USGS
science for a changing world

Prepared in cooperation with the U.S. Fish and Wildlife Service

Genetic Diversity and Population Structure in the Threatened Oregon Silverspot Butterfly (*Speyeria zerene hippolyta*) in Western Oregon and Northwestern California—Implications for Future Translocations and the Establishment of New Populations

Open-File Report 2016–1162

U.S. Department of the Interior
U.S. Geological Survey

Genetic Diversity and Population Structure in the Threatened Oregon Silverspot Butterfly (*Speyeria zerene hippolyta*) in Western Oregon and Northwestern California—Implications for Future Translocations and the Establishment of New Populations

By Mark P. Miller, Thomas D. Mullins, and Susan M. Haig

Prepared in cooperation with the U.S. Fish and Wildlife Service

Open-File Report 2016–1162

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
SALLY JEWELL, Secretary

U.S. Geological Survey
Suzette M. Kimball, Director

U.S. Geological Survey, Reston, Virginia: 2016

Suggested citation:
Miller, M.P., Mullins, T.D., and Haig, S.M., 2016, Genetic diversity and population structure in the threatened Oregon silverspot butterfly (*Speyeria zerene hippolyta*) in western Oregon and northwestern California—Implications for future translocations and the establishment of new populations: U.S. Geological Survey Open-File Report 2016-1162, 23 p., http://dx.doi.org/10.3133/ofr20161162.

ISSN 2331-1258 (online)

Contents

Figures

Tables

Conversion Factors

International System of Units to Inch/Pound

Multiply	By	To obtain
	Volume	
milliliter (mL)	0.033814	ounce, fluid (fl. oz)
microliter (µL)	0.000033814	ounce, fluid (fl. oz)
	Mass	
milligram (mg)	0.000035274	ounce, avoirdupois (oz)

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as °F = (1.8 × °C) + 32.

Datum

Horizontal coordinate information is referenced to North American Datum of 1983 (NAD 83).

Abbreviations

AMOVA	Analysis of Molecular Variance
bp	base pair
COI	cytochrome oxidase subunit I
DNA	deoxyribonucleic acid
mtDNA	mitochondrial DNA
OSB	Oregon silverspot butterfly
PCR	polymerase chain reaction
SSDr	residual sums of squared deviations
UPGMA	Unweighted Pair Group Method with Arithmetic Mean

Genetic Diversity and Population Structure in the Threatened Oregon Silverspot Butterfly (*Speyeria zerene hippolyta*) in Western Oregon and Northwestern California—Implications for Future Translocations and the Establishment of New Populations

By Mark P. Miller, Thomas D. Mullins, and Susan M. Haig

Executive Summary

We present results of population genetic analyses performed on Oregon silverspot butterflies (OSB; *Speyeria zerene hippolyta*) in western Oregon and northwestern California. We used DNA sequences from a 561-base pair region of the mitochondrial cytochrome oxidase subunit I (COI) gene for a dataset comprised of 112 *S. z. hippolyta* and 32 *S. z. gloriosa* individuals collected at 9 locations in western Oregon and northwestern California. The most pertinent findings thus far are summarized as follows:

- Among OSB populations, genetic diversity is lowest at Mount Hebo and highest at Rock Creek and Bray Point. Of the 32 haplotypes detected in OSB, only 2 were shared among populations (1 shared by Mount Hebo, Cascade Head, Bray Point, and Rock Creek, and 1 shared by Rock Creek and Lake Earl). The remaining 30 haplotypes were identified in individual populations, highlighting the strong differentiation among sites. It is unclear if the shared haplotypes represent widespread, naturally occurring genetic variation or if allele sharing among populations is due to translocation history.
- Using full siblings of individuals that were released at Rock Creek and Bray Point in 2012 as comparison standards, the analyses suggest that 54 percent of the sampled individuals from Bray Point were naturally recruited into the population and were not originating from the 2012 release of captive reared individuals. Likewise, 33 percent of the analyzed individuals from Rock Creek were naturally recruited. Both of these estimates may be underestimates if the shared alleles that we identified among populations are naturally occurring and not a product of the 2012 translocations.
- The results suggest that there are about 12–13 COI haplotypes in the Mount Hebo population. The U.S. Fish and Wildlife Service anticipates using Mount Hebo as the source of individuals when establishing new populations in the future. Nonlinear regression models based on a series of rarefaction analyses suggest that progeny from 12, 37, 109, and 326 female individuals would be required to respectively capture 25, 50, 75, and 90 percent of the allelic diversity from Mount Hebo.

- Phylogenetic analyses identified two different haplotype groups, but the two groups did not correspond to the different subspecies used in the analysis. One group included 22 *S. z. hippolyta* haplotypes and 7 haplotypes identified in *S. z. gloriosa*. The second group included eight haplotypes from *S. z. hippolyta,* three haplotypes from *S. z. gloriosa*, and one haplotype that was detected in both subspecies.

Introduction

The Oregon silverspot butterfly (*Speyeria zerene hippolyta*) is listed as "threatened" by the U.S. Fish and Wildlife Service (1980), although the Service 5-year review recommends that a change in classification to "Endangered" is warranted (U.S. Fish and Wildlife Service, 2011). Once occurring from Long Beach Peninsula, Washington, to Del Norte County, California, the subspecies is now found at only four sites in western Oregon (Mount Hebo, Cascade Head, Bray Point, and Rock Creek) and at one site in northwestern California (Lake Earl) (Schultz and others, 2011; U.S. Fish and Wildlife Service, 2011). The subspecies is believed to be extirpated in Washington (Schultz and others, 2011). In this investigation, we use molecular genetic techniques to examine genetic structure and genetic diversity patterns among populations of Oregon silverspot butterflies in Oregon and California. The analyses are intended to facilitate the design of future butterfly translocation efforts that will lead to establishment of new, self-sustaining populations on the Long Beach Peninsula of southwestern Washington and on the Clatsop Plains in northwestern Oregon, as outlined in the Oregon silverspot butterfly recovery plan (U.S. Fish and Wildlife Service, 2001).

This investigation was designed to address the following four questions:

1. **Are there extensive genetic structure and genetic diversity differences among Oregon silverspot butterfly populations?** The taxon has a disjunct distribution, and the degree of genetic differentiation among the Oregon and California populations is currently unknown. If genetic structure exists, the results may indicate relatively low demographic connectivity among populations. If demographic connectivity is low, then population reestablishment following disturbance may be less likely in the absence of naturally occurring colonists from outside of the disturbed location. Considerations of genetic structure patterns also are relevant when planning for translocations. If strong genetic structure exists, resource managers may need to carefully consider the sources of potential colonists to ensure that appropriate populations are used to establish new, self-sustainable populations and that suites of locally adapted genes are present. However, if little genetic structure exists, then managers may effectively source individuals from numerous different populations to use in the translocation process. Such a scenario would be favorable, as it would minimize the demographic impact of butterfly capture on any one population and allow for larger numbers of individual butterflies to be translocated during the early stages of population establishment.

