. *C. eurytheme* aren't capable of diapause and populations can be found at lower elevations throughout the year.

Just as individuals are affected differently by their environment based on their phenotype, the environment is also directly interacting with their genotype. It is through these genotype x phenotype x environment interactions that we are able to test a series of hypotheses to begin unraveling how natural selection is acting on a particular species, or process, of interest. To be able to understand the complexity behind these interactions, one must begin with the basic documentation of the variation found in nature. In the case of FBPase, an enzyme that appears to play a vital role in the acquisition, storage, and redistribution of carbon skeleton resources derived from larval feeding throughout an individual's lifetime, there is little known about its variation in nature.

Here, I am going to document the natural variation of FBPase in populations of *Colias meadii* that vary in elevation. Once done, I will expand my scope and compare the variation found in *C. meadii* to the natural variation of FBPase of two other species of *Colias* butterflies: *C. eriphyle* and *C. eurytheme*. To begin, I will test the functional null hypothesis of evolutionary genetics, or the hypothesis of neutral variation. If neutrality is not present, I will move forward with testing the various selective pressures acting on FBPase. If variation has a negative effect on an individual, I expect to see strong

purifying selection against deleterious variants. If not, I would expect to see either balancing or additive selection. If the variation isn't neutral and selective forces are present in FBPase, I expect to find natural variation decreasing in respect to an increasing elevational gradient. Due to more extreme environments at higher elevations, I propose stronger purifying selection will be acting on FBPase in populations found at higher elevations.

1.2 Methods

Sampling Locations and Cataloguing

Seven populations of *Colias* butterflies were sampled across the western United States between 2007 and 2021 (**Figure 1.3, Figure 1.4**). *Colias meadii* populations were collected within the Sawatch Mountain Range at Cottonwood Pass (38.84487°N, 106.41362°W, 3,694 m asl), Scarp Ridge (38.5434°N, 107.0600°W, 3,644 m asl), and Dustin Park (38.5221°N, 106.3937°W, 3,297 m asl) all within Gunnison Co., Colorado, USA. The final *C. meadii* population was from the San Juan Mountain Range at Blue Park, Saguache Co., Colorado, USA (38.09367°N, 106.87291°W, 3,260 m asl). Cottonwood Pass (CTNB2) and Scarp Ridge (SR) are both above tree-line populations with the treeline fluctuating between roughly 3,535 m and 3,688 m above sea level. Two *Colias eriphyle* populations were sampled at Jack's Cabin, Gunnison Co., Colorado, USA (38.73412°N, 106.83054°W, 2,499 m asl) and in Salina, UT along the side of U.S. Route 50, Sevier Co., Utah, USA (38.5720°N, 111.5304°W, 1,371 m asl). The final population of *Colias*, *C. eurytheme*, was sampled just off of the UC Davis campus, Yolo Co., California,

USA (37.4422°N, 121.2533°W, 5 m asl). At each location, individuals were caught and placed in a glycine envelope and then placed in the -20° C freezer.

After 24 hours, each individual was weighed and their wing length was measured. Each abdomen was then removed and split laterally. Individual abdomens were then preserved frozen in Ambion "RNAlaterIce" solution for later RNA extraction.

Sample Processing

Half of each sampled individual's abdomen was homogenized (after cuticle removal) with an IKA T8 homogenizer, and RNA was purified with a Qiagen RNeasy kit and a QiaCube processing robot. Purified RNA then underwent reverse transcription using a general primer (Gen22(15)A3end) with an Invitrogen MMLV kit and cleaned using the QiaCube DNA cleanup protocol to produce cDNA RTs.

Specific fructose-1,6-bisphosphatase primers were designed using Oligo 6 software, initially by reference to consensus sequences of other Lepidoptera, to amplify and sequence FBP cDNA (**Table 1.1**). PCR amplifications using Invitrogen HiFi Platinum Taq were run with a standard temperature cycle routine (denaturing at 94°, annealing at 58°, and elongation at 68°) in BioRad T100 thermal cyclers. RTs were amplified in two blocks: either FBP S-55 for *C. erytheme*, FBP S-48G16 for *C. eriphyle*, or FBP S-27 for *C. meadii* to FBP A855, and FBP S849C to FBP A3end. Amplicons were purified by agarose electrophoresis supported by a Qiagen kit and the QiaCube robot.

