PHYLOGENETIC AND POPULATION GENETIC ANALYSIS OF THE HUMBOLDT'S FLYING SQUIRREL USING HIGH-THROUGHPUT SEQUENCING DATA

By

Stella Yuan

A Thesis Presented to

The Faculty of Humboldt State University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Biology

Committee Membership

Dr. Catalina Cuellar-Gempeler, Committee Chair

Dr. Melissa TR Hawkins, Committee Member

Dr. Barbara Clucas, Committee Member

Dr. Nicholas Kerhoulas, Committee Member

Dr. Paul Bourdeau, Program Graduate Coordinator

May 2020

ABSTRACT

PHYLOGENETIC AND POPULATION GENETIC ANALYSIS OF THE HUMBOLDT'S FLYING SQUIRREL USING HIGH-THROUGHPUT SEQUENCING DATA

Stella Yuan

The intraspecific genetic variation and diversity within the Humboldt's flying squirrel (*Glaucomys oregonensis*) has not yet been characterized despite its elevation to full species in 2017. The San Bernardino flying squirrel (*G. o. californicus*) is thought to be the southernmost population of *G. oregonensis* and is restricted to the San Bernardino and San Jacinto Mountains in California, but recent surveys indicate they have been extirpated from the latter locality. In order to provide baseline genetic data across the geographic range of *G. oregonensis*, I had the following objectives: 1) investigate the intraspecific molecular variation in *G. oregonensis* with a focus on the subspecies distributed in California; 2) evaluate the genetic diversity within *G. o. californicus*; 3) estimate if gene flow is occurring between the rest of the species and *G. o. californicus*. Population genetic and phylogenetic analyses, incorporating nine microsatellite loci and the partial or entire mitochondrial cytochrome-*b* gene, were performed on a total of 147 samples (tissue, hair, and museum specimen) using the Illumina high-throughput sequencing (HTS) platform; thereby bioinformatically coding alleles based on read count. My results support previously published work describing a south to north colonization of

the species after the Last Glacial Maximum and highlight the genetic distinctiveness of *G. o. californicus*. The ensuing data from this study contributes valuable information toward understanding the genetic diversity within *G. oregonensis*, provides material to inform future conservation decisions for *G. o. californicus*, and has novel implications for future HTS microsatellite genotyping.

ACKNOWLEDGEMENTS

I thank everyone on my committee for their unending patience and support; Stephanie Steffen and Liz Weaver for their administrative assistance, Dave Baston for allowing me to use the Core Research Facility, Nancy McInerney and Katie Murphy for performing the sequencing runs, and all of the undergraduate and fellow graduate students who put in many hours of their own time assisting my research: Clare O'Connell, Michael Kiso, Eric Malekos, and Jack Lemke. A special thank you as well to Clark Winchell with the U.S. Fish and Wildlife Service for help locating specimens, providing historical and ecological information, and kick-starting my interest in flying squirrels. I am also incredibly grateful to the Big Bear Alpine Zoo, United States Geological Survey, Humboldt State University Vertebrate Museum, Museum of Vertebrate Zoology at University of California Berkeley, Natural History Museum of Los Angeles, University of Kansas Biodiversity Institute and Natural History Museum, University of Michigan Museum of Zoology, University of California Los Angeles Dickey Bird and Mammal Collection, and Scott Tremor at the San Diego Natural History Museum for providing the samples used in this project. And last but not least, for their financial support I thank the American Society of Mammalogists Grants-in-Aid of Research program, Sigma Xi Grants in Aid of Research program, and the Humboldt State University Department of Biology Master's Student Grant program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Flying squirrels are nocturnal rodents in the family Sciuridae and are characterized by the furred patagium that extends from their wrists to their ankles. This adaptation is what allows them to glide from tree to tree and seemingly "fly". Currently, there are 15 recognized genera of flying squirrels, but only one occurs in North America: *Glaucomys* (Thorington and Hoffmann 2005). *Glaucomys* is composed of three species: *G. volans* (southern flying squirrel), *G. sabrinus* (northern flying squirrel), and *G. oregonensis* (Humboldt's flying squirrel) (Arbogast et al. 2017). *G. volans* is found in eastern North America from southern Quebec to Florida and has isolated populations in Mesoamerica (Dolan and Carter 1977), while *G. sabrinus'* range covers most of Canada and Alaska with some populations extending into the contiguous United States (Wells-Gosling and Heaney 1984). The third species, *G. oregonensis*, resides along the Pacific Coast of British Columbia and the United States (Washington, Oregon, and California), and was only revealed to be a reproductively isolated, distinct species of flying squirrel in 2017 (Arbogast et al. 2017). This was surprising given that the estimated date of divergence between *G. sabrinus* and *G. volans* is 1.07 million years ago, while *G. oregonensis* diverged from the both of them ~1.32 million years ago (Arbogast et al. 2017).

The complex evolutionary history of *G. oregonensis* likely confounded their recognition as a separate species prior to 2017, making them a cryptic species: a species with a discrete evolutionary and genetic lineage, but no characterized morphological

differences to separate them from *G. sabrinus*. During the Pleistocene Epoch, repeated glacial cycles of global warming and cooling forced flora and fauna populations into refugia (areas containing sufficient suitable habitat) and created vicariant barriers to gene flow (Shafer et al. 2010). The types of habitat required by *Glaucomys* (deciduous hardwood forests for *G. volans* and coniferous/mixed coniferous-deciduous forests for *G. sabrinus* and *G. oregonensis*) were found south of their present distributions (Arbogast 2007; Smith 2007). Thus, all three species settled in refugia south of the Laurentide and Cordilleran ice sheets (Arbogast 1999; Brunsfeld et al. 2001; Shafer et al. 2010). *G. sabrinus* resided in a southeastern U.S. refugium, *G. volans* resided in a Central American and southeastern U.S. refugium, and *G. oregonensis* remained in a Pacific Coast refugium similar to its contemporary distribution (Arbogast 1999; Arbogast 2007). As the ice receded after the Last Glacial Maximum, each species expanded northward from their respective refugia to their current-day ranges (Arbogast 2007; Smith 2012; Arbogast et al. 2017). However, because the Laurentide ice sheet retreated more quickly than the Cordilleran ice sheet, *G. sabrinus* was able to colonize a larger area than *G. oregonensis,* and the two species eventually united in Washington and British Columbia where they now live in sympatry (Arbogast 1999; Arbogast et al. 2017). There has been no evidence of hybridization between the two to date, and it is nearly impossible to tell them apart morphologically, so DNA analysis is required to discriminate between the two species (Arbogast et al. 2017).

The present-day genetic structure and diversity of *G. oregonensis* is also likely linked to historical glacial cycles, which is commonly observed in mammalian taxa

(Arbogast et al. 2001; Shafer et al. 2010; Bozarth et al. 2011; Puckett et al. 2015; Sawyer and Cook 2016). Comparative phylogeographic studies have revealed an overall decrease in genetic variation from south to north in *G. volans* and *G. sabrinus* populations as a result of the northward post-glacial colonization (Arbogast et al. 2005; Petersen and Stewart 2006; Garroway et al. 2011). For example, the *G. volans* Mesoamerican populations contain higher mitochondrial nucleotide diversity as a whole compared to the eastern North American populations, which may have undergone one or several bottleneck and founder events (Kerhoulas and Arbogast 2010). For *G. oregonensis* the overall expansion was northward as well, but the particularly tumultuous glacial-interglacial cycles in the Pacific Northwest may have isolated ancestral populations in multiple, independent refugia, potentially leading to more genetic substructure (Arbogast 1999; Owen et al. 2003; Arbogast et al. 2017). This would not be uncommon for forest-obligate lineages (Brunsfeld et al. 2001; Shafer et al. 2010; Hope et al. 2016).

At present, there are eight named subspecies of *G. oregonensis* (Wells-Gosling and Heaney 1984; Arbogast et al. 2017), six of which are distributed in California (Fig. 1), but to date there have not been any studies evaluating the intraspecific genetic variation within *G. oregonensis.* Amongst all of the California subspecies populations, the San Bernardino population (*G. o. californicus*) stands out as they are thought to be the southernmost limit of *G. oregonensis*, and as such they may be a reservoir of genetic diversity left behind after glacial expansion as observed in *G. sabrinus* and *G. volans* (Arbogast et al. 2005; Petersen and Stewart 2006; Weigl 2007; Garroway et al. 2011). From limited sampling, Arbogast et al. (2017) showed that *G. o. californicus* contained

four unique mitochondrial haplotypes, but their study was intended as an overall phylogeographic study, and thus included only seven individuals which is not ideal for population level studies (Hale et al. 2012).

The San Bernardino subspecies is also noteworthy as it is an insular population that was historically restricted to the San Bernardino and San Jacinto Mountains, but recent surveys suggest they have been extirpated from the latter locality (Wolf 2010). The subspecies was petitioned to be listed under the Endangered Species Act in 2010 but was ultimately rejected (FWS 2016a) and is currently listed as a subspecies of special concern by the California Department of Fish and Wildlife (CNDDB 2019). Literature about this subspecies is scarce, plus they are faced with a smattering of anthropogenic threats including climate change and habitat fragmentation (Weigl 2007; Smith 2012). As individuals of *G. o. californicus* likely have no connectivity to other populations of flying squirrels (the nearest *G. o. lascivus* population is approximately 220 km away), these threats may have predisposed this population to reduced genetic diversity from forces such as genetic drift and/or founder effects. This is problematic as decreased genetic variation is correlated with a heightened risk of extinction, disease, and introgression, and may reduce the long-term viability of this subspecies (Keller and Waller 2002; Spielman et al. 2004; Marsden et al. 2016; Thatte et al. 2018; Ujvari et al. 2018). Depauperate genetic diversity due to low levels of gene flow and small effective population size have been documented in other disjunct *G. sabrinus* subspecies (Bidlack and Cook 2001, 2002; Arbogast et al. 2005), and in fact one of these subspecies, *G. s. coloratus* (Carolina

northern flying squirrel), is listed as endangered under the U.S. Endangered Species Act (Currie and Cameron 2013).

In order to bridge the knowledge gap in *G. oregonensis*, I conducted a study employing molecular markers (mitochondrial cytochrome-*b* and nuclear microsatellites) with three main objectives: 1) analyze the intraspecific genetic variation in *G. oregonensis* with a focus on the subspecific differences in California; 2) evaluate *G. o. californicus* genetic diversity; 3) estimate if gene flow is occurring between the rest of *G. oregonensis* and *G. o. californicus*. The decision to refrain listing of *G. o. californicus* was made prior to any genetic evaluation of this subspecies and by assessing their genetic diversity we can generate crucial data to help inform future conservation decisions and state management of this population. Moreover, their genetic diversity may prove to be crucial to the rest of the species because *G. o. californicus* was originally a subspecies of the widely distributed *G. sabrinus*, but is now known to be included within *G. oregonensis* which inhabit a reduced range in comparison to *G. sabrinus*. The genetic distinctiveness of several *G. sabrinus* subspecies have already been reported as essential to the species genetic variability (Bidlack and Cook 2001, 2002; Arbogast 2007; Kerhoulas and Arbogast 2010; Arbogast et al. 2017). In addition, *G. oregonensis* subspecies are defined based solely on morphology (Howell 1918). However, a combination of genetic and morphological evidence for subspecific delineations would be more appropriate (Balakirev et al. 2017; Jayat et al. 2017) since at times the genetic restructuring of subspecies can beget consequences for conservation (Puckett et al. 2015; Butynski and de Jong 2017; Gamage et al. 2017).

METHODS

Sample collection

A total of 147 *G. oregonensis* samples were obtained for this project from across the U.S. Pacific West Coast with a focus on the state of California (Fig. 1, Appendix A). The samples collected include six out of eight *G. oregonensis* subspecies; and of the 147 samples, 51 were from *G. o. californicus* individuals. Subspecies classifications were based on geographic location, morphology, and/or designated by collectors. The U.S. Geological Survey (USGS), Humboldt State University Vertebrate Museum (HSU), Museum of Vertebrate Zoology at University of California Berkeley (MVZ), Natural History Museum of Los Angeles (LACM), University of Kansas Biodiversity Institute and Natural History Museum (KU), University of Michigan Museum of Zoology (UMMZ), and the University of California Los Angeles Dickey Bird and Mammal Collection (UCLA) provided 97 museum specimen samples (clips of fur, bone, or adherent muscle tissue) and 44 tissue samples; the San Diego Natural History Museum provided one hair sample from the field and two degraded tissue samples; and the Big Bear Alpine Zoo contributed three hair samples under an approved IACUC permit (HSU no. 18/19. B.40-A) from animals housed in captivity. Total genomic DNA from hair, degraded tissue, and museum samples were isolated using Qiagen QIAamp DNA Mini Kits (Qiagen, Valencia, CA) in a lab designated for degraded and ancient DNA, while DNA from tissue samples were processed in a designated, high quality DNA facility

using Qiagen DNeasy Blood and Tissue Kits (Qiagen, Valencia, CA). Hair samples were collectively referred to and treated as degraded museum samples (Appendix G).