2. **Among populations in Oregon where translocations have already occurred, what proportion of each population's ancestry can be attributed to the translocation source population versus the population's natural ancestry?** Protocols have been developed by the Oregon Zoo (Portland, Oregon) and the Woodland Park Zoo (Seattle, Washington) for the capture and captive rearing of Oregon silverspot butterflies (Andersen and others, 2010). Using these techniques, butterflies from the Mount Hebo, Cascade Head, and Rock Creek populations in Oregon have been translocated at different points in time to help supplement declining populations at Cascade Head, Bray Point, and Rock Creek (Van Buskirk, 2010). If genetic structure exists among the natural Oregon silverspot populations, then the translocation events potentially will generate admixed populations of mixed ancestry. If repeated translocations occur, then the possibility exists that the native genetic diversity within each population will be swamped out and lost. By examining the development of admixed populations in Oregon, insights may be gained about the utility of using mixed populations as sources of individuals for population establishment. Likewise, such investigations also may lead to a better understanding of the effects of additional translocations used to supplement newly established populations in the future.
3. **How many individuals need to be collected and translocated to ensure that newly established populations possess adequate genetic diversity?** The Oregon silverspot butterfly recovery plan specifies the establishment of at least one viable Oregon silverspot butterfly population in the Long Beach Peninsula of Washington and the Oregon Clatsop Plains among the recovery criteria (U.S. Fish and Wildlife Service, 2001). Population viability and persistence are linked with genetic diversity; therefore it is essential to ensure that translocated individuals used to establish new populations harbor sufficient genetic diversity that represents levels found in stable, viable source populations.
4. **Based on genetic data, is the Oregon silverspot butterfly readily discernible from *S. z. gloriosa*, a subspecies that inhabits Oregon and California?** The range of the Oregon silverspot butterfly between the southern Oregon and northwestern California populations is intersected by the distribution of *Speyeria zerene gloriosa*, a taxon that is reportedly distributed throughout southwestern Oregon and northwestern California (fig. 1; also see fig. 3 in U.S. Fish and Wildlife Service, 2001, for historical distribution and overlap of five subspecies of *S. zerene*). Understanding the degree of distinction of these two taxa is relevant for several reasons. If clear genetic differences are noted between *S. z. hippolyta* and *S. z. gloriosa*, then precautions may be warranted when Oregon silverspot butterflies from central Oregon or northern California are used for translocation purposes to ensure that the correct subspecies is collected in the field for captive breeding. Alternatively, if no genetic differences between subspecies are identified, results may call into question the taxonomic designations applied to different silverspot butterfly subspecies and populations, which likewise may have implications for their protected status. Although this study includes only Oregon silverspots and the non-listed *S. z. gloriosa*, note that four other closely related subspecies occur in Washington, Oregon, and California, including the listed subspecies *S. z. behrensii* and *S. z. myrtlaea*.

This document provides detailed information regarding progress on this project thus far (June 2016). The original proposal was designed to provide inferences based on the joint analysis of a mitochondrial DNA (mtDNA) sequence dataset and a nuclear microsatellite dataset. To date, we only have been able to complete work on the mtDNA. We are continuing our efforts to generate the microsatellite data, and will update this document to reflect additional data as they become available.

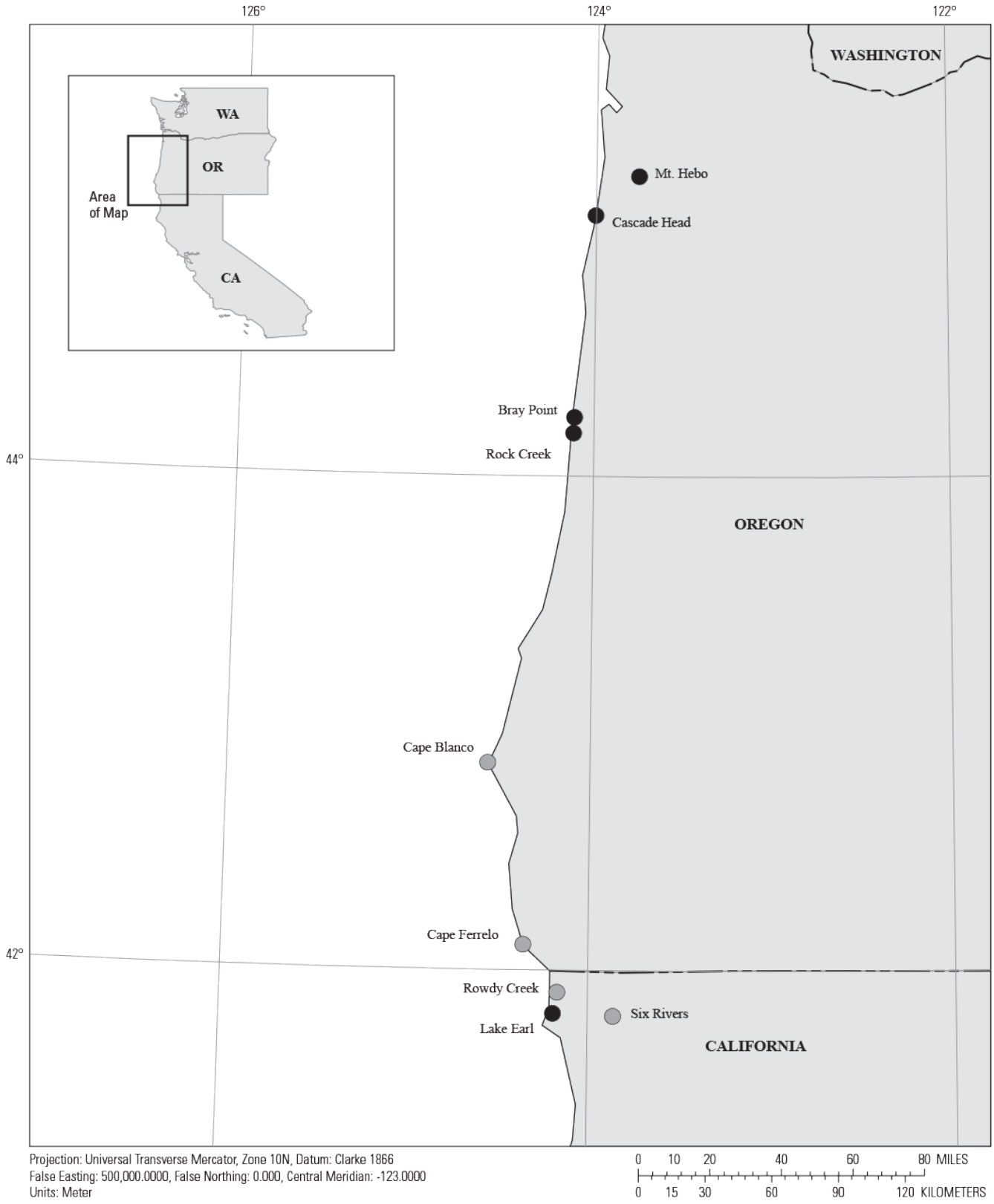

Figure 1. Map showing sampling locations of Oregon silverspot butterflies (*Speyeria zerene hippolyta*; black circles) and *Speyeria zerene gloriosa* (gray circles), western Oregon and northwestern California.

Methods

Samples

Samples were obtained from 119 *S. z. hippolyta* and 32 *S. z. gloriosa*. Tissue samples based on nonlethal wing clips (Keyghobadi and others, 2009; Hamm and others, 2010) of *S. z. hippolyta* were obtained from five sites during summer 2012 (fig. 1): Mount Hebo, Cascade Head, Rock Creek, and Bray Point, Oregon, and Lake Earl, California. Individual butterflies were captured using nets, after which small pieces of hindwing and wing vein material were removed and stored in individual envelopes that were frozen upon returning to the laboratory. All butterflies were released from their original capture location after sample acquisition occurred. Sampling at Cascade Head yielded only eight samples in 2012. Therefore, we incorporated an additional five samples from this site based on collections in 2013. Two additional deceased whole animals were included among our samples, one from Rock Creek (2009) and one from Mount Hebo (2011). Samples of the non-threatened *S. z. gloriosa* were obtained as either wing clips or whole leg from individuals sampled at Cape Ferrelo and Cape Blanco, Oregon, and Rowdy Creek and Six Rivers National Forest, California, and preserved as described earlier in this paragraph. Information associated with each sample included in these analyses is provided in appendix 1.