Purified amplicons were sequenced by Sanger's method in sense and antisense directions to check for accuracy with appropriate primers and Applied Biosystems'

"BigDye" v. 3.1. A total combination of six primers were used for complete coverage. Sequence reactions were cleaned by molecular sieving using Sephadex G-50 and sent to Functional Biosciences, Inc. (Madison, WI) for reading on an ABI 3730XL Genetic Analyzer. Returned sequences in the form of ABI trace files were checked, aligned, and assembled using BioEdit (Hall 2004). This yielded for each sampled individual a complete FBP gene sequence from its 5' untranslated region (UTR) to the 3' UTR.

Statistical Analysis

All sequence analysis was done using routines in DNaSP6 (Rozas et al. 2017). Cleaned and annotated FBP sequences were uploaded into DNaSP6 and underwent a PHASE algorithm to first estimate the haplotypes from each heterozygous individual. Various internal analyses within DNaSP6 were utilized once haplotypes had been created. Nucleotide diversity was determined using the "polymorphic sites" analysis. The "DNA polymorphism" analysis was utilized to determine the haplotype diversity. Synonymous and nonsynonymous changes were calculated with the "DNA synonymous and nonsynonymous sites" routine. Finally, an AMOVA (Excoffier et al., 1996) was performed on *C. meadii* haplotypes using Arlequin ver. 3.5.2.2.

1.3 Results

In similar studies performed on other enzymes of glycolysis, it has been found that there is a wide range for potential variation. For example, glyceraldehyde phosphate dehydrogenase (GAPDH) has almost no variation with only a single amino acid change in one individual in a sample size of 13. (Wheat et al. 2006) On the other

end of the spectrum, phosphoglucose isomerase (PGI) is highly polymorphic with 136 mutations and 17 nonsynonymous sites, as is phosphoenolpyruvate carboxykinase (PEPCK) which is comparable in its nucleotide diversity (Watt et al. 2013).

Fructose-1,6-bisphosphatase consists of 1008 coding bases resulting in 336 codons omitting the stop codon. Genetic statistics were tabulated using DNaSP 6.1 and reported in **Table 1.2** for all *Colias* populations studied.

Within Species

***Colias meadii*: Cottonwood Pass (CTNB2)**

Cottonwood Pass is the highest elevation population that was sampled for *C. meadii*. The sample collected comprises 18 individuals. Within the 36 haplotypes of CTNB2, there are 47 variants at 46 variable sites of which three are singleton sites. Of the 47 variants, 6 are nonsynonymous resulting in a change of amino acid at the protein level. CTNB2 has nucleotide diversity π_{Σ}= 1.12E-2 and Θ= 1.14E-2.

To determine if FBPase variation is neutral, a Fisher's 2x2 exact test was run using the variable and invariable synonymous and nonsynonymous sites of each population using DNaSP 6.1. If a population is neutral, the proportion of synonymous and nonsynonymous variable to invariable sites would be equal. If a population was experiencing positive selection, there would be an increase in nonsynonymous variable to invariable sites, the inverse would be true for purifying selection. For CTNB2, Fisher's Exact test resulted in p-value<0.001 (**Table 1.3**).

***Colias meadii*: Scarp Ridge (SR)**

Scarp Ridge is the second above tree line population sampled. From 44 individuals and 88 haplotypes, there are 64 variants at 60 variable sites. Eight of those variable sites are singletons. Of the 64 variants, 51 are synonymous and are considered "silent substitutions" since they do not result in a change of amino acid, and 13 are nonsynonymous. Scarp Ridge has nucleotide diversity π_Σ= 1.59E-2 and Θ= 1.25E-2. For Scarp Ridge, a Fisher's Exact test resulted in p-value<0.001 (**Table 1.4**).

***Colias meadii*: Dustin Park (DP14)**

The Dustin Park population of *C. meadii* is the highest elevation below tree line sample collected. The sample comprises 24 individuals resulting is 48 haplotypes. In DP14 there are 68 variants at 67 variable sites, 12 being singleton sites. Of the 68 variants, 12 are nonsynonymous while the remaining 56 are synonymous variants. Dustin Park has nucleotide diversity π_Σ= 1.59E-2 and Θ= 1.52E-2. In a Fisher's Exact test, DP14 had p-value<0.001 (**Table 1.5**).

***Colias meadii*: Blue Park (BLP)**

Blue Park is the lowest elevation population sampled. Of the 52 haplotypes from 26 individuals, 39 are unique. BLP has 52 variants at 51 variable sites, 11 of which are singleton sites. Of the 52 variants, 41 are synonymous and the remaining 11 are nonsynonymous resulting in a change of amino acid. Blue Park has a nucleotide diversity of π_Σ= 1.45E-2 and Θ= 1.12E-2. A Fisher's Exact test resulted in p-value<0.001 for BLP (**Table 1.6**).