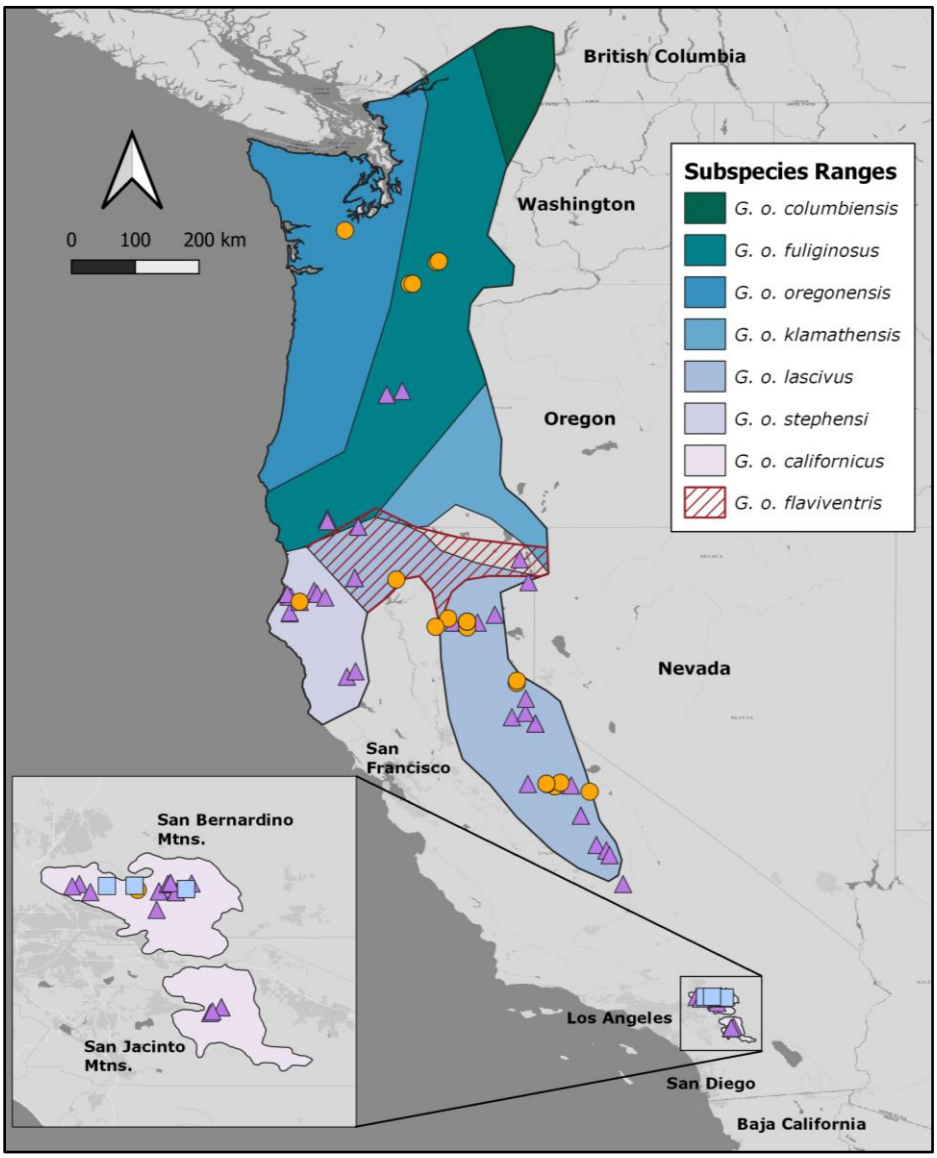


Figure 1. Map of sample collection locations across the geographic distribution of *Glaucomys oregonensis*. A total of 147 samples were obtained and include six out of the eight total *G. oregonensis* subspecies (51 *G. o. californicus*, 65 *G. o. lascivus*, 13 *G. o. fuliginosus*, 11 *G. o. stephensi*, six *G. o. flaviventris*, one *G. o. oregonensis*). Orange circles indicate tissue samples (n = 44), purple triangles indicate museum samples (n = 97), and blue squares indicate hair and degraded tissue samples (n = 6).

Polymerase Chain Reaction (PCR) amplification

The entire (~1140 bp) cytochrome-*b* gene (cyt-*b* hereafter) was amplified in PCR using universal primers (Irwin et al. 1991; Oshida et al. 2000) for all tissue samples except: HSU-7616, HSU-8184, HSU-8188, HSU-8482, HSU-VM 3095, HSU-7615, and HSU-7747. Attempts to amplify the entire gene continuously failed for the aforementioned tissue samples, so instead an approximately 300 bp region of the cyt-*b* gene was amplified with the primer L14724 (Irwin et al. 1991) and a newly designed reverse primer, GOR_R1. The GOR_R1 primer was designed from a *G. o. californicus* GenBank sequence (AF063060) using Primer3 (Untergasser et al. 2012) and was mapped to published sequences and visualized in Geneious Prime v 2020.0.4 (Kearse et al. 2012). Primers were ordered from IDT at 35.4 nmol concentration, solubilized, and diluted to 100 µM prior to PCR. An approximately 300 bp region of the cyt-*b* gene was also amplified in museum samples using L14724 and GOR_R1. Museum specimens contain highly degraded endogenous DNA which can be contaminated and in low concentration due to the preparation and age of the specimens (Hawkins et al. 2016a). Therefore, only short DNA fragments can be amplified from museum samples (Hawkins et al. 2016a). Two other primers, Gvo_R3 and Gvo_R4 (Kerhoulas and Arbogast 2010), were tested but did not amplify in PCR, probably because they were designed from *G. volans* sequences which resulted in binding site incongruence. Singleplexed PCR was performed in 16 µl reactions containing 2.0 µl of DNA template, 4.5 µl of ddH2O, 0.5 µl of each primer, and 8.5 µl of DreamTaq Green PCR Master Mix (Thermo Fisher Scientific Inc.,

Waltham, MA) with the following thermocycler profile: 95°C for 1 min; 30 cycles (tissues) or 35 cycles (museum samples) of 95°C for 30 sec, 45°C for 30 sec, 72°C for 30 sec; 72°C for 5 mins. Successful PCR amplifications were replicated twice in each sample to ensure validity, and a 1-1.5% agarose gel was run to visualize PCR products.

All samples were also genotyped at 11 polymorphic microsatellite loci using primers previously characterized for *G. sabrinus*: GS-1, GS-2, GS-4, GS-8, GS-10, GS-13, GS-16 (Zittlau et al. 2000), GLSA-12, GLSA-22, GLSA-52, and GLSA-65 (Table 1, Kiesow et al. 2011). Singleplexed PCR was performed in 16 µl reactions containing 2.0 µl of DNA template, 4.5 µl of ddH2O, 0.5 µl of each primer, and 8.5 µl of DreamTaq Green PCR Master Mix. When amplifying microsatellites with the GS-2 and GS-4 primers, 0.3 µl of ddH2O was substituted for bovine serum albumin (BSA, New England Biolabs, 12 mg) to promote reaction specificity. For all samples a touchdown PCR profile was used: 95°C for 1 min; 2 cycles of 95°C for 15 sec, 60°C for 30 sec, 72°C for 45 sec; 2 cycles changing 60°C to 58°C; 2 cycles changing 58°C to 54°C; 2 cycles changing 54°C to 52°C; 35 cycles of 95°C for 15 sec, 50°C for 30 sec, 72°C for 45 sec; 72°C for 5 mins. Successful PCR amplifications were replicated twice in tissue samples and three times in all museum samples in order to alleviate potential allelic dropout in degraded samples (De Barba et al. 2017). A summary of the performance and rates of allelic dropout specifically for museum specimens were analyzed separately for this study (Yuan et al. *in prep*). A 1.5% agarose gel was run to visualize PCR products and confirm fragments of the desired size range were recovered.

DNA sequencing

Illumina short read sequencing technologies are aimed at sequencing fragments up to 300-500 bp in length, so fragments larger than 300 bp must be sonicated or sheared before library preparation. After amplification, cyt-*b* PCR products that were over 300 bp were fragmented with NEBNext dsDNA Fragmentase (New England Biolabs, Ipswich, MA) at 37°C for 30 min (6 µl of each PCR replicate pooled together for a total of 12 µl per sample), cleaned with 1X Streptavidin magnetic (SPRI) beads (KAPA Pure Beads, KAPA Biosystems), and run on a 1% agarose gel to evaluate shearing results. Microsatellite and cyt-*b* replicates derived from the same individual were then pooled in equimolar concentrations via quantification from a NanoDrop Lite spectrophotometer and prepared for Illumina sequencing utilizing KAPA Illumina Library Prep Kits (KAPA Biosystems) with dual indexed iTru style indices (Glenn et al. 2019). The protocol closely adhered to the KAPA Illumina Library Prep Kit instructions but used a ¼ recipe for all enzyme reactions following Hawkins et al. (2016b); and PCR products were SPRI cleaned (KAPA Pure Beads, Rohland and Reich 2012) at a 1.5X concentration to remove residual primers and non-target fragments outside the desired size range. Libraries were amplified in 25 µl reactions consisting of 1.25 µl of each iTru adapter, 2.5 µl ddH2O, 7.5 µl of adapter-ligated DNA, and 12.5 µl of KAPA HiFi HotStart ReadyMix. The thermocycler conditions for library amplification were: 98°C for 45 sec; 10 cycles (tissues) or 14 cycles (museum samples) of 98°C for 15 sec, 60°C for 30 sec, 72°C for 1 min; 72°C for 5 mins. After library preparation all SPRI cleaned (1.5X) products were

quantified on an ABI QuantStudio 3 quantitative PCR machine to determine exact sample concentration and pooled in equimolar ratios across all individuals. These samples were then combined with samples from other non-related projects to fill the estimated number of sequencing reads. Amplicon sequencing of all samples was performed on an Illumina MiSeq 2x300 PE version 3 chemistry and run at the Center for Conservation Genomics, Smithsonian Conservation Biology Institute, Washington DC. Fifty-one samples had to be re-sequenced at the Laboratory of Analytical Biology, National Museum of Natural History, Smithsonian Institution on an Illumina MiSeq using a 2x250 PE version 2 Nano kit due to exceptionally low read depth. The previously amplified PCR products were pooled again for each individual, library prepped using the same aforementioned protocols, and placed in the second sequencing run. All reads were trimmed and quality filtered (Phred score ≤ 20) using FastQC v 0.11.9 (Andrews 2010) and CutAdapt v 1.18 (Martin 2011) before being utilized in downstream applications. The basic CutAdapt command used was: cutadapt --report=minimal -q 20 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT -o {R1_output.fastq} -p {R2_output.fastq} {R1_input.fastq.gz} {R2_input.fastq.gz}.

Mitochondrial DNA (mtDNA) analysis

Cytochrome-*b* sequences were mapped (Bowtie2 v 2.3.0, Langmead and Salzberg 2012) to a previously published, complete CDS of a *G. sabrinus* cyt-*b* gene sequence

from GenBank (AF030390) in Geneious Prime v 2020.0.4 (Kearse et al. 2012). The samples HSU-7746 and MVZ-132650 were excluded in downstream phylogenetic analyses because they contained too many ambiguous sites, likely due to poor original DNA quality. The data were supplemented with two more *G. sabrinus* sequences from GenBank (AF011738 and AF359210) to construct a full (~1140 bp) and a 300 bp cyt-*b* alignment. Both alignments, with and without codon partitioning, were run in PartitionFinder v 2.1.1 (Lanfear et al. 2016) to find the best fit model of nucleotide substitution using the corrected Akaike Information Criterion (AICc). After alignments were manually inspected, phylogenetic trees were generated from them in Geneious Prime using Bayesian Inference (MrBayes v 3.2.7a; Ronquist et al. 2012) and maximum likelihood (PhyML v 3.3.20180621; Guindon et al. 2010). One million MCMC chains across two runs were performed with a heated chain at a temperature of 0.25 in MrBayes, and 500 bootstrap replications were run in PhyML. Lastly, a minimum spanning haplotype network (Bandelt et al. 1999) was constructed using PopART v 1.7 (Leigh and Bryant 2015) to visualize *G. oregonensis* genealogy and haplotype diversity was calculated using GenAlEx v 6.503 (Peakall and Smouse 2006, 2012).

Microsatellite DNA analysis

The CHIIMP v 0.3.1 (Barbian et al. 2018) pipeline was employed to generate microsatellite genotypes using high-throughput sequencing data because it was designed for degraded DNA sources (fecal samples). Figure 2 outlines the process used by

CHIIMP to call alleles (procured from Barbian et al. 2018). Genotyping occurred for all loci with 5 or more reads (counts.min = 5), sequences were considered potential alleles if they made up at least 5% of the filtered read count (fraction.min = 0.05), and all loci were assigned a length buffer of 20 bp. In order to ensure accurate genotypes were reconstructed, a subset of individuals had library preparation performed on each microsatellite amplification to assess levels of allelic dropout (Yuan et al. *in prep*).

After generation of genotypes, summary statistics including F_{ST}, heterozygosity (observed and expected), and deviations from Hardy-Weinberg equilibrium were calculated in GenAlEx v 6.503 (Peakall and Smouse 2006, 2012) for each locus. The program STRUCTURE v 2.3.4 (Pritchard et al. 2000) was run to evaluate *G. oregonensis* subpopulation structure and search for admixture between *G. oregonensis* and *G. o. californicus*. This program implements a Bayesian algorithm to assign individuals to genetic groups based on allele calls without a priori information. An initial test was performed to evaluate the most likely number of genetic clusters (K), which ran 5 iterations of each K (1-8 possible populations) at a burnin length of 10,000 followed by 100,000 MCMC repetitions after burnin. For the initial test, the admixture model was used without inferring sampling location as a prior, allele frequencies were correlated, and the probability of the data was computed. The ΔK method described in Evanno et al. (2005) as a part of STRUCTURE HARVESTER v 0.6.94 (Earl and vonHoldt 2012) was utilized to find the most likely K. Following the preliminary analysis, two additional runs were performed on both the admixture and no admixture model. All of the same parameters were kept except for run length, which was modified to 10 iterations with a

burnin of 200,000 and 1,000,000 MCMC repetitions for each K (1-6). After the runs were completed, STRUCTURE HARVESTER was again run on these datasets, as well as CLUMPAK v 1.1 (Kopelman et al. 2015), to achieve a consensus result across all runs.

Additionally, the program BOTTLENECK v 1.2.02 (Piry et al. 1999) was used to scan microsatellite genotypes for signatures of a recent population bottleneck in *G. o. californicus*. A one-tailed Wilcoxon signed rank test (Cornuet and Luikart 1996; Luikart and Cornuet 1998) was computed to assess heterozygosity excess, which can be found in populations directly following population reductions for approximately 5-15 generations. The two-phase mutation (TPM) model of microsatellite evolution, infinite allele model (IAM), and stepwise mutation model (SMM) were examined in runs with 10,000 replications. For the TPM model, the mutation variance was set to 30 and the proportion of SMM in TPM was set to 5% (Huang et al. 2002; Cristescu et al. 2010).