To better understand the effects and dynamics of translocations that occurred during the 2012 field season, we also were given samples of first or second instar larvae that were captive-reared progeny of field-collected female butterflies obtained during the 2011 field season. Inseminated female butterflies were collected at Mount Hebo and Rock Creek for use in the Oregon silverspot butterfly captive propagation program (Van Buskirk, 2010). Female butterflies were housed at either the Oregon Zoo or the Woodland Park Zoo, and eggs of females that oviposited were carefully maintained and allowed to develop under controlled settings. Overall, we received a single larva from each egg-laying female (Mount Hebo, n=7; Rock Creek, n=12), with the remaining eggs all maintained and allowed to develop in preparation for field releases in 2012. Consequently, the zoo-derived samples represent full siblings of the released individuals from 2012, the relevance of which will be discussed under the methods for research objective 2 (see section, "Research Questions/Objectives and Data Analysis").

Laboratory Methods—DNA Extraction and mtDNA Sequencing

Genomic DNA was extracted from all samples. In the case of wing clips, vein tissue was excised from the samples to optimize DNA recovery, with excess wing membrane and scales being removed from around the vein to minimize the presence of polymerase chain reaction (PCR) inhibiting compounds. Extractions were performed using QIAGEN© DNeasy Blood and Tissue Kit (insect protocol), with the following modifications. Samples were placed in a -80 °C freezer for 2 hours and then crushed using a microtube pestle prior to the extraction protocol. The proteinase K incubation time was increased to 48 hours with the addition of 20 µL of 20 mg/mL proteinase K at 0- and 24-hour time points. Samples were eluted from columns after a 5-minute incubation with 100µL of buffer AE heated to 70 °C. All stocks of DNA samples were diluted 1:10, of which 4 µL was used for each PCR reaction.

Initial amplification of the mitochondrial cytochrome oxidase subunit I (COI) gene was performed with primers K693/PAT2K837 (Simon and others, 1994), which produced a 1,600-base pair (bp) fragment. Based on these initial sequences, we designed five additional primers that ideally could be used to amplify three overlapping segments corresponding to about 1,300 bp of the COI gene (CI-J-1751A/OSP2347R, OSBF2216F/OSBF2569R, OSBF2347F/PAT2K837; table 1). PCR amplifications were performed in 25 μL reactions containing 2.0 mM $MgCl_2$), 1 μM of primers, 100 μM of each dNTP, 1X PCR buffer, and 1U AmpliTaq Gold® DNA polymerase (ThermoFisher Scientific). Thermalcycling parameters included a 5 minute denaturation at 94 °C followed by 35 cycles of denaturation at 94 °C for 30 seconds, annealing at 50 °C for 30 seconds, and extension at 72 °C for 30 seconds, followed by a 10-minute 72 °C extension. All PCR products were bi-directionally sequenced with BigDye® version 3.1 dye terminator sequencing chemistry and resolved on an ABI™ 3730 automated DNA sequencer. With the exception of the 3 prime region of this fragment (amplified by OSBF2347F/PAT2K837), inconsistent amplification or poor sequences from some fragments in some individuals led to the presence of a heterogeneous sequence dataset where not all individuals possessed data from the exact same regions. Thus, despite repeated attempts to obtain the full sequence for all individuals, we ultimately were only able to analyze the final 561 bp of the target region. Resulting sequence chromatograms were aligned, edited, and trimmed using the program Geneious 7.0.6 (Kearse and others, 2012).

Table 1. Mitochondrial primer sequences developed for amplification of the mitochondrial cytochrome oxidase subunit I (COI) gene in *Speyeria zerene hippolyta* and *S. z. gloriosa*.

Primer	Sequence
CI-J-1751A[1]	GGATCACCTGATATAGCATTC
PAT2K837	TCCATTACATATAATCTGCCATATTAG
OSBF2347F[1]	CAGTTGGAATAGATATTGATACTCG
OSBF2347R[1]	CGAGTATCAATATCTATTCCAACTG
SOBF2216F	GGAATAATTTACGCTATAATAGC
SOBF2609R	CCTGTAAATAATGGGTATCAGTG

[1]Modified from primers in Simon and others (1994).

Research Questions/Objectives and Data Analysis

1. Are there extensive genetic structure and genetic diversity differences among Oregon silverspot butterfly populations?

FaBox version 1.4.1 (Villesen, 2007) was used to identify the set of unique haplotypes in the dataset and to convert the sequence alignment to an input file used for analysis with the program Arlequin version 3.5 (Excoffier and Lischer, 2010). Arlequin was subsequently used to quantify genetic diversity for each population using measures that included the number of haplotypes in each population (*A*), gene diversity (*H*), and nucleotide diversity (π).

Arlequin likewise was used to perform an Analysis of Molecular Variance (AMOVA) and to derive measures of differentiation among the set of sampled populations. Two variants of the AMOVA were performed. The first variant used only information on haplotype frequencies within populations and led to a conventional measure of genetic differentiation (F_{ST}). The second variant incorporated information on molecular distances among haplotypes, leading to a derived measure of differentiation that indicated haplotype frequencies and the degree to which haplotypes themselves are phylogenetically divergent (Φ_{ST}). MEGA version 5.2.2 (Tamura and others, 2011) was used to generate Unweighted Pair Group Method with Arithmetic Mean (UPGMA) dendrograms using analysis results as input, thereby allowing us to more readily visually interpret the genetic structure patterns in this system. This analysis was restricted to wild-caught *S. z. hippolyta* given the small sample sizes associated with several of the *S. z. gloriosa* populations (table 2). Given our interest in characterizing patterns in the wild populations, we also did not include samples received from zoos in the genetic structure analyses.

Table 2. Sample sizes (*n*) and genetic diversity levels detected in *Speyeria zerene hippolyta* and *S. z. gloriosa* populations at sites in western Oregon and northwestern California.

[Reported diversity measures include haplotype richness (*A*), haplotype diversity (*H*), and nucleotide diversity (π)]

	n	*A*	*H*	π
S. z. hippoloyta	112	32	0.712	0.0121
Mount Hebo	30	6	0.310	0.0011
Mount Hebo (Zoo)	7	1	0.000	0.0000
Cascade Head	10	3	0.511	0.0013
Bray Point	13	5	0.731	0.0080
Rock Creek	21	9	0.781	0.0205
Rock Creek (Zoo)	12	7	0.833	0.0192
Lake Earl	19	8	0.795	0.0116
S. z. gloriosa	32	11	0.839	0.0167
Cape Blanco	14	4	0.626	0.0033
Cape Ferrelo	9	4	0.583	0.0031
Six Rivers National Forest	3	2	0.667	0.0036
Rowdy Creek	6	3	0.733	0.0015

2. Among populations in Oregon where translocations have already occurred, what proportion of each population's ancestry can be attributed to the translocation source population versus the population's natural ancestry?

The original proposal described our intended use of the program STRUCTURE (Pritchard and others, 2000), which applies a Bayesian clustering procedure to identify and assign individuals to genetic clusters using codominant nuclear genetic data. The primary goal, in this research context, was to determine if evidence exists for genetic admixture in populations where translocations have previously occurred and to determine if native genetic variation still remains. As of this writing (February 2016), the nuclear microsatellite data are unavailable. We have therefore modified our analytical and conceptual framework for research objective #2 to rely on the mtDNA data that currently exist, focusing instead on identifying individuals in the samples who were naturally recruited into the population as opposed to being released through the captive propagation program.

Mitochondrial DNA is haploid and maternally inherited. Therefore, full siblings from a family group will all possess identical mtDNA sequences. Samples that we received from the Oregon and Woodland Park Zoos are full siblings (one individual per female line) of Mount Hebo and Rock Creek individuals that were reared in captivity for field releases at Bray Point and Rock Creek in 2012. Using the mtDNA sequences, we quantified the proportion of individuals at Rock Creek and Bray Point who possessed mtDNA haplotypes that were not found in the samples obtained from zoos. These individuals were inferred to have entered the population through natural recruitment rather than through the captive propagation program. Using this approach, the inferred proportion of naturally developing individuals will underestimate the true proportion if haplotypes identified in zoo samples are geographically widespread haplotypes that also naturally occur in different *S. zerene hippolyta* populations. In this scenario, individuals possessing the same haplotype as a zoo sample may actually be naturally recruited into the population, but not identified as such because of the limited ability to infer kinship using mtDNA.