Within Species Variation: *Colias meadii*

An analysis of molecular variance (AMOVA) (Excoffier et al., 1992) was run using the four *C. meadii* populations. 95.39% of variation was found to be within populations between individuals, while the remaining 4.61% of variation was found to be between populations. The AMOVA resulted in an F_{ST} = 4.61E-2 and p-value<0.001 (**Table 1.7**).

The proportion of nonsynonymous sites to total variant sites was run against elevation in a linear regression using SYSSTAT. The proportions were arcsin√ transformed to linearize and standardize the percentages resulting in an elevation coefficient of -0.12 with a p-value= 0.42 (**Table 1.8, Figure 1.5**).

Nonsynonymous variants result in a change of amino acid at the protein level. Among all *C. meadii* populations there are 22 amino acid changes in FBPase, 5 of which are shared among all populations as seen in **Figure 1.6**. CTNB2 is the only population without any population specific amino acid changes.

Among Species

Colias eriphyle and *Colias eurytheme* are members of the North American lowland complex of *Colias* butterflies with *C. meadii* being a close related outgroup (Wheat & Watt, 2008).

***Colias eriphyle*: Jack's Cabin (CJC)**

The sample of the Jack's Cabin population of *C. eriphyle* comprises 44 individuals equaling 88 haplotypes. Of the 1008 coding bases of FBPase, 908 are invariant and 100

are variable sites, of which 16 are singletons. CJC has 107 variants, 94 are synonymous and 13 are nonsynonymous. Jack's Cabin has nucleotide diversity π_Σ= 2.16E-2 and Θ= 2.1E-2. A Fisher's Exact test resulted in p-value<0.001 (**Table 1.9**).

***Colias eriphyle*: Salina, UT (SU)**

A second population of *C. eriphyle* was sampled in Salina, UT. The SU sample comprises of 48 individuals. SU has 87 variants at 86 variable sites. Of the 86 variable sites, 19 of them are singletons. The 87 variants consist of 74 synonymous variants and 13 nonsynonymous variants. Salina, UT has nucleotide diversity π_Σ= 2.02E-2 and Θ= 1.81E-2. A Fisher's Exact test resulted in p-value<0.001 (**Table 1.10**).

***Colias eurytheme*: Tracy, CA (DA)**

One population of *Colias eurytheme* was sampled in Tracy, CA. The DA sample consists of 38 individuals and 76 haplotypes. Within DA, there are 99 variants at 93 variable sites, of which 25 are singletons. There are 83 synonymous variants in the DA sample and 16 nonsynonymous variants that result in a change of amino acid. The nucleotide diversity of DA is π_Σ= 1.94E-2 and Θ= .002. Tracy, CA resulted in p-value<0.001 in a Fisher's Exact test (**Table 1.11**).

Among Species Variation: *C. meadii*, *C. eriphyle*, & *C. eurytheme*

Between all three species of *Colias* butterflies, there are 44 nonsynonymous changes that result in a change of the amino acid. Of the 44, only one is shared between all species (**Figure 1.7**). The remaining 43 variants are distributed amongst the species

with *C. meadii* having the most species specific variants and *C. eriphyle* having the fewest. Besides the one variant, *C. meadii* and *C. eurytheme* do not share any variants. Out of the 484 total haplotypes of all three species, the 44 amino acid changes resulted in 107 unique haplotypes at the amino acid level.

1.4 Discussion

Fructose-1,6-Bisphosphatase Variation

Within the range determined by GAPDH and PGI comparison for polymorphic variation within glycolytic enzymes, it is now known that fructose-1,6-bisphosphatase is moderately highly polymorphic. The null hypothesis stating that the synonymous and nonsynonymous sites will maintain the same proportion of polymorphic and invariant sites is strongly rejected. With $p < 10^{-3}$ for Fisher's exact tests in all populations of *Colias* butterflies, it has been proven that FBPase selection is not neutral and that there is strong purifying selection acting on FBPase variation regardless of population or species. This is seen by the large disproportion of nonsynonymous variant to invariant sites to synonymous variant to invariant sites. Roughly 1% of all nonsynonymous sites within *Colias meadii* are variable, while approximately 19% of all synonymous sites are variable. This is clear evidence that purifying selection is acting against nonsynonymous variation in an attempt to purge any potentially deleterious variants out of the population.