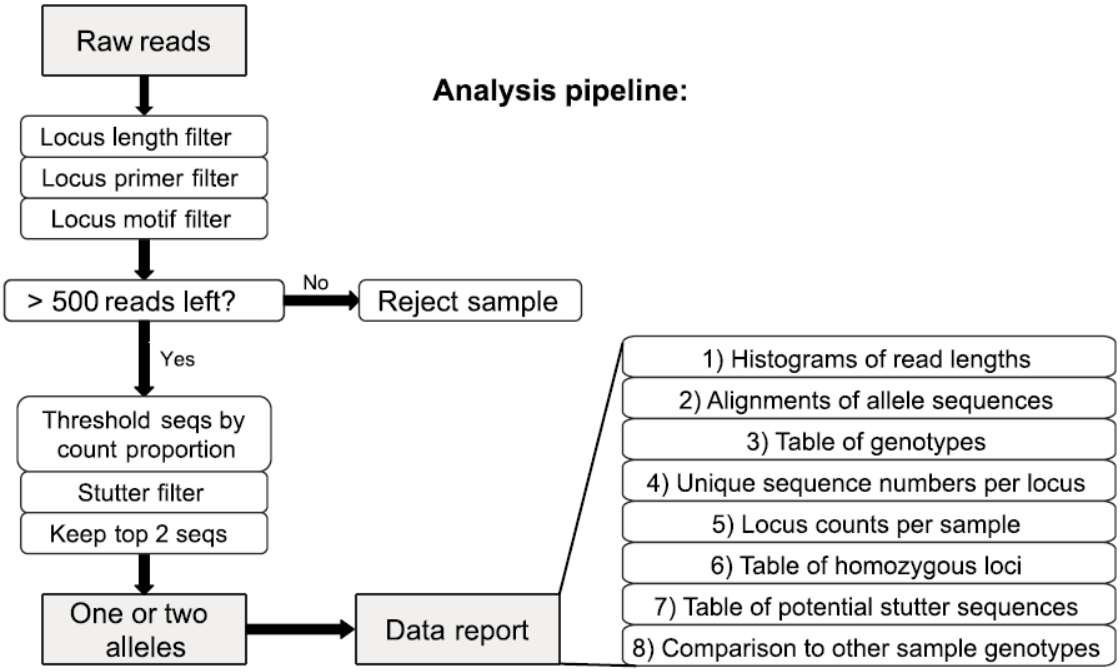


Figure 2. Flowchart illustrating the CHIIMP high-throughput sequencing bioinformatics pipeline used to genotype microsatellites (obtained from Barbian et al. 2018). Stutter was defined as sequences one motif shorter than the potential allele sequences and PCR artifacts were defined as sequences 1 bp shorter or longer than the potential allele sequences. A selection of CHIIMP output is available in Appendix D.

Table 1. List of primers used to amplify 11 microsatellite loci in 147 *Glaucomys oregonensis* samples. All of the GS primers were taken from Zittlau et al. (2000) and all of the GLSA primers were taken from Kiesow et al. (2011).

Locus	Sequence (5'- 3')	Repeat Motif	Size Range (bp)	Anneal Temp. (°C)
GS-1	F: GCTGCCCTCATTTTATCCCC R: GCTTCGTGTGTATATGTGTGTGTG	(GT)3(AT)2(GT)14	91–101	53.1
GS-2	F: AACATTCTCGCCACATCTAA R: CTACACCCCCAGCCCTACAA	(GT)16	96–124	F: 48.6 R: 55.8
GS-4	F: CTTCTTGAGTTGCTGGGGTGAC R: ATCTAAACCATAACACACACACA	(GT)15	104–112	F: 55.1 R: 49.3
GS-8	F: ATGCCATCTCCCCTCTC R: GCTGTGCTTCCAACCTGT	(GT)20	209–221	F: 50.6 R: 52.7
GS-10	F: CTATGCTGAGGAGGAGTGGTG R: CGTTTATGTGAAGAGCCTTG	(GT)18	189–205	F: 53.8 R: 48.4
GS-13	F: CTGGTCTCTTGAGTTAGGTG R: TATTCCTTTCTTCTCTCCTCCC	(GT)16	106–118	F: 49.1 R: 50.3
GS-16	F: AATGGAATAGGTATGAGTTGTC R: TGATGCTGCTTCTCTCTGG	(GT)15	92–108	F: 46.9 R: 51.4
GLSA-12	F: AGCATATGGAACCCCATATCAG R: GGCAAGATTTGTATCCTTGG	(GT)5A(TG)3TTT(GT)5	154–170	54.8
GLSA-22	F: CCTGAAAATGATGCATGTGG R: AGAGTAGGCTGTTCCTTTGAGG	(CA)15	168–188	54.8
GLSA-52	F: TCCATCCACAGTGTGTGAGC R: CCT GGA GTC CAC TCAAGCAT	(CA)16	210–254	57.5
GLSA-65	F: TTT GGG AAT TGA GGC TAT GG R: TTC ACA GTG ACA GCA GGT GAC	(GT)17	170–210	52.5

RESULTS

A total of 13,824,680 raw reads were produced from the two sequencing runs (240,917 reads mapped to the cyt-*b* gene and 8,306,496 reads mapped to microsatellites). A summary table of read depth per sample is provided in Appendix B.

Mitochondrial DNA results

Identical minimum spanning haplotype networks were recovered from the two cyt-*b* alignments (~1140 bp and ~300 bp, Fig. 3). Thirteen unique haplotypes were uncovered from the dataset, which included three GenBank sequences and 145 individuals sampled in this study (n = 148). One *G. sabrinus* sample, AF030390, grouped with *G. oregonensis* samples. According to GenBank, this individual was originally collected from Lewis County, Washington and consequently, is likely *G. oregonensis*. It is also evident from the haplotype network that *G. o. californicus* and *G. o. lascivus* do not form two discrete haplogroups but each contain a few unique haplotypes. Furthermore, in our dataset the haplotypes found in Oregon and Washington are not found in California (Fig. 4). But, due to the low sample size of individuals in Oregon and Washington, we refrain from over analyzing the results of these localities. Limited sampling of all subspecies except *G. o. lascivus* and *G. o. californicus* restricts the conclusions we can draw from the diversity statistics as well (Table 2). Nonetheless, it grants a peek into each subspecies' genetic variation. *G. o. fuliginosus* had the highest

haplotype diversity (H = 0.625), while *G. o. lascivus* had the highest number of haplotypes (h = 5) and private haplotypes (h_p = 3). Notably, *G. o. californicus* had the lowest haplotype diversity (H = 0.076) and two private haplotypes, one of which is from a San Jacinto individual (UCLA-7487).

Based on results from PartitionFinder, the best-fit model of nucleotide substitution was GTR+G for both the full and 300 bp alignment without codon partitioning. This information was used to build trees incorporating maximum likelihood estimation and Bayesian inference (Appendix C). Within *G. oregonensis*, branches were not well supported in any of the trees likely due to the limited base pairs sequenced. The species level divergence between *G. sabrinus* and *G. oregonensis*, however, was well supported (100 in maximum likelihood bootstrap support and 1.0 posterior probability).

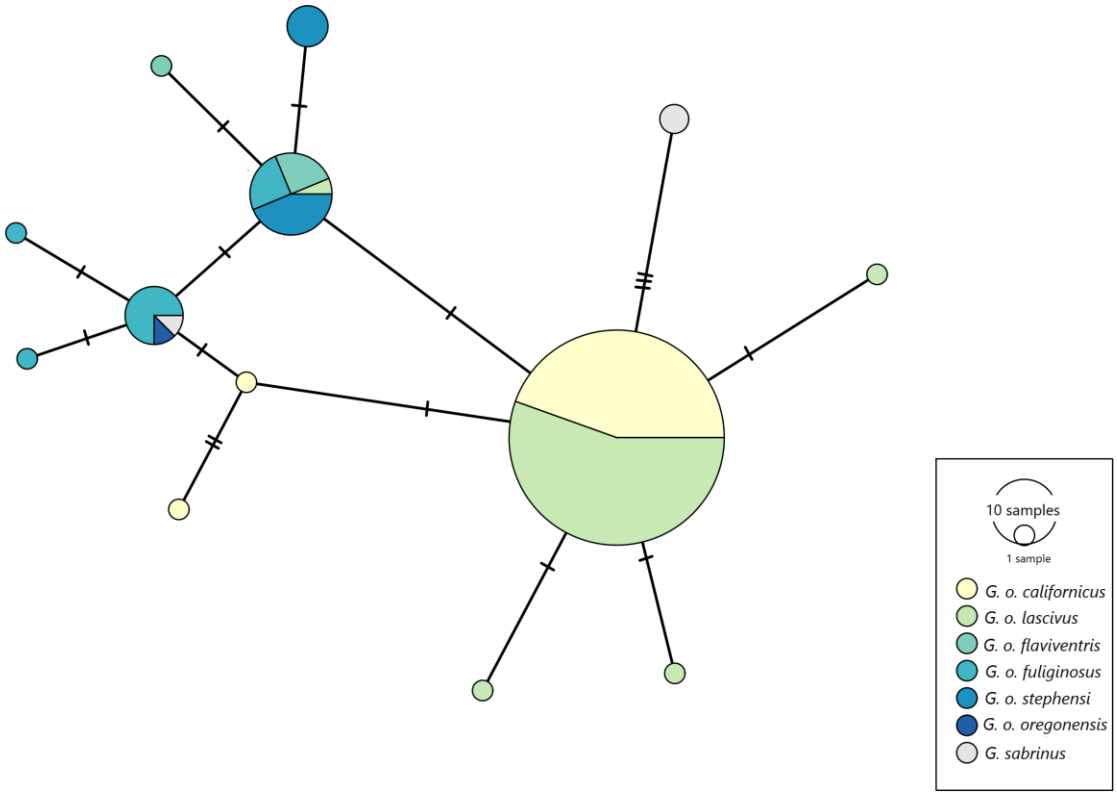


Figure 3. Cytochrome-*b* minimum spanning network constructed based on ~300 bp for 108 *Glaucomys oregonensis* samples, ~1140 bp for 37 *G. oregonensis* samples, and ~1140 bp for three *G. sabrinus* samples (n=148). Circle sizes are proportional to the number of samples and each dash mark represents a nucleotide substitution between the recovered haplotypes. From the original 147 *G. oregonensis* samples used in this study, two were excluded in phylogenetic analyses because they contained too many ambiguous sites (HSU-7746 and MVZ-132650).

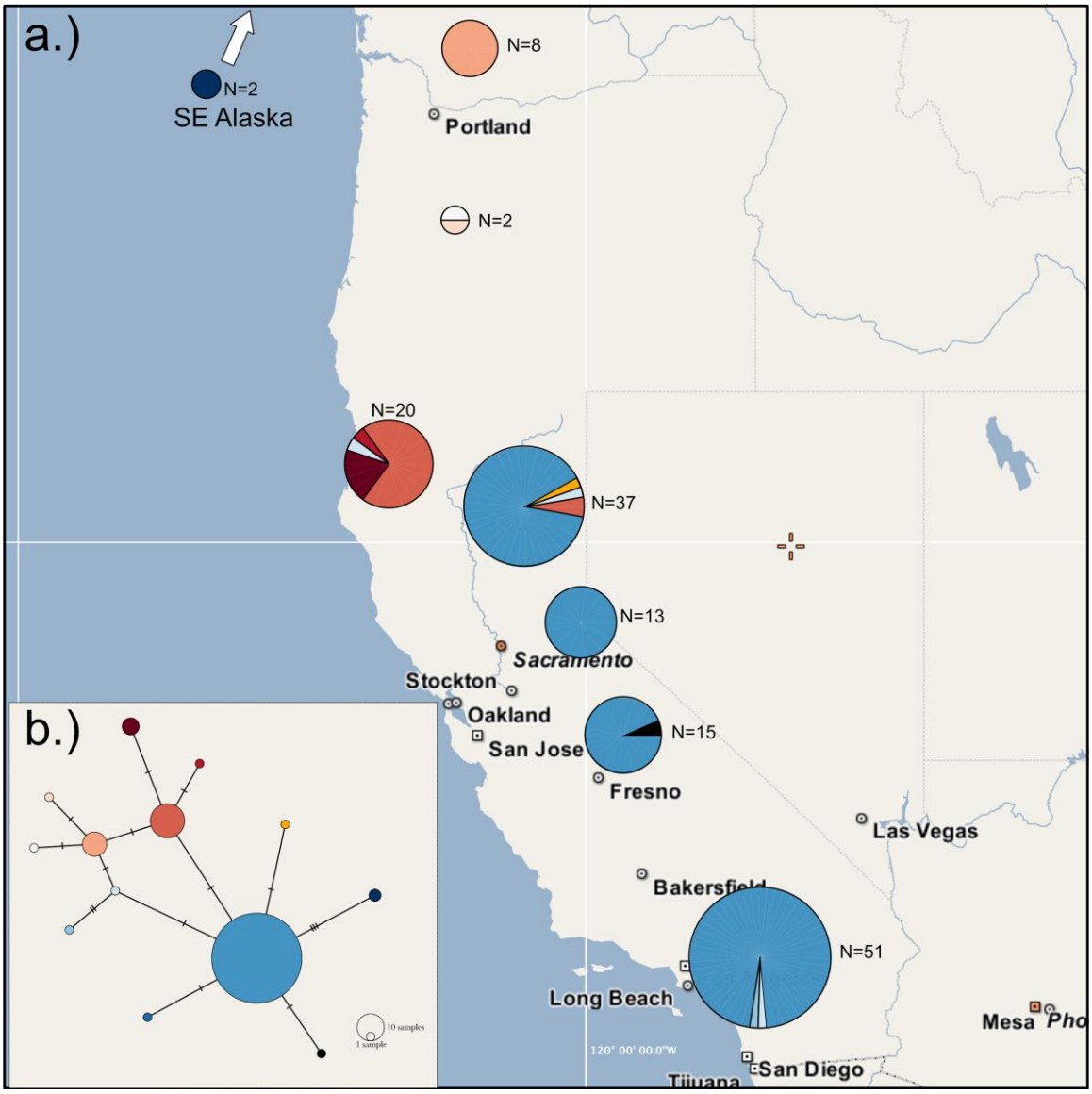


Figure 4. a.) Distribution of *Glaucomys oregonensis* and *G. sabrinus* (N=2 from SE Alaska) cytochrome-*b* haplotypes. Each haplotype is designated with a different color and circle sizes are proportional to the number of samples. **b.)** Depicts the minimum spanning network and relationship between the recovered haplotypes. Each dash mark represents a nucleotide substitution between the recovered haplotypes.

Table 2. Mitochondrial genetic diversity of 145 *Glaucomys oregonensis* samples separated by subspecies. Summary statistics were calculated in GenAlEx (Peakall and Smouse 2006, 2012). From the original 147 *G. oregonensis* samples used in this study, two were excluded in phylogenetic analyses because they contained too many ambiguous sites (HSU-7746 and MVZ-132650). The abbreviations listed are as follows: number of individuals (n), number of haplotypes (h), number of private haplotypes (h_p), and haplotype diversity (H).