3. How many individuals need to be collected and translocated to ensure that newly established populations possess adequate genetic diversity?

As the site of the largest known *S. zerene hippolyta* population, Mount Hebo is the planned source of individuals to use for establishment of new butterfly populations in Oregon and Washington (Anne Walker, U.S. Fish and Wildlife Service, written commun.). We, therefore, performed analyses similar to those presented in Jackson and others (2008) and Zhang and others (2015) to (1) estimate the total number of unique haplotypes at Mount Hebo, and (2) determine the relationship between sample sizes and the total number of haplotypes expected in a sample of Mount Hebo individuals. Analyses used rarefaction (equations 1, 2a, and 2b of Kalinowski, 2004) to quantify the number of haplotypes that would be expected to be observed in samples of size n=1 through n=N-1, where N is the original number of analyzed individuals. Nonlinear regression models were fit to the rarefaction data to estimate the rate at which alleles accumulate in a sample as sample sizes vary from 1 to N. We fit four models to the rarefaction curves including:

1. A hyperbolic function of the form $y = (a*x)/(1 + (b*x))$,
2. An exponential function of the form $y = a*(1-b^x)$,
3. A power function of the form $y = x^a$, and
4. A logarithmic function of the form $y = 1 + a*\ln(x)$.

We calculated the residual sums of squared deviations (SSDr) from each model and assumed that the model with the smallest SSDr value would provide the best estimates of the true number of alleles in the population (Jackson and others, 2008). The hyperbolic and exponential functions possess asymptotes at y=a/b and y=a, respectively. We assume that the asymptotic value approximates the true value when either model had the lowest SSDr. The power and logarithmic functions have no asymptotes. Thus, if either of the latter two models has the smallest SSDr, extrapolation is required to infer values based on estimated model parameters and by assuming a known population size (x). For the purposes of these analyses, we assumed underlying population sizes at Mount Hebo of 500, 1,000, and 2,000 when generating extrapolation-based estimates.

4. Based on genetic data, is the Oregon silverspot butterfly readily discernible from *S. z. gloriosa*, a subspecies that inhabits Oregon and California?

This topic was addressed by performing a phylogenetic analysis of all unique haplotypes detected among all samples from both subspecies that were included in the analysis. This analysis allowed us to determine if haplotypes from different subspecies naturally separated into different haplotype groups. Such a pattern, if identified, would provide additional support for the existence of different subspecies (Miller and others, 2010), although subspecies may be defined based on different criteria besides genetics (Haig and others, 2006). Bayesian phylogenetic analysis was performed using the program MrBayes 3.2.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) with a GTR+I nucleotide substitution model chosen by jModelTest 2.1.5 (Darriba and others, 2012). The analysis was based on four replicate runs that used 6×10^6 Markov-Chain Monte-Carlo steps per replicate. Trees were sampled every 2,000 steps (3,000 total recorded trees), with the first 300 steps discarded as burn-in. Convergence among runs was assessed by recording the average standard deviation of split frequencies, as recommend by the program authors. The final tree was visualized using the program MEGA 5.2.2 (Tamura and others, 2011).

Results

Genetic Structure and Diversity

We were able to obtain mtDNA sequences from 112 out of 119 (94 percent) of the *S. z. hippolyta* samples and from all 32 of the *S. z. gloriosa* samples that were collected for this study. We detected 42 haplotypes among the 144 individuals included in the analysis (appendix 2). Thirty-two haplotypes were detected among the *S. z. hippolyta* samples, whereas 11 haplotypes were detected among the *S. z. gloriosa* individuals (table 2; Genbank accession numbers KX074016-KX074057). One haplotype (H3) was identified in both subspecies (appendix 2). With two exceptions, each *S. z. hippolyta* population possessed its own unique sets of haplotypes (appendix 2). One shared haplotype (haplotype H1) was identified at Mount Hebo, Cascade Head, Bray Point, and Rock Creek. A second shared haplotype (haplotype H3) was detected at Rock Creek, Lake Earl, and also among some *S. z. gloriosa* samples from Cape Ferrelo.

Genetic diversity among *S. z. hippolyta* populations was highly variable. Despite the largest sample size of all populations, genetic diversity was least among Mount Hebo samples and greatest among Bray Point, Rock Creek, and Lake Earl samples (table 2). Genetic diversity at Cascade Head was intermediate among *S. z. hippolyta* populations, albeit less than the diversity observed among *S. z. gloriosa* populations (table 2).

Although high levels of genetic differentiation were observed among *S. z. hippolyta* populations (Global $F_{ST} = 0.230$, $P < 0.0001$; Global $\Phi_{ST} = 0.322$, $P < 0.0001$), examination of pairwise patterns indicated that some population pairs are more highly differentiated than others (table 3). When F_{ST} was used as the measure of differentiation, pairwise values ranged from 0.035 to 0.477, whereas pairwise Φ_{ST} values ranged from 0.013 to 0.509. Regardless of the specific measure used, nonsignificant (after sequential Bonferroni corrections) differentiation was identified between Mount Hebo and Cascade Head, Cascade Head and Bray Point/Rock Creek, and Bray Point and Rock Creek (table 3). UPGMA dendrograms derived from the two measures had subtle, but important, differences (fig. 2). In particular, when F_{ST} was used as the differentiation measure, the Bray Point and Rock Creek populations were more similar to one another than they were to other populations. However, when pairwise Φ_{ST} values were used to construct the dendrogram, the Rock Creek population was instead distantly paired with the Lake Earl population. The difference between dendrograms can be attributed to the existence of two separate mtDNA lineages among Rock Creek samples and will be discussed in the context of the phylogenetic analyses performed under research objective 4.

Table 3. Global and pairwise measures of genetic differentiation of Oregon silverspot butterfly populations based on mitochondrial DNA sequences, at sites in western Oregon and northwestern California.

[Values below the diagonal indicate the differentiation measure, whereas values above the diagonal represent P-values. Values in bold are significant at the α=0.05 level after sequential Bonferroni corrections. A) Genetic differentiation measures based on calculations using haplotype frequencies. B) Differentiation measures that include information on the molecular distances among haplotypes in calculations. Unweighted Pair Group Method with Arithmetic Mean (UPGMA) dengdrograms derived from each distance matrix are shown in figure 2]

A) Global F_{ST} = 0.230; P < 0.0001

	MH	CH	BP	RC	LE
Mount Hebo	-	0.195	**0.003**	**0.000**	**0.000**
Cascade Head	0.035	-	0.101	0.092	**0.000**
Bray Point	**0.194**	0.077	-	0.071	**0.000**
Rock Creek	**0.167**	0.063	0.056	-	**0.000**
Lake Earl	**0.477**	**0.327**	**0.235**	**0.192**	-

B) Global Φ_{ST} = 0.322; P < 0.0001

	MH	CH	BP	RC	LE
Mount Hebo	-	0.136	**0.001**	**0.000**	**0.000**
Cascade Head	0.045	-	0.301	0.020	**0.000**
Bray Point	**0.083**	0.013	-	0.024	**0.000**
Rock Creek	**0.435**	0.297	0.209	-	**0.007**
Lake Earl	**0.509**	**0.375**	**0.277**	**0.214**	-