Tajima's D is a statistical test used to test selection. It is not a high-powered test, so with smaller sample sizes like these, it is not very reliable. It becomes increasingly

informative as the sample size grows. Tajima's D is meant to test the relationship between segregating sites and nucleotide diversity and is the strongest test available for this purpose (Nei & Kumar 2000). If Tajima's D=0, a population is considered neutral. A positive Tajima's D would indicate a population is undergoing balancing selection, while a negative Tajima's D would signify that a population is neither neutral or in balancing selection, but rather undergoing purifying or positive selection. Signatures of positive and purifying selection are virtually indistinguishable using Tajima's D. In our case, it is already known that purifying selection is acting on FBPase variation based on the Fisher's 2x2 exact tests, so even though the samples are too small to get a significant Tajima's D value, it is safe to say that FBPase variation is not neutral and to reject the master null hypothesis of neutrality.

Within *C. meadii* Variation

Four populations of *Colias meadii* were sampled at various elevations throughout the Gunnison Basin of the Rocky Mountains in Colorado, USA. The highest elevation population sampled was Cottonwood Pass (CTNB2) at 3,694 m above sea level. CTNB2 is located above tree line in a tundra environment with small shrubs and grasses. Another above tree line population sampled was Scarp Ridge (SR). SR is located at 3,644 m above sea level, and while lower in elevation than CTNB2, is still a similar tundra environment. Below tree line, two other populations of *C. meadii* were sampled. Dustin Park (DP14) is located at 3,297 m above sea level and individuals can be found flying throughout grassy patches within an alpine forest. The final *C. meadii* population sampled was Blue Park (BLP). BLP is located at 3,260 m above sea level and is the only

population sampled from the south side of the basin. Individuals can also be found here flying in the open grassy meadows of the alpine forest.

Elevational changes bring tradeoffs in various climatic pressures and habitat compositions. For example, as elevation increases, initially it would seem that the temperature drops. While true, it isn't the entire picture. Higher elevations undergo more extreme temperature fluxes throughout the day than lower elevation environments. This means that populations at higher elevations will experience lower lows and higher highs in one day than populations at a lower elevation. Individuals living in higher elevation habitats will then need to adapt to these fluxes faster than individuals of the same species, but from lower elevation populations. Besides temperature, elevational changes can also change the patterns of precipitation, wind patterns, and in the case of herbivorous insects, host plant communities to name a few (Mayor et al. 2017, Ohler et al. 2020).

Based on the AMOVA performed in **Table 1.7**, it is known that individuals are more variable to others within their population than their population is to other populations. 4.61% of the total variation of FBPase is found among the four populations of *C. meadii* with a p-value < 0.001 signifying that the ~5% difference is significant. While this may seem trivial at first glance, one must remember that FBPase is an enzyme that is a part of a highly conserved metabolic pathway. Being under such strong conservation, any variation to the system must be considered carefully.

To determine if the ~5% of variation among populations was a result of the elevational differences of the populations, a linear regression was run in SYSSTAT comparing the proportion of nonsynonymous variants to total variants against elevation (**Table 1.8, Figure 1.5**). With a p-value=0.42, variation between populations cannot be attributed to elevation. One must also note that this regression was run using only four points of data. Generally speaking, having only four points of data would not give a regression enough power to result in a significant p-value. As more populations are added, there is a chance that elevation might turn into a significant effector on FBPase variation. Unfortunately, *C. meadii* are not the easiest species to sample. Due to their limited distribution, few populations are presently known that are large enough to sample without detrimentally affecting the population. Beyond that, some populations at one time were accessible but are now inaccessible due to dangerous conditions in an extreme environment. Regardless, as the data currently stands, elevation does not influence FBPase variation. This suggests that each population is locally adapted to their specific environments regardless of their elevation and that something within each population is driving the selection acting on FBPase variation.

Among Species Variation

Fructose-1,6-bisphosphatase variation was compared across three species of *Colias* butterflies, *C. meadii* (CME)*, C. eriphyle* (CER)*,* and *C. eurytheme* (CEU). As previously determined, *C. eriphyle* and *C. eurytheme* are two members of the lowland complex of North American *Colias* butteflies, with *C. philodice* and *C. vitabunda* rounding out the group (Wheat & Watt 2008). It was because of this that it was initially

believed that FBPase variation would be more similar between CER and CEU with CME FBPase variation being more distinct. Interestingly, FBPase variation does not follow a phylogentic relationship. There are no fixed differences between any of the three species of interest. This makes sense since glycolysis is a highly conserved metabolic process. In terms of amino acid variants, CEU shares as many amino acid variants with CME as it does with CEU (**Figure 1.7**). If there had been a phylogenetic correlation, CER and CEU would have had an increased number of variants between the two species and fewer shared variants with CME.