Population	**n**	**h**	**h_p**	**H**
G. o. californicus	51	3	2	0.076
G. o. flaviventris	5	2	1	0.320
G. o. fuliginosus	12	4	2	0.625
G. o. lascivus	65	5	3	0.118
G. o. oregonensis	1	1	0	0.000
G. o. stephensi	11	2	1	0.463

Microsatellite results

Nine loci were successfully genotyped in CHIIMP with 1260 genotypes present (63 missing genotypes) across all samples and loci. A selection of outputs from CHIIMP are provided in Appendix D. The two loci with imperfect repeats, GS-1 and GLSA-12, were taken out of the dataset as they could not be analyzed by the program, leaving only dinucleotide microsatellite data for subsequent analyses. Overall, *G. o. lascivus* and *G. o. californicus* are quite diverse when examining mean allelic richness (A = 9.333 and A = 6.222 respectively, Table 3 and 4). Both populations were also out of

HWE at 7 out of 9 loci. Appendix E summarizes the genetic variation found in the other four sampled *G. oregonensis* subspecies.

Table 3. Nuclear microsatellite summary statistics for *Glaucomys oregonensis californicus* at nine loci calculated in GenAlEx (Peakall and Smouse 2006, 2012). The abbreviations listed are as follows: number of individuals successfully genotyped at that locus (n), total number of alleles (A), number of private alleles (A_p), observed heterozygosity (H_o), expected heterozygosity (H_e), fixation index (F), and Hardy-Weinberg Equilibrium deviation (HWE). Significant P-values are displayed in bold

Locus	**n**	**A**	**A_p**	**H_o**	**H_e**	**F**	**HWE**
GS-2	51	4	1	0.078	0.076	-0.028	1.000
GS-4	49	9	1	0.612	0.721	0.151	**<0.001**
GS-8	46	6	1	0.478	0.418	-0.144	**<0.001**
GS-10	49	6	1	0.918	0.574	-0.599	**0.004**
GS-13	50	8	2	0.900	0.597	-0.507	**<0.001**
GS-16	48	4	0	0.896	0.544	-0.645	**<0.001**
GLSA-22	50	6	1	0.980	0.676	-0.450	**<0.001**
GLSA-52	33	5	0	0.485	0.582	0.167	**<0.001**
GLSA-65	49	8	1	0.714	0.589	-0.214	0.427
Mean	47.222	6.222	-	0.674	0.531	-0.252	-
SE	1.839	0.596	-	0.098	0.063	0.104	-

Table 4. Nuclear microsatellite summary statistics for *Glaucomys oregonensis lascivus* at nine loci calculated in GenAlEx (Peakall and Smouse 2006, 2012). The abbreviations listed are as follows: number of individuals successfully genotyped at that locus (n), total number of alleles (A), number of private alleles (A_p), observed heterozygosity (H_o), expected heterozygosity (H_e), fixation index (F), and Hardy-Weinberg Equilibrium deviation (HWE). Significant P-values are displayed in bold.

Locus	**n**	**A**	**A_p**	**H_o**	**H_e**	**F**	**HWE**
GS-2	65	9	2	0.231	0.229	-0.009	0.181
GS-4	65	9	2	0.677	0.755	0.103	**0.001**
GS-8	65	8	2	0.277	0.427	0.351	**<0.001**
GS-10	65	12	3	0.877	0.799	-0.098	**<0.001**
GS-13	65	13	5	0.892	0.809	-0.103	**<0.001**
GS-16	65	11	4	0.938	0.741	-0.266	**<0.001**
GLSA-22	64	7	0	0.984	0.773	-0.273	**<0.001**
GLSA-52	52	6	0	0.712	0.753	0.055	0.118
GLSA-65	64	9	2	0.516	0.577	0.106	**<0.001**
Mean	63.333	9.333	-	0.678	0.651	-0.015	-
SE	1.424	0.764	-	0.094	0.067	0.066	-

Pairwise F_{ST} calculations roughly reflect an isolation by distance model even when *G. o. oregonensis* is excluded due to low sample size (Table 5). For instance, *G. o. californicus* is most closely related to *G. o. lascivus* (F_{ST} = 0.053) and there is more differentiation between *G. o. californicus* and *G. o. flaviventris* (F_{ST} = 0.165). On the other hand, the STRUCTURE analysis placed the majority of *G. o. californicus* and *G. o. lascivus* individuals in different genetic clusters (Fig. 5). The mixture of white within the

green genetic group could signify recent gene flow or mixed ancestry. The most likely number of genetic clusters for the data in the admixture model was K=2 and K=3 in the no admixture model (Appendix F). Thus, the STRUCTURE analysis supports at most three genetic groups.

BOTTLENECK did not detect population bottleneck signatures in *G. o. californicus, G. o. lascivus,* or *G. o. flaviventris* (Table 6). Interestingly, bottlenecks were detected in *G. o. fuliginosus* and *G. o. stephensi* under the IAM (P = 0.003 and P = 0.001 respectively) and TPM (P = 0.01 and P = 0.001 respectively) models of microsatellite evolution.

Table 5. Pairwise population F_{st} values calculated in GenAlEx (Peakall and Smouse 2006, 2012) for 51 *Glaucomys oregonensis californicus,* 65 *G. o. lascivus,* 13 *G. o. fuliginosus,* 11 *G. o. stephensi,* six *G. o. flaviventris* samples, and one *G. o. oregonensis* sample.

	G. o. californicus	*G. o. flaviventris*	*G. o. fuliginosus*	*G. o. lascivus*	*G. o. stephensi*
G. o. californicus	0.000				
G. o. flaviventris	0.165	0.000			
G. o. fuliginosus	0.157	0.068	0.000		
G. o. lascivus	0.053	0.093	0.093	0.000	
G. o. stephensi	0.127	0.034	0.046	0.064	0.000

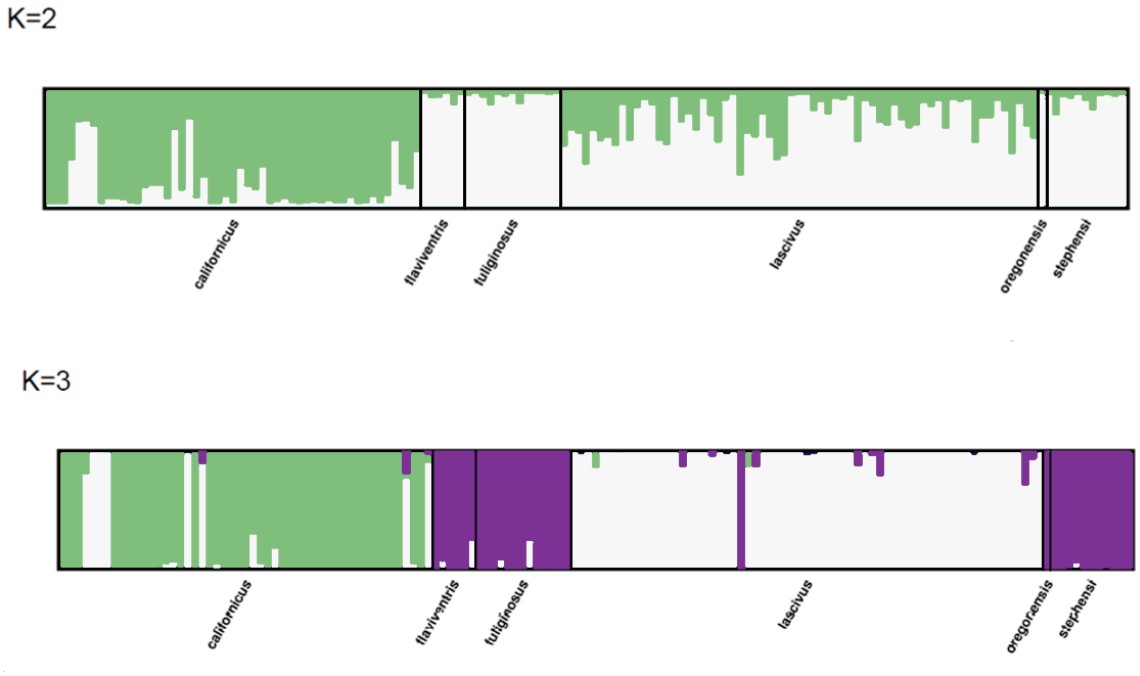


Figure 5. Q-plots from STRUCTURE generated in CLUMPAK (Kopelman et al. 2015) illustrating which genetic cluster individuals were assigned to in the admixture (top) and no admixture models (bottom). Individuals are grouped by subspecies on the x-axis. The most likely number of genetic clusters presented in the admixture (K=2) and no admixture (K=3) models were computed in STRUCTURE HARVESTER (Evanno et al. 2005, Earl and vonHoldt 2012).

Table 6. One-tailed Wilcoxon signed rank tests calculated in BOTTLENECK v 1.2.02 (Piry et al. 1999) for each subspecies population and under each microsatellite mutation model. Significant P-values are indicated in bold with asterisks (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

Population	**IAM**	**TPM**	**SMM**
G. o. californicus	0.674	0.898	1.000
G. o. flaviventris	0.248	0.248	0.590
G. o. fuliginosus	**0.003****	**0.010****	0.500
G. o. lascivus	0.500	0.715	0.997
G. o. stephensi	**0.001*****	**0.001*****	0.082

DISCUSSION

Genetic structure

My study had three main objectives: 1) analyze the intraspecific genetic variation in *G. oregonensis* with a focus on the subspecific differences in California, 2) evaluate *G. o. californicus* genetic diversity, and 3) estimate if gene flow is occurring between the rest of *G. oregonensis* and *G. o. californicus*. In regard to the first objective, our microsatellite data supports previously published work describing a chiefly south to north migration of *G. oregonensis* (Arbogast 2007; Smith 2012; Arbogast et al. 2017). As populations colonized new land post-Last Glacial Maximum, founder effects would have induced a genetic diversity cline, with lower diversity observed in regions of newer colonization and more private alleles in the original populations (Hewitt 1996, 2004). Despite limited sampling of the more northern subspecies, we did find the same general cline in genetic diversity as found in *G. sabrinus* and *G. volans* (Arbogast et al. 2005; Petersen and Stewart 2006; Garroway et al. 2011). *G. o. californicus* and *G. o. lascivus* have a total of 28 private microsatellite alleles, more than their northern counterparts, and are sorted into two genetically different groups in STRUCTURE. The subsequent observation of increased mitochondrial haplotypes, especially exclusive haplotypes, in northern California (Fig. 4) suggests a pattern of multiple, independent refugia during glacial cycles (Brunsfeld et al. 2001; Shafer et al. 2010; Hope et al. 2016). Genetic variation would have increased in the area when slightly diverged populations interbred

after the ice receded (Petit et al. 2003; Shafer et al. 2010), thereby confounding the signatures of glaciation and potentially leading to outbreeding depression if local adaptations were diluted by admixture. Accordingly, the most common mitochondrial haplotype (depicted with the medium shade of blue in Fig. 4) appears to be an ancestral haplotype that may have originated in the south and moved northward as individuals expanded their range after the Last Glacial Maximum.

Considering *G. o. fuliginosus'* small sample size (n = 12), they are surprisingly diverse (H = 0.625), which implies there is more diversity to be discovered with increased sampling of this subspecies. Or alternatively, that the current subspecies delineations should be revised. This statement is introduced cautiously though, as more samples would be necessary to resolve subspecific designations. Our STRUCTURE analysis clearly supports up to three genetically recognizable populations: a *G. o. lascivus* population, a *G. o. californicus* population, and a population that is a mixture of the other sampled subspecies. Excluding *G. o. oregonensis* (n = 1), perhaps *G. o. stephensi*, *G. o. flaviventris*, and *G. o. fuliginosus* should be considered one subspecies with distinct population segments (DPS) pending morphological evaluation. Or they could be split into merely two subspecies. *G. o. fuliginosus* would encompass Oregon and Washington and *G. o. stephensi* and *G. o. flaviventris* would represent one subspecies.

Nevertheless, it is apparent from our study that *G. o. californicus* is important to overall species diversity as stated in Arbogast et al. (2017). There are many exclusive alleles and haplotypes in this population, in addition to representing a single genetic cluster. It is assumed to be a fragmented and isolated population, but we did not detect a

prominent decrease in genetic diversity compared to the larger, more continuous *G. o. lascivus* population. A study conducted by Smith and Person (2007) showed that the minimum habitat size required to sustain isolated *G. sabrinus* populations in Alaska for 100 years was 78,935 ha; and the San Bernardino Mountains are about 202,638 ha when calculated in Google Earth Pro v 7.3.2 by drawing a polygon over entirely forested areas. So, assuming the area is completely forested and contains other preferred habitat parameters, that location may be sufficient to support the current *G. o. californicus* population.

We also did not detect a bottleneck in the San Bernardino population, though this may have been due to a variety of different factors. BOTTLENECK is only able to detect recent events so the ability to identify a bottleneck is strongly influenced by a species' generation time and the time passed since the event (Piry et al. 1999; Cristescu et al. 2010). As such, it is more difficult to identify deep bottleneck events in taxa with short generation times and when the proposed bottleneck may have occurred thousands of years ago. Cristescu et al. (2010) also noticed that bottlenecks were not detected when the number of loci used in the program was less than 11, and they further hypothesized that loci with imperfect repeats may encourage better bottleneck detection because they are less prone to DNA slipped strand mispairing (slippage). We used nine loci in our study, but all of them were dinucleotide microsatellites as the imperfect repeats could not be genotyped. For these reasons, there is a possibility that we were not able to detect a bottleneck because it occurred too many generations ago and/or our genetic markers did not provide enough resolution.

The third objective of the study was to estimate if gene flow was transpiring between the rest of *G. oregonensis* and *G. o. californicus*, and even though we were not able to answer this question in our study, we are able to offer some insight. The San Bernardino population does share some of its mitochondrial haplotypes with *G. o. lascivus*, but from nuclear DNA markers one would infer that they are two, very separate populations. In STRUCTURE there were a few individuals who were assigned to the white population (as depicted in Fig. 5) even though most of the San Bernardino subspecies was assigned to the green population. This gives the impression of gene flow on the surface, but six out of the seven individuals who were assigned to the white population (Fig. 5) were museum specimens collected from the San Jacinto Mountains in the early 1900s. The seventh individual was a museum sample collected in 1905 from the San Bernardino Mountains. We speculate that either *G. o. lascivus* retained San Jacinto, ancestral genotypes and *G. o. californicus* diverged, or that they were isolated in the San Bernardino Mountains at some point during glacial cycles.

Finally, the GenBank sequence AF030390 was of a *G. oregonensis* not a *G. sabrinus* individual. This sample had the same cyt-*b* haplotype as other *G. oregonensis* samples from Washington and the species split with *G. sabrinus* is well-supported (Appendix C). Seeing as *G. oregonensis* was described fairly recently in 2017, there could be more samples housed in GenBank that are not *G. sabrinus* but *G. oregonensis*.