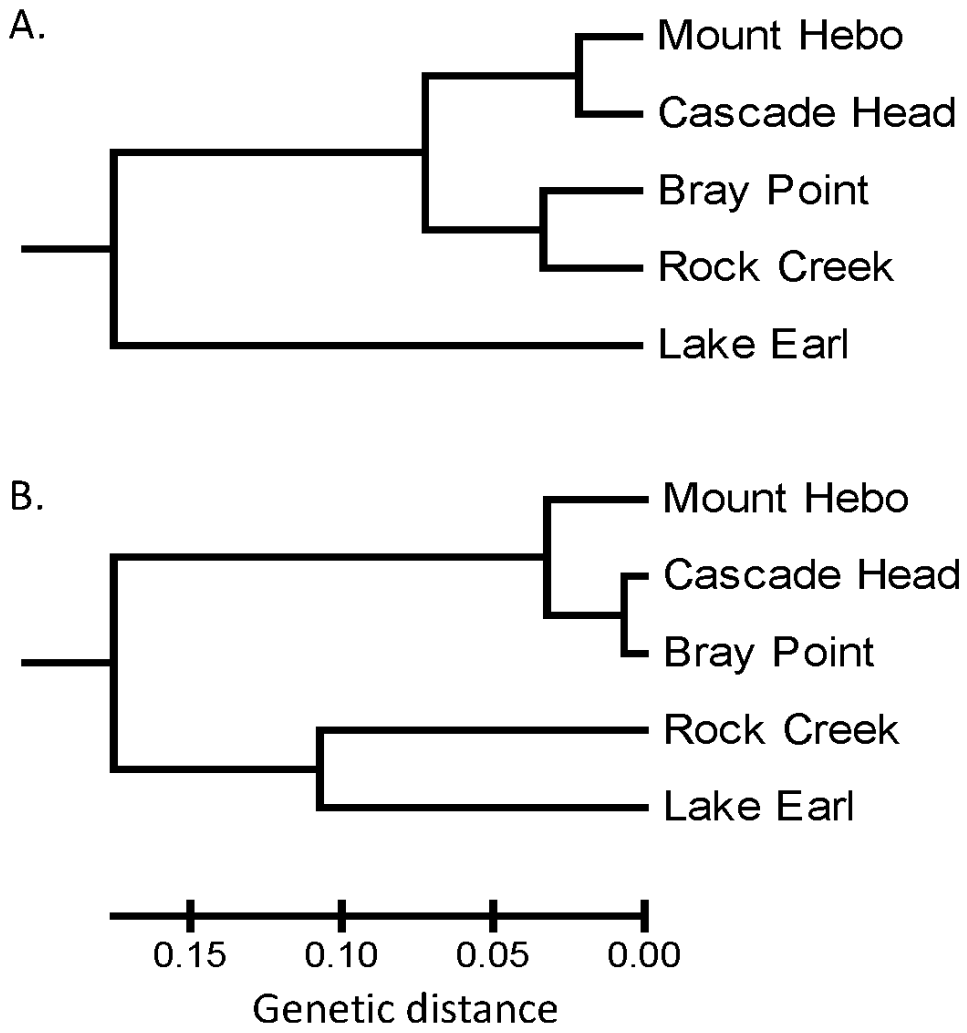

Figure 2. Unweighted Pair Group Method with Arithmetic Mean (UPGMA) dendrograms showing genetic differentiation patterns among Oregon silverspot butterfly populations using haplotype frequencies (F_ST; panel A) or molecular distances among haplotypes ($\Phi_{ST;}$ panel B) as the measure of differentiation. See table 2 for actual distance matrices used as inputs.

Natural Compared to Translocated Genetic Variation in Populations

Seven haplotypes (H1–H7) were identified among the 19 samples obtained from the two zoos (table 4). Of the 13 Bray Point butterflies included in the analyses, 7 (54 percent) possessed haplotypes that were not found among zoo samples, and, therefore, are consistent with being the product of natural recruitment into the population. Seven of the 21 wild-caught Rock Creek samples (33 percent) also did not match any haplotype detected among zoo samples (table 4), and similarly were inferred to be non-translocated individuals.

Table 4. Subset of information from appendix 2 highlighting observed counts of haplotypes detected among Rock Creek and Mount Hebo zoo samples and among wild-caught butterflies from Rock Creek and Bray Point sites, western Oregon.

[Haplotypes observed in wild-caught samples, but not in zoo samples, indicate the detection of a non-translocated individual. See text for more information]

Haplotype	Bray Point	Rock Creek	Rock Creek (Zoo)	Mount Hebo (Zoo)
H1	6	9	5	7
H2			1	
H3		5	2	
H4			1	
H5			1	
H6			1	
H7			1	
H8	4			
H9		1		
H10		1		
H11		1		
H12		1		
H13		1		
H14		1		
H15		1		
H16				
H17	1			
H18	1			
H19	1			
Total individuals	13	21	12	7
Percentage non-translocated	53.85	33.33		

Sample Sizes Required to Capture Genetic Diversity in Populations

The regression analyses indicated that the hyperbolic function provided a superior fit (SSDr=1.97) to the rarefaction data from Mount Hebo relative to the other nonlinear functions considered (exponential, SSDr=2.19; power, SSDr=3.53; logarithmic, SSDr=15.22). Using the parameters inferred from the regression (a=0.348, b=0.028), the results suggest that there are about 12.6 haplotypes present at Mount Hebo based on the calculated asymptote for the best-fit model (fig. 3). Assuming the regression line presented in figure 3, samples of 12, 37, 109, and 326 individuals would be required to respectively capture 25, 50, 75, and 90 percent of the allelic diversity from Mount Hebo. Because this analysis is based on samples of mtDNA, these sample sizes correspond to numbers of females needed for captive rearing or translocation because males do not pass mtDNA on to their progeny.

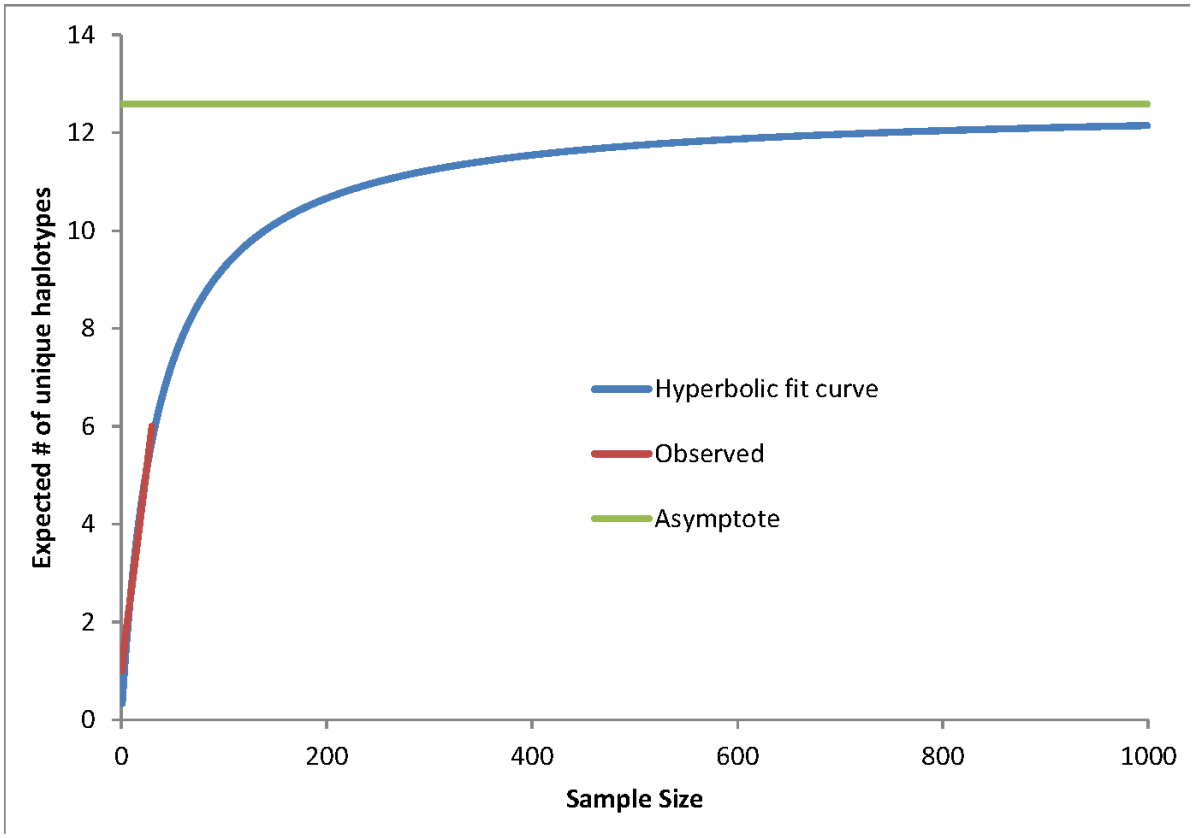

Figure 3. Extrapolated rarefaction curve showing predicted effect of sample size on number of haplotypes expected in samples of *Speyeria zerene hippolyta* from Mount Hebo, western Oregon. Asymptote of the best-fit nonlinear regression (a hyperbolic function; see text for more information) indicates total number of haplotypes estimated for the whole population.