Outside of phylogenetic relationships, there are a number of other factors that could be influencing the variation of FBPase among species. *C. meadii* and *C. eriphyle* share a more similar geographic distribution than CME and *C. eurytheme* do. Since CER and CME share similar variants that aren't shared with CEU, there could be some relationship with geographic location and variation. While CEU are considered a migratory species since they have large dispersal rates in the warmer months, in cooler months, they are pushed back into more temperate environments. Due to this, *C. eurytheme* have lost the ability to diapause whereas *C. meadii* and *C. eriphyle* must diapause to survive through the winter. As mentioned previously, FBPase is responsible for converting unused amino acids carbon skeletons of glycolysis into glycogen storage. These glycogen stores are then used during pupation and diapause as nutrient stores. Since CER can not diapause, while CER and CME must, there is reason to believe that a relationship between FBPase variation and diapause exists.

Conclusion

As has been proven above, FBPase variation is not neutral in *Colias* butterflies, but rather is under strong purifying selection. Within *C. meadii* variation was found to not be correlated with elevation which implies that each population is locally adapted to their own environment. Future field studies looking to correlate FBPase variation to microclimate variation are needed to determine if the differences among populations are abiotically controlled or if some other factor is at play. Beyond external factors controlling selection of FBPase, future kinetic studies are needed to determine how internal differences caused by changes in amino acids are shifting the functionality of the enzyme itself. Once finished, it will be possible to relate genotypes to performance and to be able to watch adaptation at work.

Table 1.1: Primers for FBPase PCR Amplification

Primers for FBPase PCR Amplification	
FBP S-55	ACG GTT GCG CGG TAT CGG CTT G
FBP S-48G16	CGG CTT GGC GCG GTT G
FBP S-27	GCC GCC TTT TTT GTG TTT GA
FBP S352	AGA AAC GCG GGA AAT ACG TAG
FBP S849C	GGA AAG CTC CGC CTA CTC TAC
FBP A399	GAT TGA ACC GAC GGA GAC GA
FBP A855	TAC TAT GTA GGA CAT CGG GTT G
A – 3end	AGG TAG GTA GGT AAC AAA AAT AAA AG
Gen22(15)A3end	GAC CAC GCT GAT GCA TAA CGA CTT TTT TTT TTT TTT T

Table 1.2: General FBPase Genetic Statistics of *Colias* Butterflies

	C. meadii (CME)				C. eriphyle (CER)		C. eurytheme (CEU)
	CTNB2	**SR**	**DP14**	**BLP**	**CJC**	**SU**	**DA**
Elevation (m)	3,694	3,644	3,297	3,260	2,499	1,371	5
Number of Individuals	18	44	24	26	44	48	38
Unique Haplotype Counts	29	79	42	39	82	86	74
Variable Site Counts							
Invariant	962	951	941	960	908	925	918
Variable	46	60	67	51	100	86	93
Singleton	3	8	12	11	16	19	25
Variant Counts							
n_{Σ}	47	64	68	52	107	87	99
n_{SS}	41	51	56	41	94	74	83
n_{NSS}	6	13	12	11	13	13	16
Nucleotide diversity							
π	0.01116	0.01589	0.01591	0.01446	0.02155	0.02021	0.01942
π_{SS}	0.05473	0.05814	0.0582	0.05176	0.07921	0.07305	0.07211
π_{NSS}	0.00183	0.00229	0.00224	0.00238	0.00287	0.00317	0.00241
Θ	0.0114	0.01254	0.0152	0.01138	0.02102	0.0181	0.01998

Table 1.3: Fisher's 2x2 Exact Test for Cottonwood Pass

CTNB2	Variable Sites	Invariable Sites	Total
NSS	6	756	762
SS	40	206	246
Total	46	962	1008

Table 1.4: Fisher's 2x2 Exact Test for Scarp Ridge

SR	Variable Sites	Invariable Sites	Total
NSS	13	749	762
SS	47	199	246
Total	60	948	1008

Table 1.5: Fisher's 2x2 Exact Test for Dustin Park

DP14	Variable Sites	Invariable Sites	Total
NSS	12	750	762
SS	56	190	246
Total	68	940	1008

Table1.6: Fisher's 2x2 Exact Test for Blue Park

BLP	Variable Sites	Invariable Sites	Total
NSS	11	751	762
SS	40	206	246
Total	51	957	1008

Table 1.7: Analysis of Molecular Variance (AMOVA) Table for *Colias meadii* Populations

Source of Variation	d.f.	Sum of Squares	Variance Components	Percentage of Variation	Fst	P-value
Among Populations	3	143.257	0.37237 Va	4.61	0.04605	0.000*
Within Populations	444	3424.716	7.71333 Vb	95.39		
Total	447	3567.973	8.0857			

Table 1.8: Regression Coefficients for Nonsynonymous Variation vs. Elevation of *C. meadii*