Genotyping by sequencing (GBS)

Microsatellites have been the traditional DNA marker of choice for many population genetic studies but suffer from several issues like fragment size homology, ascertainment bias, and low throughput as a result of capillary electrophoresis genotyping (Darby et al. 2016; De Barba et al. 2017; Zhan et al. 2017). Some of these issues can be alleviated, however, if the process of genotyping microsatellites is converted to fit exact sequence data produced by high-throughput sequencing (HTS) platforms like Illumina (Vartia et al. 2016). Raw sequence data was generated in this study, but the analysis conducted used total fragment length as input, similar to traditional capillary sequencing. The data incorporating precise sequence content from this study is still being analyzed (Yuan et al. *in prep*), but we discuss some of the implications of GBS below.

Access to raw sequences from HTS allows precise sizing of microsatellite fragments (Darby et al. 2016; Barbian et al. 2018; Šarhanová et al. 2018; Tibihika et al. 2019). In other words, researchers are able to discern the exact length of a fragment down to the last base pair, which was difficult to do in capillary sequencing as machine calibrations often affected the genotyping process (Darby et al. 2016). Access to raw sequences also allows easy detection of variation either within the repeat motif or the flanking regions (Barbian et al. 2018; Šarhanová et al. 2018). Both of these factors contribute to what is arguably the greatest advantage of GBS: the reduction in microsatellite fragment size (length) homoplasy (Vartia et al. 2016; Barbian et al. 2018; Šarhanová et al. 2018). Size homoplasy arises when fragments of the same length are

considered the same alleles but there are hidden, underlying nucleotide differences between them (Darby et al. 2016). For example, in our dataset both LACM-920 and LACM-871 have a genotype of 84/84 at the GS-2 locus (Appendix D). But, while LACM-871 is truly homozygous, LACM-920 is not (Fig. 6). This allele variation would have been missed if we had genotyped the samples using capillary electrophoresis, but we were able to detect this variation using the CHIIMP pipeline (Barbian et al. 2018).

LACM 920 GS2-1 84 bp

ATCTGCTCTGGGGTGTGTGTGTGTGTGTGTTTCTTTTTAGTCTA

LACM 920 GS2-2 84 bp

GTCTGCTCTGGGGTGTGTGTGTGTGTGTGTTTCTTTTTAGTCTA

Figure 6. GS-2 genotype for LACM-920. Both alleles are shown with the locus primers cut from the sequence. Using traditional capillary electrophoresis methods, the underlying nucleotide difference (highlighted in yellow) would have been missed.

Along with this increased detection of previously masked alleles, researchers will inevitably notice a trend toward higher allelic diversity, more loci deviating from HWE, and increased heterozygosity in future studies (Tibihika et al. 2019). Some studies have shown that this helps clarify population genetic structure (Darby et al. 2016; Barbian et al. 2018; Šarhanová et al. 2018; Tibihika et al. 2019), and we are exploring that result with the samples from this study (Yuan et al. *in prep*). This allelic diversity increase is supposed to represent a more accurate dataset since it increases detail and resolution in

analyses without changing global genetic diversity statistics entirely (Barbian et al. 2018; Šarhanová et al. 2018). More research should be conducted on this subject, but if that holds true, GBS would become even more valuable for conservation and wildlife studies that use degraded DNA samples such as museum specimens and fecal samples.

Microsatellites are still the preferred marker to test in conservation and wildlife studies because their codominant mode of inheritance, neutrality, and highly polymorphic character provide high resolution and statistical power to studies of genetic structure and kinship (forensics) (Lampa et al. 2013; Putnam and Carbone 2014; De Barba et al. 2017; Zhan et al. 2017; Pimentel et al. 2018). In recent years single nucleotide polymorphisms (SNPs) have increased in popularity, but SNP studies usually require many loci (Zhan et al. 2017) and restriction enzyme digesting (Darby et al. 2016), which makes the transition to testing SNPs on degraded samples difficult as the DNA template is already fragmented and of low-quality (Lampa et al. 2013; Hawkins et al. 2016a). Moreover, the large amount of data already available from traditionally analyzed microsatellites facilitates future dataset comparisons, with the exception of homoplasy presence discussed above, if future studies shift toward GBS pipelines (Lampa et al. 2013; Putnam and Carbone 2014; De Barba et al. 2017).

Alongside the potential benefits of GBS, there are a few cons to consider. For one, it is still not possible to detect homoplasy that arises from convergence on the same length and sequence (Šarhanová et al. 2018), and it will likely become more difficult to compare across different types of datasets (GBS produced data versus traditional capillary sequencing). For instance, Barbian et al. (2018) discovered consistent allele

length differences (1 - 3 bp) between their capillary and GBS dataset. On top of this, the fundamental way alleles are named may need to change. Questions that should be discussed by researchers include: should total fragment length continue to be used as the allele definition, or should a shift be made toward naming alleles based on the true number of repeat motifs, or should every unique sequence be declared an allele (Darby et al. 2016; Šarhanová et al. 2018; Tibihika et al. 2019)?

In addition, older programs that analyzed capillary sequencing data will now need to update their software and/or algorithms (Darby et al. 2016). Examples include how our raw CHIIMP output required extra processing to be run in programs such as STRUCTURE, and how the data should not be run in BOTTLENECK because it is analyzing allele mutation models. In line with this, once more research has been conducted on allelic dropout it will also be important to standardize the genotyping pipeline to a certain extent and increase efficiency. For example, De Barba et al. (2017), Barbian et al. (2018), Pimentel et al. (2018), and Tibihika et al. (2019), all have slightly different parameters for calling alleles and different pipelines, so what is the most accurate way to genotype the samples?

Ecological and conservation implications

Glaucomys oregonensis is a crucial component of its ecosystem as its association with important mycorrhizal fungi (truffles) and the endangered *Strix occidentalis caurina* (Northern spotted owl) is well documented (Carey 2002; Weigl 2007; Holloway et al.

2012; Smith 2012). The species is also a prey item for three other populations of concern; *Pekania pennanti* (fisher, CNDDB 2019; FWS 2016b), *S. o. occidentalis* (California spotted owl), which is a California subspecies of special concern that was rejected last year to be listed as endangered (CNDDB 2019; FWS 2019), and *Martes caurina humboldtensis* (Humboldt marten), which is listed as endangered under the California Endangered Species Act (FWS 2018; CNDDB 2019). Ergo, the continued wellbeing of *G. oregonensis* is essential to the health of its ecological community, and since subpopulations with unique genetic diversity like *G. o. californicus* are important contributors to the species gene pool, their continued survival is valuable as well.

Evaluation of the San Bernardino population suggested the intensity of the threats they were exposed to, such as urban air pollution, habitat fragmentation, and wildfire treatment, was low, and the request for protection was denied (FWS 2016a). This was however, before a genetic profile was built, and stochastic events in the future may still endanger this population. Ecological studies have shown how *G. oregonensis* can be behaviorally plastic (Rosenberg and Anthony 1992; Smith 2007; Weigl 2007), possibly because they require only certain habitat necessities to persist, like nest sites and food (Waters and Zabel 1995; Carey 2002). Older forests usually have these essential habitat parameters, for example a higher abundance and variety of fungi and lichen, two staples in their diet, and can thus better sustain flying squirrel populations (Waters and Zabel 1995; Pyare and Longland 2001; Meyer et al. 2005; Smith 2007, 2012). Therefore, the *G. o. californicus* population may be a self-perpetuating population at the moment, but to ensure the long term viability of this population, their habitat needs to remain largely

intact as increased stand turnover and habitat loss could lead to population declines (Holloway et al. 2012).

When *G. o. californicus* was proposed for listing on the Endangered Species Act, the U.S. Fish and Wildlife Service began evaluating the population in and around the San Bernardino Mountains. During that time, they conducted surveys and expanded to citizen science measures to determine presence/absence data for the population. Around 76% of the San Bernardino Mountains and 65% of the San Jacinto Mountains are managed by the U.S. Forest Service, yet there have been no confirmed sightings of flying squirrels in the San Jacinto Mountains since the 1990s (FWS 2016a). Due to the unique mitochondrial haplotype recovered in the San Jacinto population from a sample collected in 1916 (UCLA-7487) and the genetic relationship with *G. o. lascivus*, it is unnerving to consider if this population has been extirpated. While it is optimistic that the overall *G. o. californicus* genetic diversity is not in a current bottleneck, there are no other local populations available should genetic rescue or intervention need to be implemented in the future, so we recommend persistent vigilance and long-term monitoring of the population as small mammals are prone to boom-bust population cycles (Weldy et al. 2019). If flying squirrels are intended to be reintroduced to the San Jacinto habitat, the genetic data recovered here suggest that individuals with the *G. o. lascivus* nuclear profile, plus the most common (blue) mitochondrial haplotype, best represent that historical population although more data may be beneficial to confirm these findings.

CONCLUSIONS

Overall, this study represents the most in-depth analysis of *G. oregonensis* to date, especially with regard to the subspecies in California. One hundred and forty-seven new mitochondrial and microsatellite profiles have been generated, with 51 of those taken from the San Bernardino flying squirrel. I recovered a pattern of mitochondrial diversity unlike similar studies in the eastern United States, likely due to the Pacific Coast retaining refugial populations of flying squirrels during the most recent retreat of the Laurentide and Cordilleran ice sheets (Arbogast 1999; Arbogast 2007). This has resulted in higher haplotypic diversity in northern California and fewer in the south, opposite of traditional northward colonization following glacial retreat (Kerhoulas and Arbogast 2010; Bozarth et al. 2011). These results cannot be completely untangled from signatures of population bottlenecks or founder effects, however, baseline genetic data have now been established for this species and will likely be useful for comparison in future studies.

And although the majority of my samples were museum specimen samples (Appendix G), which contain degraded endogenous DNA and possible exogenous DNA contamination (Hawkins et al. 2016a), a study using exclusively fresh tissue would not add much more to the data resolution. Previous studies have shown that loci amplification success varies more in museum samples and there may be an increase in the amount of nuclear copies of mitochondrial genes, but with a combination of stringent bioinformatics and lab protocols, data generated from degraded samples have proven to be quite reliable

(den Tex et al. 2010; Hawkins et al. 2016a; De Barba et al. 2017; Barbian et al. 2018; McDonough et al. 2018). Analyses investigating the specific amounts of allelic dropout in these samples is also being conducted to validate recovered genotypes (Yuan et al. *in prep*).

In addition, as discussed previously, *G. o. californicus* appears to have suitable habitat for long-term viability even while isolated, so long as no additional stochastic events disrupt the current ecosystem and global climate change does not alter the current forest composition. The genetic diversity of this subspecies is still low, even with a robust sample size (including 44 out of the 50 known U.S. museum specimens), and it is possible that ancient bottlenecks did in fact occur in this population, despite our BOTTLENECK analysis not identifying a recent change. This is probably due to the limited ability of BOTTLENECK to reveal events beyond about 15 generations (maximally ~45 years in flying squirrels; Villa et al. 1999). Additionally, the now seemingly extirpated San Jacinto population appears to be quite genetically distinct from the San Bernardino population, a mere 40 kilometers away, which may indicate a significant amount of isolation between the two forms. Further surveys for the San Jacinto subpopulation should be performed, and if possible, sampled for a genetic profile to evaluate if our historical signatures exist in any remnant squirrels.

REFERENCES

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data [Internet]. Cambridge (UK): Babraham Institute; [cited 2020 Mar 1]. Available from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arbogast BS. 1999. Mitochondrial DNA phylogeography of the New World flying squirrels (*Glaucomys*): implications for Pleistocene biogeography. J Mammal. 80(1):142-155.

Arbogast BS, Browne RA, Weigl PD. 2001. Evolutionary genetics and Pleistocene biogeography of North American tree squirrels (*Tamiasciurus*). J Mammal. 82(2):302-319.

Arbogast BS, Browne RA, Weigl PD, Kenagy GJ. 2005. Conservation genetics of endangered flying squirrels (*Glaucomys*) from the Appalachian Mountains of eastern North America. Anim Conserv. 8(2):123–133.

Arbogast BS. 2007. A brief history of the New World flying squirrels: phylogeny, biogeography, and conservation genetics. J Mammal. 88(4):840–849.

Arbogast BS, Schumacher KI, Kerhoulas NJ, Bidlack AL, Cook JA, Kenagy GJ. 2017. Genetic data reveal a cryptic species of New World flying squirrel: *Glaucomys oregonensis*. J Mammal. 98(4):1027-1041.

Balakirev AE, Abramov AV, Rozhnov VV. 2017. The phylogeography of red spiny rats *Maxomys surifer* (Rodentia, Muridae) in Indochina with comments on taxonomy and description of new subspecies. Zool Stud. 56(6):2017-2056.

Bandelt H, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 16(1):37–48.

Barbian HJ, Connell AJ, Avitto AN, Russell RM, Smith AG, Gundlapally MS, Shazad AL, Li Y, Bibollet-Ruche F, Wroblewski EE, Mjungu D. 2018. CHIIMP: An automated high-throughput microsatellite genotyping platform reveals greater allelic diversity in wild chimpanzees. Ecol Evol. 8(16):7946-7963.

Bidlack AL, Cook JA. 2001. Reduced genetic variation in insular northern flying squirrels (*Glaucomys sabrinus*) along the North Pacific Coast. Anim Conserv. 4(4):283–290.

Bidlack AL, Cook JA. 2002. A nuclear perspective on endemism in northern flying squirrels (*Glaucomys sabrinus*) of the Alexander Archipelago, Alaska. Conserv Genet. 3(3):247–259.

Bozarth CA, Lance SL, Civitello DJ, Glenn JL, Maldonado JE. 2011. Phylogeography of the gray fox (*Urocyon cinereoargenteus*) in the eastern United States. J Mammal. 92(2):283-294.

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS. 2001. Comparative phylogeography of northwestern North America: a synthesis. In: Silvertown J, Antonovics J, editors. Integrating ecology and evolution in a spatial context. Oxford (UK): Blackwell Science. p. 319–340.

Butynski TM, de Jong YA. 2017. The Mount Kenya potto is a subspecies of the eastern potto *Perodicticus ibeanus*. Primate Conserv. 31(1):49-52.

Carey AB. 2002. Ecology of northern flying squirrels: implications for ecosystem management in the Pacific Northwest, USA. In: Goldingay RL, Scheibe JS, editors. Proceedings of the international theriological congress. Furth, Germany: Filander Verlag. p. 45-66.