Differentiation of Subspecies

The Bayesian phylogenetic analysis identified two main haplotype groups that were comprised of haplotypes from both *S. z. hippolyta and S. z. gloriosa* (fig. 4). Group I had a posterior probability of more than 99 percent and included 22 *S. z. hippolyta* haplotypes and 7 haplotypes identified in *S. z. gloriosa*. Group II was characterized by a posterior probability of 100 percent, and included eight haplotypes from *S. z. hippolyta,* three haplotypes from *S. z. gloriosa* , and one haplotype (haplotype H3) that was detected in both subspecies. One additional haplotype (haplotype H10), detected in a single individual from Rock Creek, was intermediate to the two groups and could not be clearly categorized. Samples from Mount Hebo and Cascade Head only possessed Group I haplotypes, whereas samples from the other populations indicated mixtures of Groups I and II haplotypes (Bray Point, 12 Group I, 1 Group II; Rock Creek, 11 Group I, 9 Group II, 1 uncategorized; Lake Earl, 16 Group I, 3 Group II). The average standard deviation of split frequencies across the four replicate runs was 0.0059, providing strong evidence that the independent analyses were converging on the same solutions.

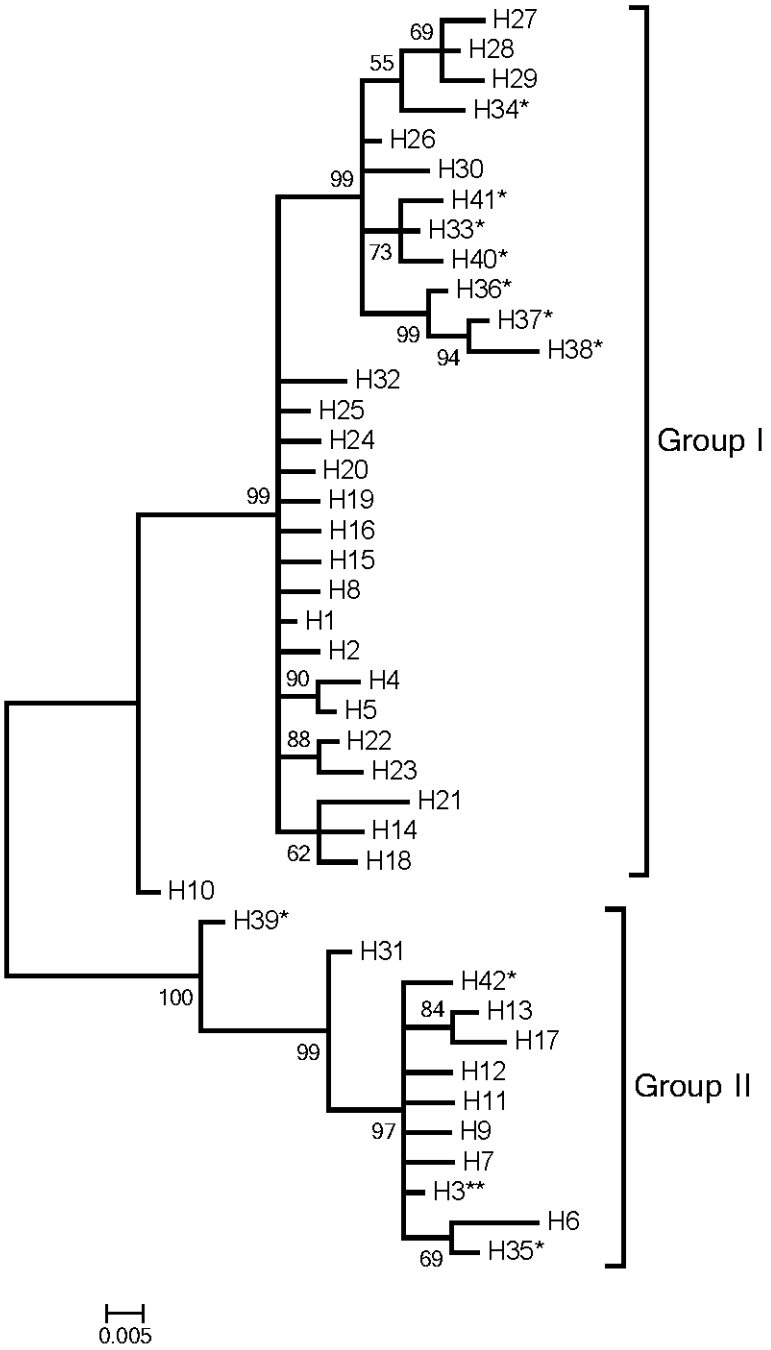

Figure 4. Midpoint-rooted phylogenetic tree showing relationships between haplotypes detecting in *Speyeria zerene hippolyta* (no asterisks) and *S. z. gloriosa* (indicated by single asterisk). One haplotype (H3) was identified in both subspecies and is highlighted with double asterisks. Numbers indicate Bayesian posterior probabilities that quantify the degree of support for individual nodes. Frequencies of each haplotype at different sampling locations are provided in appendix 2.

Discussion

Genetic Diversity and Structure

The analyses provide new insights about genetic structure and diversity in Oregon silverspot butterflies while also providing information that will assist with planning for the establishment of new butterfly populations in Oregon and Washington. Van Buskirk (2000; summarized in Van Buskirk, 2010) performed previous genetic analyses of *S. z. hippolyta* using museum samples from extinct populations and field-caught samples collected during 1995–96 (all samples collected prior to any translocations). Using a 456–613 bp region of the mitochondrial cytochrome oxidase subunit II (COII) gene and smaller sample sizes than those used here, analyses identified very low genetic diversity in most populations. Single haplotypes were identified in the Cascade Head (n=12), Mount Hebo (n=11), Rock Creek (n=9), and Bray Point (n=10) populations, whereas two genetically divergent haplotypes were detected at Lake Earl (n=19). Using a different mitochondrial gene and larger sample sizes, the analyses identified much greater genetic diversity, with three to nine haplotypes detected at each *S. z. hippolyta* population (table 2). However, the Mount Hebo population had the least genetic diversity of any site examined, primarily because a single haplotype (haplotype H1) was numerically dominant in the samples (appendix 2). The remaining populations from the analyses indicated much greater diversity, particularly among the Bray Point, Rock Creek, and Lake Earl samples that possessed both Group I and II haplotypes (table 2, fig. 4, appendix 2).