Regression Coefficients $B = (X'X)^{-1}X'Y$						
Effect	**Coefficient**	**Standard Error**	**Std. Coefficient**	**Tolerance**	**t**	**p-Value**
CONSTANT	0.8670	0.4278	0.0000	.	2.0268	0.1799
ELEVATION	-0.1238	0.1229	-0.5803	1.0000	-1.0076	0.4197

Table 1.9: Fisher's 2x2 Exact Test for Jack's Cabin

CJC	Variable Sites	Invariable Sites	Total
NSS	13	749	762
SS	87	159	246
Total	100	908	1008

Table 1.10: Fisher's 2x2 Exact Test for Salina, UT

SU	Variable Sites	Invariable Sites	Total
NSS	13	749	762
SS	82	164	246
Total	95	913	1008

Table 1.11: Fisher's 2x2 Exact Test for Tracy, CA

DA	Variable Sites	Invariable Sites	Total
NSS	16	746	762
SS	77	169	246
Total	93	915	1008

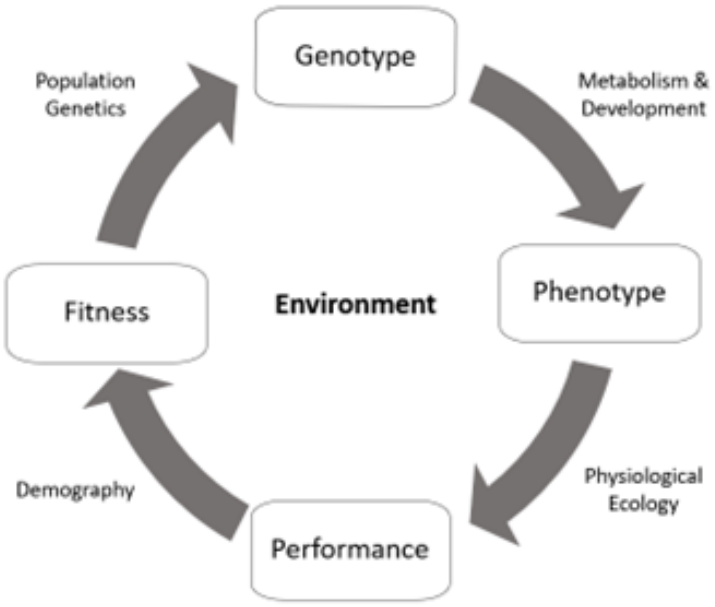

Figure 1.1: The Evolutionary Recursion

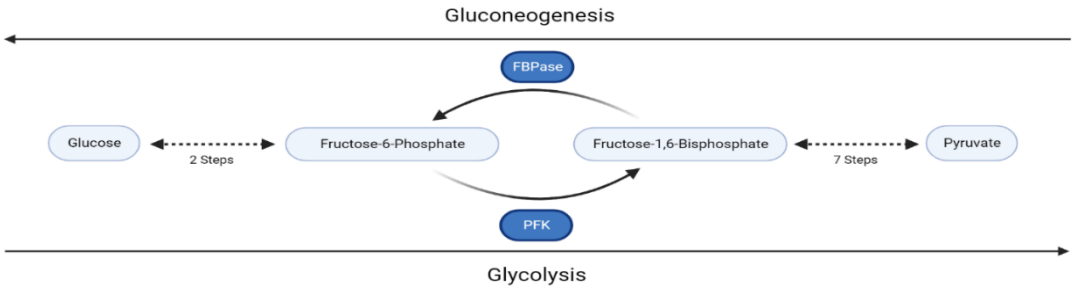

Figure 1.2: Glycolytic Limiting Step Breakdown

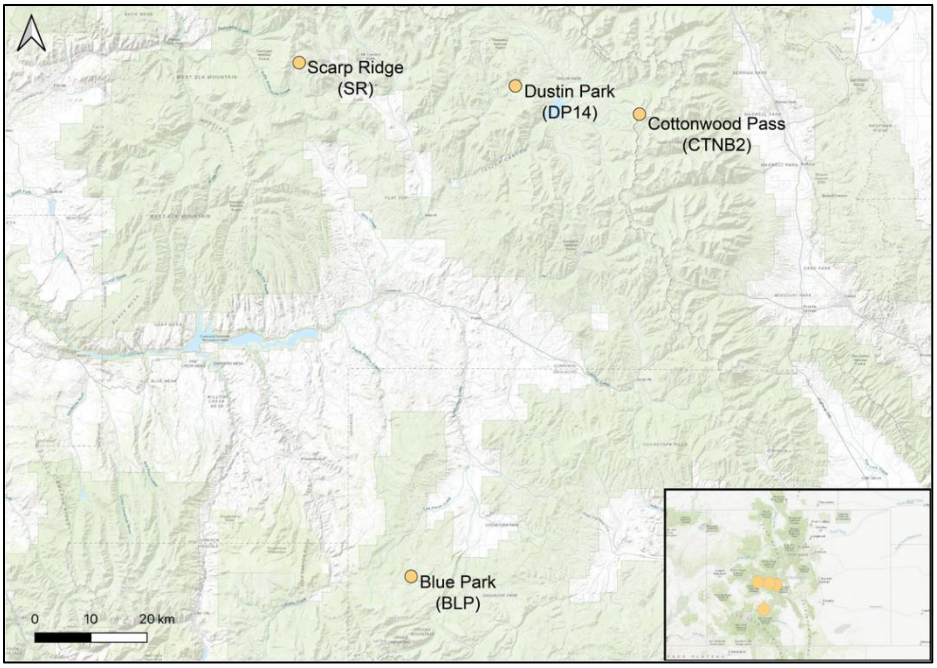

Figure 1.3: Map of *Colias meadii* Sampling Populations

Figure 1.4: Map of *Colias* Sampling Populations

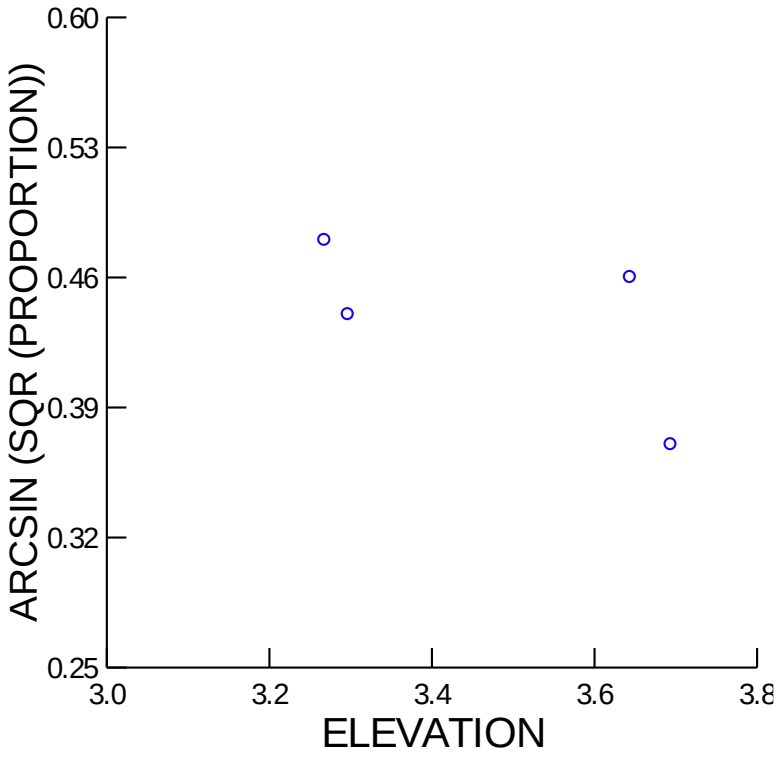

Figure 1.5: Linear Regression of Proportion of Nonsynonymous Variants vs. Elevation for *Colias meadii*