CNDDB (California Natural Diversity Database), CDFW. 2019. Special Animals List [Internet]. Sacramento (CA): California Department of Fish and Wildlife; [cited 2020 Feb 19]. Available from https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=109406&inline

Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics. 144(4):2001-2014.

Cristescu R, Sherwin WB, Handasyde K, Cahill V, Cooper DW. 2010. Detecting bottlenecks using BOTTLENECK 1.2.02 in wild populations: the importance of the microsatellite structure. Conserv Genet. 11(3):1043-1049.

Currie R, Cameron S. 2013. Carolina northern flying squirrel (*Glaucomys sabrinus coloratus*) 5-year review: summary and evaluation [Internet]. Asheville (NC): U.S. Fish and Wildlife Service; [cited 2020 Feb 19]. Available from https://ecos.fws.gov/docs/five_year_review/doc4177.pdf

Darby BJ, Erickson SF, Hervey SD, Ellis-Felege SN. 2016. Digital fragment analysis of short tandem repeats by high-throughput amplicon sequencing. Ecol Evol. 6(13):4502-4512.

De Barba M, Miquel C, Lobréaux S, Quenette PY, Swenson JE, Taberlet P. 2017. High-throughput microsatellite genotyping in ecology: improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. Mol Ecol Resour. 17(3):492-507.

Den Tex RJ, Maldonado JE, Thorington R, Leonard JA. 2010. Nuclear copies of mitochondrial genes: another problem for ancient DNA. Genetica. 138:979-984.

Dolan PG, Carter DC. 1977. *Glaucomys volans*. Mamm Species. (78):1-6.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 4(2):359-361.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 14(8):2611-2620.

FWS (U.S. Fish and Wildlife Service). 2016a. 12-month findings on petitions to list island marble butterfly, San Bernardino flying squirrel, spotless crake, and Sprague's pipit as endangered or threatened species; notice of 12-month petition findings [Internet]. Carlsbad (CA): U.S. Fish and Wildlife Service; [cited 2020 Feb 19]. Available from https://www.govinfo.gov/content/pkg/FR-2016-04-05/pdf/2016-07809.pdf#page=1

FWS (U.S. Fish and Wildlife Service). 2016b. Final species report: Fisher (*Pekania pennanti*), west coast population [Internet]. Federal Register; [cited 2020 Apr 25]. Available from https://www.regulations.gov/docket?D=FWS-R8-ES-2018-0105

FWS (U.S. Fish and Wildlife Service). 2018. Endangered and threatened wildlife and plants; threatened species status for coastal distinct population segment of the Pacific marten [Internet]. Federal Register; [cited 2020 Mar 28]. Available from https://www.federalregister.gov/documents/2018/10/09/2018-21794/endangered-and-threatened-wildlife-and-plants-threatened-species-status-for-coastal-distinct

FWS (U.S. Fish and Wildlife Service). 2019. Endangered and threatened wildlife and plants; 12-month finding for the California spotted owl [Internet]. Federal Register; [cited 2020 Mar 28]. Available from https://www.federalregister.gov/documents/2019/11/08/2019-24336/endangered-and-threatened-wildlife-and-plants-12-month-finding-for-the-california-spotted-owl

Gamage S, Groves CP, Manikar FM, Turner CS, Padmalal KU, Kotagama SW. 2017. The taxonomy, distribution, and conservation status of the slender loris (Primates, Lorisidae: *Loris*) in Sri Lanka. Primate Conserv. 31(1):83-106.

Garroway CJ, Bowman J, Holloway GL, Malcolm JR, Wilson PJ. 2011. The genetic signature of rapid range expansion by flying squirrels in response to contemporary climate warming. Glob Change Biol. 17(5):1760–1769.

Glenn TC, Nilsen RA, Kieran TJ, Sanders JG, Bayona-Vásquez NJ, Finger JW, Pierson TW, Bentley KE, Hoffberg SL, Louha S, Garcia-De Leon FJ. 2019. Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). PeerJ. 7:e7755.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59(3):307-321.

Hale ML, Burg TM, Steeves TE. 2012. Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. PLoS ONE [Internet]. [cited 2020 Feb 19];7(9):e45170. Available from https://doi.org/10.1371/journal.pone.0045170

Hawkins MT, Hofman CA, Callicrate T, McDonough MM, Tsuchiya MT, Gutiérrez EE, Helgen KM, Maldonado JE. 2016a. In-solution hybridization for mammalian mitogenome enrichment: Pros, cons and challenges associated with multiplexing degraded DNA. Mol Ecol Resour. 16(5):1173-1188.

Hawkins MT, Leonard JA, Helgen KM, McDonough MM, Rockwood LL, Maldonado JE. 2016b. Evolutionary history of endemic Sulawesi squirrels constructed from UCEs and mitogenomes sequenced from museum specimens. BMC Evol Biol. 16(1):80.

Hewitt GM. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc. 58(3):247-276.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philos T R Soc B. 359(1442):183-195.

Holloway GL, Smith WP, Halpern CB, Gitzen RA, Maguire CC, West SD. 2012. Influence of forest structure and experimental green-tree retention on northern flying squirrel (*Glaucomys sabrinus*) abundance. For Ecol Manag. 285:187-194.

Hope AG, Malaney JL, Bell KC, Salazar-Miralles F, Chavez AS, Barber BR, Cook JA. 2016. Revision of widespread red squirrels (genus: *Tamiasciurus*) highlights the

complexity of speciation within North American forests. Mol Phylogenet Evol. 100:170-182.

Howell AH. 1918. Revision of the American flying squirrels. N Am Fauna. 13:1-64.

Huang QY, Xu FH, Shen H, Deng HY, Liu YJ, Liu YZ, Li JL, Recker RR, Deng HW. 2002. Mutation patterns at dinucleotide microsatellite loci in humans. Am J Hum Genet. 70(3):625-634.

Irwin DM, Kocher TD, Wilson AC. 1991. Evolution of the cytochromeb gene of mammals. J Mol. 32(2):128-144.

Jayat JP, D'Elía G, Torres R, Pacheco SE, Ortiz PE, Salazar-Bravo J, Patterson BD. 2017. Integration of morphological, ecological, and genetic evidence suggests that the genus *Andinomys* (Rodentia, Cricetidae) is monospecific. J Mammal. 98(4):1060-1077.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12):1647-1649.

Keller LF, Waller DM. 2002. Inbreeding effects in wild populations. Trends Ecol Evol. 17(5):230–241.

Kerhoulas NJ, Arbogast BS. 2010. Molecular systematics and Pleistocene biogeography of Mesoamerican flying squirrels. J Mammal. 91(3):654-667.

Kiesow AM, Wallace LE, Britten HB. 2011. Characterization and isolation of five microsatellite loci in northern flying squirrels, *Glaucomys sabrinus* (Sciuridae, Rodentia). West N Am Nat. 71(4):553–556.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour. 15(5):1179-1191.

Lampa S, Henle K, Klenke R, Hoehn M, Gruber B. 2013. How to overcome genotyping errors in non-invasive genetic mark-recapture population size estimation—A review of available methods illustrated by a case study. J Wildlife Manage. 77(8):1490-1511.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol. 34(3):772-773.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9(4):357.

Leigh JW, Bryant D. 2015. PopART: Full-feature software for haplotype network construction. Methods Ecol Evol. 6(9):1110-1116.

Luikart G, Cornuet JM. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conserv Biol. 12(1):228-237.

Marsden CD, Ortega-Del Vecchyo D, O'Brien DP, Taylor JF, Ramirez O, Vilà C, Marques-Bonet T, Schnabel RD, Wayne RK, Lohmueller KE. 2016. Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. Proc Natl Acad Sci USA. 113(1):152–157.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal [Internet]. [cited 2020 Mar 5];17(1):10-12. Available from https://doi.org/10.14806/ej.17.1.200

McDonough MM, Parker LD, Rotzel McInerney N, Campana MG, Maldonado JE. 2018. Performance of commonly requested destructive museum samples for mammalian genomic studies. J Mammal. 99(4):789-802.

Meyer MD, North MP, Kelt DA. 2005. Fungi in the diets of northern flying squirrels and lodgepole chipmunks in the Sierra Nevada. Can J Zool. 83(12):1581-1589.

Oshida T, Lin LK, Masuda R, Yoshida MC. 2000. Phylogenetic relationships among Asian species of *Petaurista* (Rodentia, Sciuridae), inferred from mitochondrial cytochrome b gene sequences. Zool Sci. 17(1):123-128.

Owen LA, Finkel RC, Minnich RA, Perez AE. 2003. Extreme southwestern margin of late Quaternary glaciation in North America: Timing and controls. Geology. 31(8):729–732.

Peakall RO, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes. 6(1):288-295.

Peakall RO, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 28(19):2537–2539.

Petersen SD, Stewart DT. 2006. Phylogeography and conservation genetics of southern flying squirrels (*Glaucomys volans*) from Nova Scotia. J Mammal. 87(1):153–160.

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. Science. 300(5625):1563-1565.

Pimentel JS, Carmo AO, Rosse IC, Martins AP, Ludwig S, Facchin S, Pereira AH, Brandão-Dias PF, Abreu NL, Kalapothakis E. 2018. High-throughput sequencing strategy for microsatellite genotyping using neotropical fish as a model. Front Genet [Internet]. [cited 2020 Mar 28];9:73. Available from https://doi.org/10.3389/fgene.2018.00073

Piry S, Luikart G, Cornuet JM. 1999. BOTTLENECK: a program for detecting recent effective population size reductions from allele data frequencies. J Hered. 90(4):502-503.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics. 155(2):945-959.

Puckett EE, Etter PD, Johnson EA, Eggert LS. 2015. Phylogeographic analyses of American black bears (*Ursus americanus*) suggest four glacial refugia and complex patterns of postglacial admixture. Mol Biol Evol. 32(9):2338-2350.

Putman AI, Carbone I. 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecol Evol. 4(22):4399-4428.

Pyare S, Longland WS. 2001. Patterns of ectomycorrhizal-fungi consumption by small mammals in remnant old-growth forests of the Sierra Nevada. J Mammal. 82(3):681-689.

Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22(5):939-946.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 61(3):539-542.

Rosenberg DK, Anthony RG. 1992. Characteristics of northern flying squirrel populations in young second-and old-growth forests in western Oregon. Can J Zool. 70(1):161-166.

Šarhanová P, Pfanzelt S, Brandt R, Himmelbach A, Blattner FR. 2018. SSR-seq: Genotyping of microsatellites using next-generation sequencing reveals higher level of polymorphism as compared to traditional fragment size scoring. Ecol Evol. 8(22):10817-10833.

Sawyer YE, Cook JA. 2016. Phylogeographic structure in long-tailed voles (Rodentia: Arvicolinae) belies the complex Pleistocene history of isolation, divergence, and recolonization of northwest North America's fauna. Ecol Evol. 6(18):6633-6647.

Shafer ABA, Cullingham CI, Côté SD, Coltman DW. 2010. Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. Mol Ecol. 19(21):4589–4621.

Smith WP. 2007. Ecology of *Glaucomys sabrinus*: habitat, demography, and community relations. J Mammal. 88(4):862–881.

Smith WP. 2012. Sentinels of ecological processes: the case of the northern flying squirrel. BioScience. 62(11):950–961.

Smith WP, Person DK. 2007. Estimated persistence of northern flying squirrel populations in temperate rain forest fragments of Southeast Alaska. Biol Conserv. 137(4):626–636.

Spielman D, Brook BW, Frankham R. 2004. Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci USA. 101(42):15261–15264.

Thatte P, Joshi A, Vaidyanathan S, Landguth E, Ramakrishnan U. 2018. Maintaining tiger connectivity and minimizing extinction into the next century: Insights from landscape genetics and spatially-explicit simulations. Biol Conserv. 218:181–191.

Thorington RWJ, Hoffmann RS. 2005. Family Sciuridae. In: Wilson DE, Reeder DM, editors. Mammal species of the world: a taxonomic and geographic reference. Baltimore (MD): The John Hopkins University Press. p. 754–818.

Tibihika PD, Curto M, Dornstauder-Schrammel E, Winter S, Alemayehu E, Waidbacher H, Meimberg H. 2019. Application of microsatellite genotyping by sequencing (SSR-GBS) to measure genetic diversity of the East African *Oreochromis niloticus*. Conserv Genet. 20(2):357-372.

Ujvari B, Klaassen M, Raven N, Russell T, Vittecoq M, Hamede R, Thomas F, Madsen T. 2018. Genetic diversity, inbreeding and cancer. Proc Royal Soc B. 285(1875):20172589.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities and interfaces. Nucleic Acids Res [Internet]. [cited 2020 Feb 12];40(15):e115. Available from https://doi.org/10.1093/nar/gks596

Vartia S, Villanueva-Cañas JL, Finarelli J, Farrell ED, Collins PC, Hughes GM, Carlsson JE, Gauthier DT, McGinnity P, Cross TF, FitzGerald RD. 2016. A novel method of microsatellite genotyping-by-sequencing using individual combinatorial barcoding. R Soc Open Sci. 3(1):150565.

Villa LJ, Carey AB, Wilson TM, and Glos KB. 1999. Maturation and reproduction of northern flying squirrels in Pacific Northwest forests [Internet]. Portland (OR): USDA For. Serv. Gen. Tech. Rep. PNW-GTR-444; [cited 2020 Mar 28]. Available from https://doi.org/10.2737/PNW-GTR-444

Waters JR, Zabel CJ. 1995. Northern flying squirrel densities in fir forests of northeastern California. J Wildlife Manage. 59(4):858-866.

Weigl PD. 2007. The northern flying squirrel (*Glaucomys sabrinus*): a conservation challenge. J Mammal. 88(4):897–907.

Weldy MJ, Epps CW, Lesmeister DB, Manning T, Linnell MA, Forsman ED. 2019. Abundance and ecological associations of small mammals. J Wildlife Manage. 83(4):902-915.

Wells-Gosling N, Heaney LR. 1984. *Glaucomys sabrinus*. Mamm Species. (229):1–8.