Strong genetic structure was detected among *S. z. hippolyta* populations (table 3). However, inferences regarding genetic structure and diversity in this subspecies are confounded by the fact that translocations occurred at Rock Creek and Bray Point in the year that most of the samples were collected and that translocations have occurred at these sites and at Cascade Head at various points in time since 2000 (Van Buskirk, 2010). Thus, if translocated individuals were able to survive and reproduce, the Van Buskirk (2000) analyses would provide a better depiction of pre-translocation genetic structure because samples from that study were obtained in or prior to 1996. As with our study, Van Buskirk (2000) identified strong genetic structure between populations, with the primary exception being that Mount Hebo and Cascade Head possessed a single identical haplotype. Our study, based on a different mitochondrial gene and larger sample sizes, provided evidence of more extensive allele sharing, as haplotype H1 was identified at Mount Hebo, Cascade Head, Rock Creek, and Bray Point (appendix 2). Thus, the presence of this more widespread haplotype may be a consequence of recent translocations. Although the Lake Earl population was significantly differentiated from all the Oregon *S. z. hippolyta* populations (table 3), this isolated population nonetheless shared the second-most common haplotype with *S. z. hippolyta* from Rock Creek and *S. z. gloriosa* from Cape Ferrelo (haplotype H3; appendix 2). This result also is consistent with results presented in Van Buskirk (2000), who also identified shared haplotypes between Lake Earl and other Oregon populations. Lake Earl has not been the recipient of translocated individuals, suggesting that the pattern is naturally occurring and not a product of any prior management activities. Finally, our larger sample sizes relative to Van Buskirk (2000) also identified numerous low-frequency haplotypes in most populations (appendix 2). Given that these low-frequency haplotypes were usually unique to specific populations, they may be reasonable indications that naturally occurring genetic variation still exists at each population despite the relatively recent translocations that have occurred.

Natural Compared to Translocated Genetic Variation in Populations

In 2012, the year in which most of the samples were collected, population supplementation occurred at Rock Creek and Bray Point using the progeny of mated females collected at Rock Creek and Mount Hebo in 2011. Despite the introduction of about 1,000–1,100 larvae and pupae at each of these sites, we were able to conservatively estimate that 53.85 percent of Bray Point and 33.33 percent of Rock Creek samples originated through natural recruitment into the population (table 4). This inference represents a lower bounds estimate because some additional naturally occurring variation may have gone undetected at these sites. Furthermore, some individuals who possessed the geographically widespread shared haplotype H1 also may have originated from natural recruitment rather than translocations. This problem is further confounded at Rock Creek, because Rock Creek individuals from 2011 were used to produce some of the caterpillars and pupae that were released back to that site in 2012. Formally addressing this issue requires the existence of a nuclear genetic marker data set, which we are continuing to try to develop using sets of microsatellite loci that we have identified within the Oregon silverspot butterfly genome. If the data set is generated for the current samples, we will be able to use the data to identify wild-caught individuals that are full siblings of the zoo samples that were provided for analysis. If identified, we will be able to unequivocally determine the origin of all butterflies in the data set (wild compared to translocated) and provide a more refined understanding of the degree of natural recruitment that is occurring at Bray Point and Rock Creek. The original objective for this component of the project involved identifying the extent of genetic admixture within the wild butterfly populations (see research objective 2 in section, "Research Questions/Objectives and Data Analysis"). This analysis will only become possible if the microsatellite data are successfully generated from the samples.

Sample Sizes Required to Capture Genetic Diversity in Populations

A series of rarefaction and nonlinear regression analyses were performed to better understand the relationships between sample sizes and the expected number of alleles in a sample (fig. 3). The motivation for this analysis included the need to plan for the establishment of new *S. z. hippolyta* populations in Oregon and Washington. When establishing a new population, founding individuals should ideally represent a large fraction of the genetic diversity from source populations to maximize population health and viability. The analyses suggest that samples of 12, 37, 109, and 326 individuals are required to respectively capture 25, 50, 75, and 90 percent of the allelic diversity from Mount Hebo (fig. 3), the planned source of individuals for new population establishments. The analyses assume that the proportional values reported also apply to other genetic loci within the *S. z. hippolyta* genome. However, in the absence of a microsatellite data set, we are currently unable to evaluate this assumption. Furthermore, if the current captive propagation system is used to produce individuals for population establishment, then the number of individual wild-caught females used to produce progeny represents the actual sample size. For example, if the captive propagation program were able produce 1,000 caterpillars from 10 initial egg-laying females, the true sample size would be 10, not 1,000, because the genetic diversity could never exceed the diversity of the initial 10 founders. Thus, multiple introductions performed over several sequential years may be required depending on the level of genetic diversity desired in any newly established populations.

Differentiation of Subspecies

The phylogenetic analyses indicated the presence to two primary haplotype groups among the samples included in this study (fig. 4). Although only a single haplotype was shared between *S. z. hippolyta* and *S. z. gloriosa* (appendix 2), the two haplotype groups were not correlated with the two different subspecies that we investigated. Instead, each haplotype group contained DNA sequences that were derived from both subspecies included in the analyses. This outcome is consistent with the results from Van Buskirk (2000), who also identified two main mitochondrial haplotype groups among *S. z. hippolyta* populations. McHugh and others (2013) placed sequences from Van Buskirk (2000) into a broader phylogenetic context by incorporating data from other *Speyeria* species and *S. zerene* subspecies. Like our analysis, this more extensive data set reiterated that *S. z. hippolyta* haplotypes were divided into two main groups, and also indicated that these groups contained haplotypes from many other *S. zerene* subspecies including *bremnerii*, *gloriosa*, *platina*, *sinope*, and *gunderi*.

McHugh and others (2013) outlined numerous different hypotheses that account for the discrepancy between subspecies identities and molecular phylogenies. Primary mechanisms distill to (1) recent evolution and insufficient time for lineage sorting, and (2) recent hybridization of previously isolated groups. Regardless of the specific mechanism, there are two primary implications of these results. First, these results suggest that mitochondrial DNA sequences cannot be used to document subspecies identities. Thus, if any new *S. zerene* populations are discovered in the current geographic range of *S. z. hippolyta*, careful morphological analyses will be required to properly assign individuals to appropriate subspecific units. Such analyses will be especially important given the disjunct distribution of *S. z. hippolyta* and the subtle morphological and life history variation that distinguishes among *S. zerene* subspecies (Grey and Moeck, 1962; McCorkle and Hammond, 1988). Second, the absence of clear genetic signals that distinguish among subspecies has led some to advocate for the elimination of this taxonomic rank (Zink, 2004; Zink and others, 2013). Others, however, recognize that multiple different criteria (genetics, morphology, life history, etc.) may provide a basis for justifying the existence of different subspecies (Haig and others, 2006), with a primary emphasis being on the existence of characters or attributes that allows for reliable diagnosis of different subspecific groups (Patten and Unitt, 2002).

Acknowledgments

We thank Anne Walker and Gary Falxa, U.S. Fish and Wildlife Service, for providing samples for use in the analysis.

Referenced Cited

Andersen, M.J., Arnold, M., Barclay, E., Myers, L., Shepherdson, D., and Sullivan, E., 2010, Husbandry manual—Oregon silverspot *Speyeria zerene hippolyta*: Oregon Zoo, Portland, Oregon.

Darriba, D., Taboada, G., Doallo, R., and Posada, D., 2012, jModeltest 2—More models, new heuristics and parallel computing: Nature Methods, v. 9, p. 772.

Excoffier, L., and Lischer, H.E.L., 2010, Arlequin suite version 3.5—A new series of programs to perform population genetics analyses under Linux and Windows: Molecular Ecology Resources, v. 10, p. 564–567.

Grey, L.P., and Moeck, A.H., 1962, Notes on overlapping subspecies—I, An example in *Speyeria zerene* (Nymphalidae): Journal of the Lepidopterists' Society, v. 16, p. 81–97.

Haig, S.M., Beever, E.A., Chambers, S.M., Draheim, H.M, Dugger, B.D., Dunham, S., Elliot-Smith, E., Fontaine, J.B., Kesler, D.C., Knaus, B.J., Lopes, I.F., Loschl, P., Mullins, T.D., and Sheffield, L.M., 2006, Taxonomic considerations in listing subspecies under the U.S. Endangered Species Act: Conservation Biology, v. 20, p. 1,584–1,594.