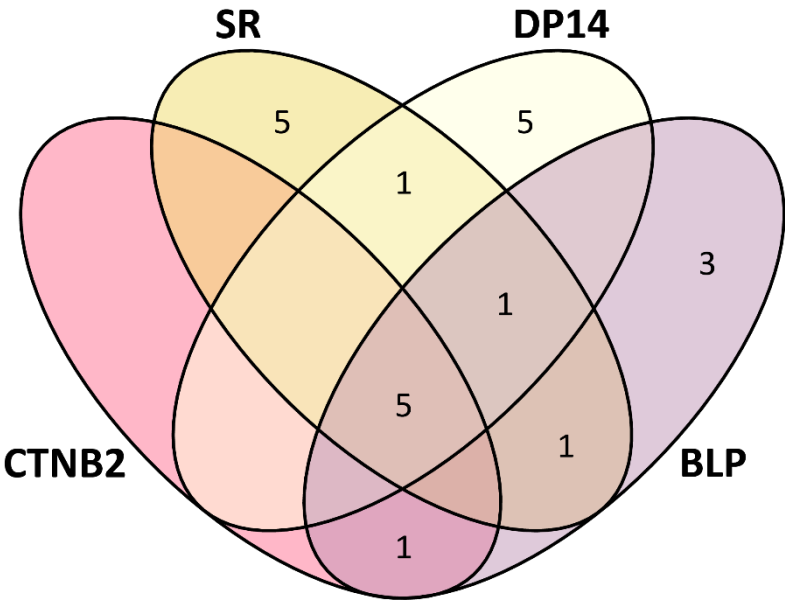

Figure 1.6: Amino Acid Variants Among Populations of *C. meadii*

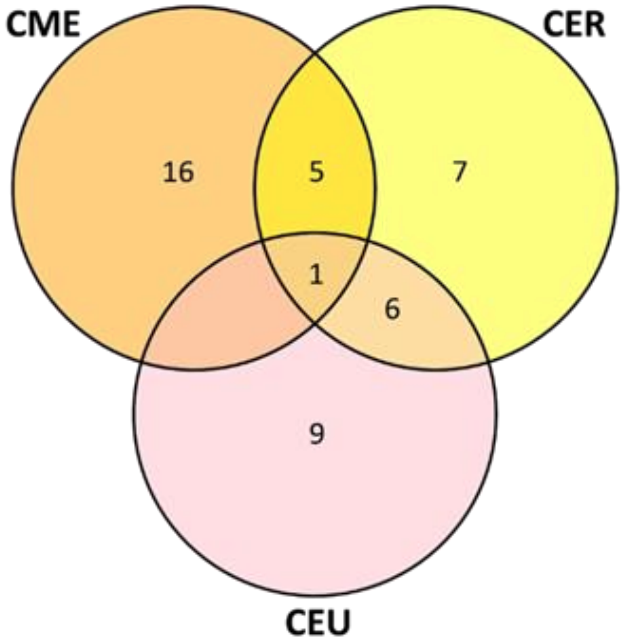

Figure 1.7: Amino Acid Variants Among Species of *Colias* Butterflies

References

Darwin, C., (1859). *The origin of species*. New York: PF Collier & son.

Excoffier, L., Smouse, P. E., & Quattro, J. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479-491.

Feder, M.E., Watt, W. B. (1992). Functional Biology of Adaptation. In *Genes in Ecology* (pp. 365–392). Blackwell Scientific Publications.

Hall, T. (2004). BioEdit: biological sequence alignment editor. URL: www.mbio. ncsu.edu/BioEdit/bioedit.html.

Longo, M.D. (2014). "Exploring the complexity of pathway-level evolution." PhD thesis, Stanford University.

Nei, M., Kumar, S. (2000). *Molecular evolution and phylogenetics*. Oxford university press.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. *Mol. Biol. Evol.* **34:** 3299-3302. DOI: 10.1093/molbev/msx248

Sherman, P. W., & Watt, W. B. (1973). The thermal ecology of someColias butterfly larvae. *Journal of Comparative Physiology 1973 83:1*, *83*(1), 25–40. https://doi.org/10.1007/BF00694570

Smolinski, M. B., Green, S. R., & Storey, K. B. (2020). Characterizing the regulation of pyruvate kinase in response to hibernation in ground squirrel liver (Urocitellus richardsonii). *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *248–249*, 110466. https://doi.org/10.1016/J.CBPB.2020.110466

Storey, K. B. (1997). Metabolic regulation in mammalian hibernation: Enzyme and protein adaptations. *Comparative Biochemistry and Physiology Part A: Physiology*, *118*(4), 1115–1124. https://doi.org/10.1016/S0300-9629(97)00238-7

Tabashnik, B. E. (1983). Host range evolution: the shift from native legume hosts to alfalfa by the butterfly, Colias philodice eriphyle. *Evolution*, 150-162.

Watt, W. B. (1968). Adaptive Significance of Pigment Polymorphisms in Colias Butterflies. I. Variation of Melanin Pigment in Relation to Thermoregulation. *Evolution*, *22*(3), 437. https://doi.org/10.2307/2406873

Watt, W. B., Han, D., & Tabashnik, B. T. (1979). Population structure of pierid butterflies. II. A" native" population of Colias philodice eriphyle in Colorado. *Oecologia*, 44-52.

Watt, W. B., Hoch, P. C., & Mills, S. G. (1974). Nectar resource use by Colias butterflies. *Oecologia*, *14*(4), 353-374.

Watt, W. B., Hudson, R. R., Wang, B., & Wang, E. (2013). A genetic polymorphism evolving in parallel in two cell compartments and in two clades. *BMC evolutionary biology*, *13*(1), 1-15.

Wheat, C. W., Watt, W. B., Pollock, D. D., & Schulte, P. M. (2006). From DNA to fitness differences: sequences and structures of adaptive variants of Colias