Wolf S. 2010. Petition to list the San Bernardino flying squirrel (*Glaucomys sabrinus californicus*) as threatened or endangered under the United States Endangered Species Act [Internet]. San Francisco (CA): Center for Biological Diversity; [cited 2020 Feb 19]. Available from https://www.biologicaldiversity.org/species/mammals/San_Bernardino_flying_squirrel/pdfs/San_Bernardino_Flying_Squirrel_Petition.pdf

Yuan SC, Malekos E, Hawkins MT. Assessing the levels of microsatellite dropout in museum specimens using high-throughput sequencing technology. *In preparation for Molecular Ecology Resources.*

Zhan L, Paterson IG, Fraser BA, Watson B, Bradbury IR, Nadukkalam Ravindran P, Reznick D, Beiko RG, Bentzen P. 2017. MEGASAT: automated inference of microsatellite genotypes from sequence data. Mol Ecol Resour. 17(2):247-256.

Zittlau KA, Davis CS, Strobeck C. 2000. Characterization of microsatellite loci in northern flying squirrels (*Glaucomys sabrinus*). Mol Ecol. 9(6):826-827.

Appendix A

List of *Glaucomys oregonensis* samples analyzed in this study (n = 147), consisting of 51 *G. o. californicus*, 65 *G. o. lascivus,* 13 *G. o. fuliginosus,* 11 *G. o. stephensi,* six *G. o. flaviventris* samples, and one *G. o. oregonensis* sample. The museum and zoo abbreviations are as follows: BBAZ = Big Bear Alpine Zoo, Big Bear Lake, California; HSU = Humboldt State University Vertebrate Museum, Arcata, California; KU = University of Kansas Biodiversity Institute and Natural History Museum, Lawrence, Kansas; LACM = Natural History Museum of Los Angeles, Los Angeles, California; MVZ = Museum of Vertebrate Zoology at University of California Berkeley, Berkeley, California; SDNHM = San Diego Natural History Museum, San Diego, California; UCLA = University of California Los Angeles Dickey Bird and Mammal Collection, Los Angeles, California; UMMZ = University of Michigan Museum of Zoology, Ann Arbor, Michigan. Samples were from frozen tissue, degraded tissue, hair, or museum specimens (clips of fur, bone, adherent muscle tissue). Locality and year refer to collection locations and date, respectively.

*HSU sample that has not been accessioned yet.

Subspecies	**Catalog Number**	**Type of Sample**	**State**	**Locality**	**Year**	**Latitude**	**Longitude**	**Datum**
G. o. californicus	BBAZ-M16002	Hair in RNALater	California	San Bernardino Co.	2019	34.22951	-116.8585	WGS84
G. o. californicus	BBAZ-M16003	Hair in RNALater	California	San Bernardino Co.	2019	34.22951	-116.8585	WGS84
G. o. californicus	BBAZ-M16001	Hair in RNALater	California	San Bernardino Co.	2019	34.22951	-116.8585	WGS84

Subspecies	Catalog Number	Type of Sample	State	Locality	Year	Latitude	Longitude	Datum
G. o. californicus	SDNHM-001	Hair in RNALater	California	San Bernardino Co.	2019	34.24183	-117.1823	WGS84
G. o. californicus	SDNHM-002	Degraded Tissue (Whole Tail)	California	San Bernardino Co.	2016	34.24315	-117.0695	WGS84
G. o. californicus	SDNHM-003	Degraded Tissue (Whole Tail)	California	San Bernardino Co.	2016	34.24315	-117.0695	WGS84
G. o. californicus	UMMZ-79755	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79758	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79760	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79761	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79763	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79753	Museum	California	San Bernardino Co.	1920	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79754	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79756	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79757	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79759	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79762	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UMMZ-79764	Museum	California	San Bernardino Co.	1926	34.25387	-116.9227	NAD27
G. o. californicus	UCLA-17237	Museum	California	San Bernardino Co.	1920	34.25	-116.9333	WGS84
G. o. californicus	UCLA-2566	Museum	California	San Bernardino Co.	1927	34.21667	-116.9	WGS84
G. o. californicus	UCLA-2576	Museum	California	San Bernardino Co.	1920	34.21667	-116.9	WGS84
G. o. californicus	UCLA-2583	Museum	California	San Bernardino Co.	1920	34.21667	-116.9	WGS84
G. o. californicus	UCLA-2598	Museum	California	San Bernardino Co.	1920	34.21667	-116.9	WGS84
G. o. californicus	UCLA-2931	Museum	California	San Bernardino Co.	1920	34.21667	-117.25	WGS84
G. o. californicus	UCLA-17238	Museum	California	San Bernardino Co.	1920	34.25	-116.9333	WGS84
G. o. californicus	UCLA-2573	Museum	California	San Bernardino Co.	1920	34.21667	-116.9	WGS84
G. o. californicus	UCLA-2597	Museum	California	San Bernardino Co.	1920	34.21667	-116.9	WGS84
G. o. californicus	UCLA-7487	Museum	California	Riverside Co.	1916	33.72926	-116.7509	WGS84
G. o. californicus	KU-46261	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27
G. o. californicus	KU-46262	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27
G. o. californicus	KU-46263	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27

Subspecies	Catalog Number	Type of Sample	State	Locality	Year	Latitude	Longitude	Datum
G. o. californicus	KU-46264	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27
G. o. californicus	KU-46265	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27
G. o. californicus	KU-46266	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27
G. o. californicus	KU-46267	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27
G. o. californicus	KU-46268	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27
G. o. californicus	KU-46269	Museum	California	San Bernardino Co.	1926	34.22971	-116.9191	NAD27
G. o. californicus	LACM-921	Museum	California	Riverside Co.	1916	33.74617	-116.7145	NAD27
G. o. californicus	LACM-95619	Museum	California	Riverside Co.	1919	33.74617	-116.7145	NAD27
G. o. californicus	LACM-871	Museum	California	Riverside Co.	1915	33.74617	-116.7145	NAD27
G. o. californicus	LACM-8094	Museum	California	San Bernardino Co.	1942	34.25224	-116.8362	NAD27
G. o. californicus	LACM-920	Museum	California	Riverside Co.	1916	33.74617	-116.7145	NAD27
G. o. californicus	LACM-4826	Museum	California	San Bernardino Co.	1935	34.2495	-117.2952	NAD27
G. o. californicus	MVZ-176127	Museum	California	San Bernardino Co.	1973	34.14433	-116.9791	NAD27
G. o. californicus	MVZ-176128	Museum	California	San Bernardino Co.	1973	34.2405	-117.3244	NAD27
G. o. californicus	MVZ-2088	Museum	California	Riverside Co.	1908	33.72238	-116.7605	NAD27
G. o. californicus	MVZ-176126	Museum	California	San Bernardino Co.	1973	34.21944	-116.9697	NAD27
G. o. californicus	MVZ-5210	Museum	California	San Bernardino Co.	1905	34.25387	-116.9227	NAD27
G. o. californicus	MVZ-5211	Museum	California	San Bernardino Co.	1905	34.21944	-116.9697	NAD27
G. o. californicus	MVZ-7007	Museum	California	San Bernardino Co.	1905	34.25387	-116.9227	NAD27
G. o. californicus	HSU-VM 3095*	Frozen Tissue (Whole Organism)	California	San Bernardino Co.	2002	34.22619	-117.0564	WGS84
G. o. flaviventris	MVZ-11002	Museum	California	Modoc Co.	1910	41.46038	-120.243	NAD27
G. o. flaviventris	MVZ-13303	Museum	California	Siskiyou Co.	1911	41.15268	-122.9632	NAD27
G. o. flaviventris	MVZ-13306	Museum	California	Siskiyou Co.	1911	41.15268	-122.9632	NAD27
G. o. flaviventris	MVZ-13309	Museum	California	Siskiyou Co.	1911	41.15268	-122.9632	NAD27
G. o. flaviventris	MVZ-132650	Museum	California	Mendocino Co.	1951	39.61266	-122.9493	NAD27
G. o. flaviventris	MVZ-132653	Museum	California	Mendocino Co.	1951	39.5228	-123.0911	NAD27
G. o. fuliginosus	MVZ-69214	Museum	California	Siskiyou Co.	1935	41.99797	-122.9055	NAD27
G. o. fuliginosus	MVZ-69216	Museum	California	Siskiyou Co.	1935	41.99797	-122.9055	NAD27
G. o. fuliginosus	HSU-4333	Museum	Oregon	Lane Co.	1983	44.23669	-122.1787	WGS84

Subspecies	Catalog Number	Type of Sample	State	Locality	Year	Latitude	Longitude	Datum
G. o. fuliginosus	HSU-4336	Museum	Oregon	Lane Co.	1983	44.17751	-122.4364	WGS84
G. o. fuliginosus	HSU-4563	Museum	Oregon	Josephine Co.	1985	42.13429	-123.4142	WGS84
G. o. fuliginosus	HSU-4870	Museum	Oregon	Josephine Co.	1984	42.08951	-123.4156	WGS84
G. o. fuliginosus	HSU-7615	Frozen Tissue	Washington	Skamania Co.	1999	46.00769	-122.0106	WGS84
G. o. fuliginosus	HSU-7746	Frozen Tissue	Washington	Skamania Co.	1999	46.37827	-121.5765	WGS84
G. o. fuliginosus	HSU-7747	Frozen Tissue	Washington	Skamania Co.	1999	46.37862	-121.5768	WGS84
G. o. fuliginosus	HSU-7748	Frozen Tissue	Washington	Skamania Co.	1999	46.37814	-121.5767	WGS84
G. o. fuliginosus	HSU-7749	Frozen Tissue	Washington	Skamania Co.	1999	46.36283	-121.6083	WGS84
G. o. fuliginosus	HSU-7750	Frozen Tissue	Washington	Skamania Co.	1999	46.00351	-122.0556	WGS84
G. o. fuliginosus	HSU-7751	Frozen Tissue	Washington	Skamania Co.	1999	46.00702	-122.0071	WGS84
G. o. lascivus	UMMZ-79766	Museum	California	El Dorado Co.	1926	38.92833	-120.1478	NAD27
G. o. lascivus	UMMZ-79769	Museum	California	Placer Co.	1926	39.16021	-120.141	NAD27
G. o. lascivus	UCLA-1913	Museum	California	El Dorado Co.	1919	38.75	-119.9833	WGS84
G. o. lascivus	UCLA-1915	Museum	California	El Dorado Co.	1919	38.75	-119.9833	WGS84
G. o. lascivus	UCLA-1920	Museum	California	El Dorado Co.	1919	38.75	-119.9833	WGS84
G. o. lascivus	UCLA-1923	Museum	California	El Dorado Co.	1919	38.75	-119.9833	WGS84
G. o. lascivus	UCLA-1899	Museum	California	El Dorado Co.	1919	38.75	-119.9833	WGS84
G. o. lascivus	UCLA-1911	Museum	California	El Dorado Co.	1919	38.75	-119.9833	WGS84
G. o. lascivus	UCLA-1912	Museum	California	El Dorado Co.	1919	38.75	-119.9833	WGS84
G. o. lascivus	UCLA-1928	Museum	California	El Dorado Co.	1919	38.75	-119.9833	WGS84
G. o. lascivus	KU-133429	Museum	California	Lassen Co.	1977	41.08368	-120.097	NAD27
G. o. lascivus	KU-142386	Museum	California	El Dorado Co.	1973	38.85444	-120.3699	NAD27
G. o. lascivus	KU-142906	Museum	California	Lassen Co.	1958	40.4164	-121.3345	NAD27
G. o. lascivus	KU-142911	Museum	California	Lassen Co.	1958	40.4164	-121.3345	NAD27
G. o. lascivus	KU-142912	Museum	California	Lassen Co.	1958	40.4164	-121.3345	NAD27
G. o. lascivus	KU-142916	Museum	California	Lassen Co.	1958	40.4164	-121.3345	NAD27
G. o. lascivus	KU-142921	Museum	California	Lassen Co.	1957	40.4164	-120.9363	NAD27
G. o. lascivus	KU-142927	Museum	California	Lassen Co.	1959	40.54684	-120.6519	NAD27
G. o. lascivus	LACM-10643	Museum	California	Tulare Co.	1938	36.10332	-118.5365	NAD27

Subspecies	Catalog Number	Type of Sample	State	Locality	Year	Latitude	Longitude	Datum
G. o. lascivus	LACM-4859	Museum	California	Fresno Co.	1936	36.7477	-118.9759	WGS84
G. o. lascivus	MVZ-21863	Museum	California	Mariposa Co.	1914	37.73669	-119.5982	NAD27
G. o. lascivus	MVZ-23457	Museum	California	Mariposa Co.	1915	37.73082	-119.3923	NAD27
G. o. lascivus	MVZ-24368	Museum	California	Fresno Co.	1916	37.2317	-119.235	NAD27
G. o. lascivus	MVZ-24376	Museum	California	Fresno Co.	1916	37.2317	-119.235	NAD27
G. o. lascivus	MVZ-30092	Museum	California	Tulare Co.	1919	36.57978	-118.7574	NAD27
G. o. lascivus	MVZ-31805	Museum	California	Mariposa Co.	1920	37.75431	-120.108	NAD27
G. o. lascivus	MVZ-109021	Museum	California	Tulare Co.	1934	36.65525	-118.8116	NAD27
G. o. lascivus	MVZ-201571	Frozen Tissue	California	Mariposa Co.	2003	37.71515	-119.665	WGS84
G. o. lascivus	MVZ-201572	Frozen Tissue	California	Mariposa Co.	2003	37.75139	-119.7915	WGS84
G. o. lascivus	MVZ-207299	Frozen Tissue	California	Mariposa Co.	2004	37.77441	-119.5692	WGS84
G. o. lascivus	MVZ-218050	Frozen Tissue	California	Tehama Co.	2006	40.34688	-121.6293	WGS84
G. o. lascivus	MVZ-222759	Frozen Tissue	California	Nevada Co.	2008	39.42091	-120.291	WGS84
G. o. lascivus	MVZ-228422	Frozen Tissue	California	Plumas Co.	2005	40.33739	-121.1094	WGS84
G. o. lascivus	MVZ-216223	Frozen Tissue	California	Mariposa Co.	2005	37.75817	-119.8003	WGS84
G. o. lascivus	MVZ-216224	Frozen Tissue	California	Mariposa Co.	2005	37.75817	-119.8003	WGS84
G. o. lascivus	MVZ-217498	Frozen Tissue	California	Shasta Co.	2004	40.48484	-121.4212	WGS84
G. o. lascivus	MVZ-222760	Frozen Tissue	California	Sierra Co.	2008	39.45815	-120.2862	WGS84
G. o. lascivus	MVZ-232971	Frozen Tissue	California	Madera Co.	2016	37.62826	-119.0868	WGS84
G. o. lascivus	MVZ-225094	Frozen Tissue	California	Shasta Co.	2008	41.12669	-122.2768	NAD27
G. o. lascivus	HSU-6328	Museum	California	Plumas Co.	1990	40.43144	-121.1087	WGS84
G. o. lascivus	HSU-6329	Museum	California	Plumas Co.	1990	40.43145	-121.1086	WGS84
G. o. lascivus	HSU-6347	Museum	California	Plumas Co.	1991	40.42127	-121.1169	WGS84
G. o. lascivus	HSU-6348	Museum	California	Plumas Co.	1991	40.42223	-121.1173	WGS84
G. o. lascivus	HSU-8177	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8178	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8179	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8180	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8181	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8182	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84