Hamm, C.A., Aggarwal, D., and Landis, D.A., 2010, Evaluating the impact of non-lethal DNA sampling on two butterflies, *Vanessa cardui* and *Satyrodes eurydice*: Journal of Insect Conservation 14, p. 11–18.

Huelsenbeck, J.P., and Ronquist, F., 2001, MRBAYES—Bayesian inference of phylogeny: Bioinformatics, v. 17, p. 754–755.

Jackson, J.A., Patenaude, N.J., Carroll, E.L., and Baker, C.S., 2008, How few whales were there after whaling?—Inference from contemporary mtDNA diversity: Molecular Ecology, v. 17, p. 236–251.

Kalinowski, S.T., 2004, Counting alleles with rarefaction: private alleles and hierarchical sampling designs: Conservation Genetics, v. 5, p. 539–543.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., and Drummond, A., 2012, Geneious Basic—An integrated and extendable desktop software platform for the organization and analysis of sequence data: Bioinformatics, v. 28, p. 1,647–1,649.

Keyghobadi, N., Crawford, L.A., and Maxwell, S.A., 2009, Successful analysis of AFLPs from non-lethally sampled wing tissues in butterflies: Conservation Genetics, v. 10, p. 2,021–2,024.

McCorkle, D.V., and Hammond, P.C., 1988, Biology of *Speyeria zerene hippolyta* (Nymphalidae) in a marine-modified environment: Journal of the Lepidopterists' Society, v. 42, p. 184–195.

McHugh, A., Bierzychudek, P., Greever, C., Marzulla, T., VanBuskirk, R., and Binford, G., 2013, A molecular phylogenetic analysis of *Speyeria* and its implications for the management of the threatened *Speyeria zerene hippolyta*: Journal of Insect Conservation, v. 17, p. 1,237–1,253.

Miller, M.P., Haig, S.M., Gratto-Trevor, C.L., and Mullins, T.D., 2010, Subspecies status and population genetic structure in piping plover (*Charadrius melodus*): The Auk, v. 127, p. 57–71.

Patten, M.A., and Unitt, P., 2002, Diagnosability versus mean differences of sage sparrow subspecies: The Auk, v. 119, p. 26–35.

Pritchard, J.K., Stephens, M., and Donnelly, P., 2000, Inference of population structure using multilocus genotype data: Genetics, v. 155, p. 945–959.

Ronquist, F., and Huelsenbeck, J.P., 2003, MrBayes 3—Bayesian phylogenetic inference under mixed models: Bioinformatics, v. 12, p. 1,572–1,574.

Schultz, C.B., Henry, E., Carleton, A., Hicks, T., Thomas, R., Potter, A., Collins, M., Linders, M., Fimbel, C., Black, S., Anderson, H.E., Diehl, G., Hamman, S., Gilbert, R., Foster, J., Hayes, D., Wilderman, D., Davenport, R., Steel, E., Page, N., Lilley, P.L., Heron, J., Kroeker, N., Webb, C., and Reader, B., 2011, Conservation of prairie-oak butterflies in Oregon, Washington, and British Columbia: Northwest Science, v. 85, p. 361–388.

Simon, C., Frati, F., Breckenbach, A., Crespi, B., Liu, H., and Flook, P., 1994, Evolution, weighting and phylogenetic utility of mitochondrial gene sequences, and compilation of conserved PCR primers: Annals of the Entomological Society of America, v. 87, p. 651–701.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S., 2011, MEGA5—Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony method: Molecular Biology and Evolution, v. 28, p. 2,731–2,739.

U.S. Fish and Wildlife Service, 1980, Listing the Oregon silverspot butterfly as a threatened species with critical habitat: Federal Register, v. 45, p. 44,935–44,939.

U.S. Fish and Wildlife Service, 2001, Oregon silverspot butterfly (*Speyeria zerene hippolyta*) revised recovery plan: U.S. Fish and Wildlife Service, Portland, Oregon, 113 p.

U.S. Fish and Wildlife Service, 2011, Oregon silverspot butterfly (*Speyeria zerene hippolyta*)—5-year review summary and evaluation: U.S. Fish and Wildlife Service, Portland, Oregon, 30 p.

Van Buskirk, R.W., Jr., 2000, Phylogeography, monitoring and conservation of *Speyeria zerene hippolyta*, the Oregon Silverspot Butterfly: Davis, University of California, Davis, Ph.D. dissertation.

Van Buskirk, R., 2010, Controlled propagation and reintroduction plan for the Oregon silverspot butterfly (*Speyeria zerene hippolyta*): Forest Grove, Oregon, Pacific University.

Villesen, P., 2007, FaBox—An online toolbox for fasta sequences: Molecular Ecology Notes, v. 7, p. 965–968.

Zhang, H., Miller, M.P., Yang, F., Chan, H.K., Gaubert, P., Ades, G., and Fischer, G.A., 2015, Genetic investigation of confiscated pangolin scales for conservation and illegal trade monitoring: Global Ecology and Conservation, v. 4, p. 414–422.

Zink, R.M., 2004, The role of subspecies in obscuring avian biological diversity and misleading conservation policy: Proceedings of the Royal Society of London B, v. 271, p. 561–564.

Zink, R.M., Groth, J.G., Vazquez-Miranda, H., and Barrowclough, G.F., 2013, Phylogeography of the California Gnatcatcher (*Polioptila californica*) using multilocus DNA sequences and ecological niche modeling—Implications for conservation: The Auk, v. 130, p. 449–458.

Appendix 1. Sample Information for Oregon Silverspot Butterfly Specimens Included in This Analysis

Appendix 1is a Microsoft® Excel file and is available for download at http://dx.doi.org/10.3133/ofr20161162.

Appendix 2. Haplotype Frequencies Among Populations of Oregon Silverspot Butterflies (*Speyeria zerene hippolyta*) and *S. z. gloriosa*

[CA, California; OR, Oregon]

Population	Haplotype																					
	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22
S. z. hippolyta																						
Mt. Hebo, OR	25																			1	1	1
Mt. Hebo (Zoo), OR	7																					
Cascade Head, OR	7															2						
Bray Point, OR	6							4									1	1	1			
Rock Creek, OR	9		5						1	1	1	1	1	1	1							
Rock Creek (Zoo), OR	5	1	2	1	1	1	1															
Lake Earl, CA			2																			
S. z. gloriosa																						
Cape Blanco, OR																						
Cape Ferrelo, OR			6																			
Six Rivers National Forest, CA																						
Rowdy Creek, CA																						
Total	59	1	15	1	1	1	1	4	1	1	1	1	1	1	1	2	1	1	1	1	1	1

Population	H23	H24	H25	H26	H27	H28	H29	H30	H31	H32	H33	H34	H35	H36	H37	H38	H39	H40	H41	H42	Total
S. z. hippolyta																					
Mt. Hebo, OR	1	1																			30
Mt. Hebo (Zoo) , OR																					7
Cascade Head, OR										1											10
Bray Point, OR																					13
Rock Creek, OR																					21
Rock Creek (Zoo), OR																					12
Lake Earl, CA			1	8	1	4	1	1	1												19
S. z. gloriosa																					
Cape Blanco, OR											4			8	1	1					14
Cape Ferrelo, OR													1				1			1	9
Six Rivers National Forest, CA											2	1									3
Rowdy Creek, CA											3							2	1		6
Total	1	1	1	8	1	4	1	1	1	1	9	1	1	8	1	1	1	2	1	1	144

Publishing support provided by the U.S. Geological Survey
Science Publishing Network, Tacoma Publishing Service Center

For more information concerning the research in this report, contact the
Director, Forest and Rangeland Ecosystem Science Center
U.S. Geological Survey
777 NW 9th St., Suite 400
Corvallis, Oregon 97330
http://fresc.usgs.gov/

ISSN 2331-1258 (online)
http://dx.doi.org/10.3133/ofr20161162