Subspecies	Catalog Number	Type of Sample	State	Locality	Year	Latitude	Longitude	Datum
G. o. lascivus	HSU-8183	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8184	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-7891	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-7915	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8186	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8187	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8188	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8189	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8190	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8191	Frozen Tissue	California	Plumas Co.	no date	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8192	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8193	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8194	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8195	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8196	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. lascivus	HSU-8197	Frozen Tissue	California	Plumas Co.	1992	40.43322	-121.1094	WGS84
G. o. oregonensis	HSU-7616	Frozen Tissue	Washington	Thurston Co.	1999	46.88667	-123.1238	WGS84
G. o. stephensi	HSU-2102	Museum	California	Humboldt Co.	1976	40.7611	-123.8716	WGS84
G. o. stephensi	HSU-4199	Museum	California	Humboldt Co.	1982	40.93957	-123.6314	WGS84
G. o. stephensi	HSU-4200	Museum	California	Trinity Co.	1983	40.89175	-123.5828	WGS84
G. o. stephensi	HSU-4239	Museum	California	Trinity Co.	1982	40.83152	-123.4506	WGS84
G. o. stephensi	HSU-7142	Museum	California	Humboldt Co.	1992	40.846	-124.0536	WGS84
G. o. stephensi	HSU-7461	Museum	California	Humboldt Co.	1996	40.56975	-124.0352	WGS84
G. o. stephensi	HSU-7462	Museum	California	Humboldt Co.	1996	40.58429	-124.0356	WGS84
G. o. stephensi	HSU-7611	Museum	California	Humboldt Co.	2000	40.84623	-124.0155	WGS84
G. o. stephensi	HSU-8482	Frozen Tissue	California	Humboldt Co.	2011	40.7611	-123.8716	WGS84
G. o. stephensi	HSU-8481	Museum	California	Humboldt Co.	2012	40.88777	-124.0759	WGS84
G. o. stephensi	HSU-1836	Museum	California	Humboldt Co.	1975	40.7611	-123.8716	WGS84

Appendix B

Table of read depth per sample. A total of 13,824,680 raw reads were produced from two Illumina MiSeq sequencing runs (240,917 cyt-*b* reads, and 8,306,496 microsatellite reads).

*HSU sample that has not been accessioned yet.

Catalog Number	**Total raw reads**	**mtDNA reads**	**Microsatellite reads**	**Other/non-mapped reads**
BBAZ-M16001	22,778	279	10,534	11,965
BBAZ-M16002	30,744	1,034	13,870	15,840
BBAZ-M16003	30,792	1,153	14,754	14,885
HSU-4563	32,264	618	13,739	17,907
HSU-4870	68,334	2,597	36,200	29,537
HSU-6328	9,784	130	5,879	3,775
HSU-6329	39,320	488	29,133	9,699
HSU-6347	25,470	414	13,236	11,820
HSU-6348	13,224	189	8,915	4,120
HSU-7142	18,984	547	12,721	5,716
HSU-7461	31,336	385	17,201	13,750
HSU-7462	60,998	1,942	36,315	22,741
HSU-7611	19,890	672	14,019	5,199
HSU-7615	322,690	22,606	201,994	98,090
HSU-7616	347,194	36,176	240,331	70,687

Catalog Number	Total raw reads	mtDNA reads	Microsatellite reads	Other/non-mapped reads
HSU-2102	6,800	96	3,425	3,279
HSU-7746	331,310	267	24,318	306,725
HSU-7747	312,856	27,624	182,592	102,640
HSU-7748	106,734	648	56,840	49,246
HSU-7749	369,620	544	237,483	131,593
HSU-7750	427,904	869	271,405	155,630
HSU-7751	362,172	2,819	232,463	126,890
HSU-8177	210,226	1,268	130,812	78,146
HSU-8178	296,366	3,378	170,993	121,995
HSU-8179	173,566	497	107,433	65,636
HSU-8180	158,834	981	101,670	56,183
HSU-8181	96,746	180	63,149	33,417
HSU-8182	520,486	4,201	326,342	189,943
HSU-8183	168,298	228	127,204	40,866
HSU-8184	50,842	481	23,354	27,007
HSU-7891	125,342	688	67,422	57,232
HSU-7915	298,970	1,877	164,379	132,714
HSU-8186	40,276	127	20,604	19,545
HSU-8187	243,656	1,566	164,395	77,695
HSU-8188	259,634	20,431	181,589	57,614

Catalog Number	Total raw reads	mtDNA reads	Microsatellite reads	Other/non-mapped reads
HSU-8189	64,562	185	51,620	12,757
HSU-8190	180,092	169	131,759	48,164
HSU-8191	208,424	73	154,432	53,919
HSU-8192	119,262	136	95,572	23,554
HSU-8193	246,582	2,112	183,815	60,655
HSU-8194	89,102	97	54,929	34,076
HSU-8195	225,412	983	172,587	51,842
HSU-8196	92,424	431	64,354	27,639
HSU-8197	154,962	662	102,091	52,209
HSU-8482	410,454	24,968	284,708	100,778
HSU-8481	18,916	643	11,496	6,777
HSU VM-3095*	237,678	10,505	184,750	42,423
HSU-4199	13,160	263	8,104	4,793
HSU-4200	16,928	303	11,667	4,958
HSU-4239	6,988	190	3,443	3,355
HSU-4333	9,110	211	4,572	4,327
HSU-4336	34,326	495	20,028	13,803
HSU-1836	6,548	88	3,317	3,143
KU-133429	36,564	988	21,509	14,067
KU-142386	32,382	597	16,039	15,746

Catalog Number	Total raw reads	mtDNA reads	Microsatellite reads	Other/non-mapped reads
KU-142906	37,262	765	20,838	15,659
KU-142911	31,756	394	17,573	13,789
KU-142912	35,508	827	17,648	17,033
KU-142916	31,730	850	17,893	12,987
KU-142921	42,720	755	25,121	16,844
KU-142927	39,478	779	21,363	17,336
KU-46261	155,984	211	23,633	132,140
KU-46262	55,132	855	34,655	19,622
KU-46263	64,666	2,571	41,385	20,710
KU-46264	103,032	5,149	46,846	51,037
KU-46265	46,862	830	21,251	24,781
KU-46266	38,420	547	18,043	19,830
KU-46267	38,172	592	14,672	22,908
KU-46268	37,008	586	18,648	17,774
KU-46269	32,718	904	15,532	16,282
LACM-10643	88,598	425	55,761	32,412
LACM-4826	77,968	128	59,684	18,156
LACM-4859	130,896	422	82,022	48,452
LACM-8094	134,338	131	86,878	47,329
LACM-871	88,088	137	55,244	32,707

Catalog Number	Total raw reads	mtDNA reads	Microsatellite reads	Other/non-mapped reads
LACM-920	118,516	76	80,079	38,361
LACM-921	57,358	106	49,531	7,721
LACM-95619	61,226	342	46,439	14,445
MVZ-109021	60,338	585	51,133	8,620
MVZ-11002	93,992	5,496	51,861	36,635
MVZ-132650	72,014	38	46,233	25,743
MVZ-132653	24,232	398	18,722	5,112
MVZ-13303	94,652	473	61,959	32,220
MVZ-13306	147,456	471	96,819	50,166
MVZ-13309	65,664	210	48,152	17,302
MVZ-176126	49,152	965	41,025	7,162
MVZ-176127	26,326	308	20,875	5,143
MVZ-176128	18,376	501	13,039	4,836
MVZ-201571	69,494	53	36,303	33,138
MVZ-201572	34,646	63	21,852	12,731
MVZ-207299	92,676	134	56,354	36,188
MVZ-2088	118,054	83	40,934	77,037
MVZ-216223	180,588	311	92,674	87,603
MVZ-216224	118,208	175	72,767	45,266
MVZ-217498	66,184	72	40,635	25,477

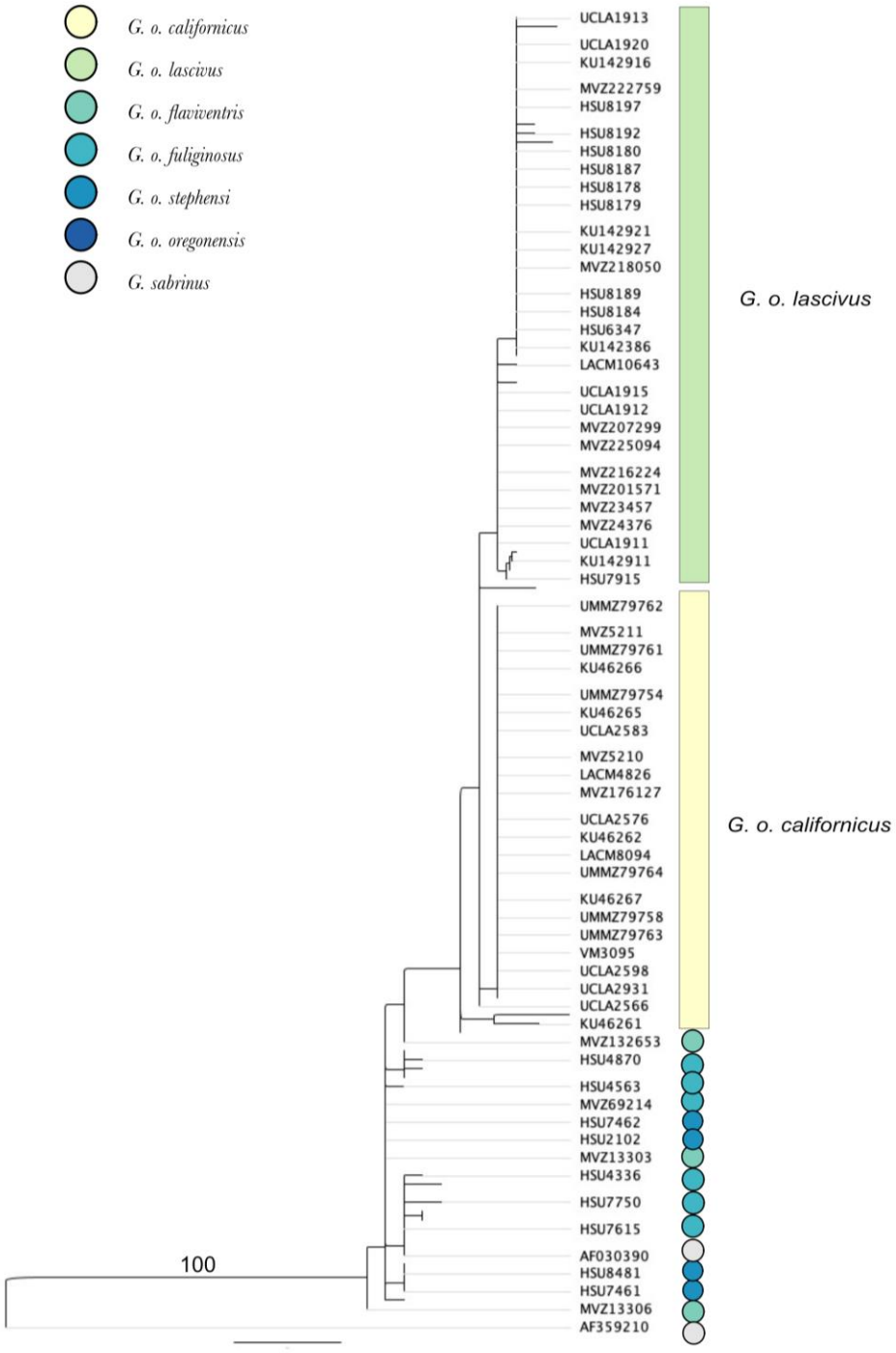
Catalog Number	Total raw reads	mtDNA reads	Microsatellite reads	Other/non-mapped reads
MVZ-218050	104,662	155	44,928	59,579
MVZ-21863	89,452	532	54,573	34,347
MVZ-222759	149,288	159	79,156	69,973
MVZ-222760	93,266	111	58,717	34,438
MVZ-225094	153,356	141	106,636	46,579
MVZ-228422	163,362	212	97,650	65,500
MVZ-232971	92,482	62	68,999	23,421
MVZ-23457	134,782	4,605	86,212	43,965
MVZ-24368	50,368	366	40,506	9,496
MVZ-24376	68,050	130	53,483	14,437
MVZ-30092	60,456	1,315	43,424	15,717
MVZ-31805	80,960	939	51,146	28,875
MVZ-5210	129,212	1,179	72,461	55,572
MVZ-5211	140,970	1,676	1,235	138,059
MVZ-69214	86,142	140	65,507	20,495
MVZ-69216	43,208	81	31,874	11,253
MVZ-7007	135,060	2,004	82,788	50,268
SDNHM-001	28,908	957	14,436	13,515
SDNHM-002	27,552	692	17,139	9,721
SDNHM-003	29,596	972	18,131	10,493

Catalog Number	Total raw reads	mtDNA reads	Microsatellite reads	Other/non-mapped reads
UCLA-17237	45,540	255	18,868	26,417
UCLA-17238	22,918	301	14,775	7,842
UCLA-1899	41,370	545	27,506	13,319
UCLA-1911	52,782	247	25,521	27,014
UCLA-1912	44,420	187	21,382	22,851
UCLA-1913	49,254	148	28,875	20,231
UCLA-1915	41,200	299	17,679	23,222
UCLA-1920	37,344	115	20,223	17,006
UCLA-1923	42,186	83	22,390	19,713
UCLA-1928	52,342	253	32,836	19,253
UCLA-2566	36,312	150	21,333	14,829
UCLA-2573	45,686	467	24,420	20,799
UCLA-2576	48,570	874	30,331	17,365
UCLA-2583	41,492	573	24,758	16,161
UCLA-2597	41,084	1,896	17,448	21,740
UCLA-2598	28,546	347	16,216	11,983
UCLA-2931	31,790	860	17,917	13,013
UCLA-7487	40,548	330	2,519	37,699
UMMZ-79753	46,750	309	25,734	20,707
UMMZ-79754	23,060	439	9,615	13,006

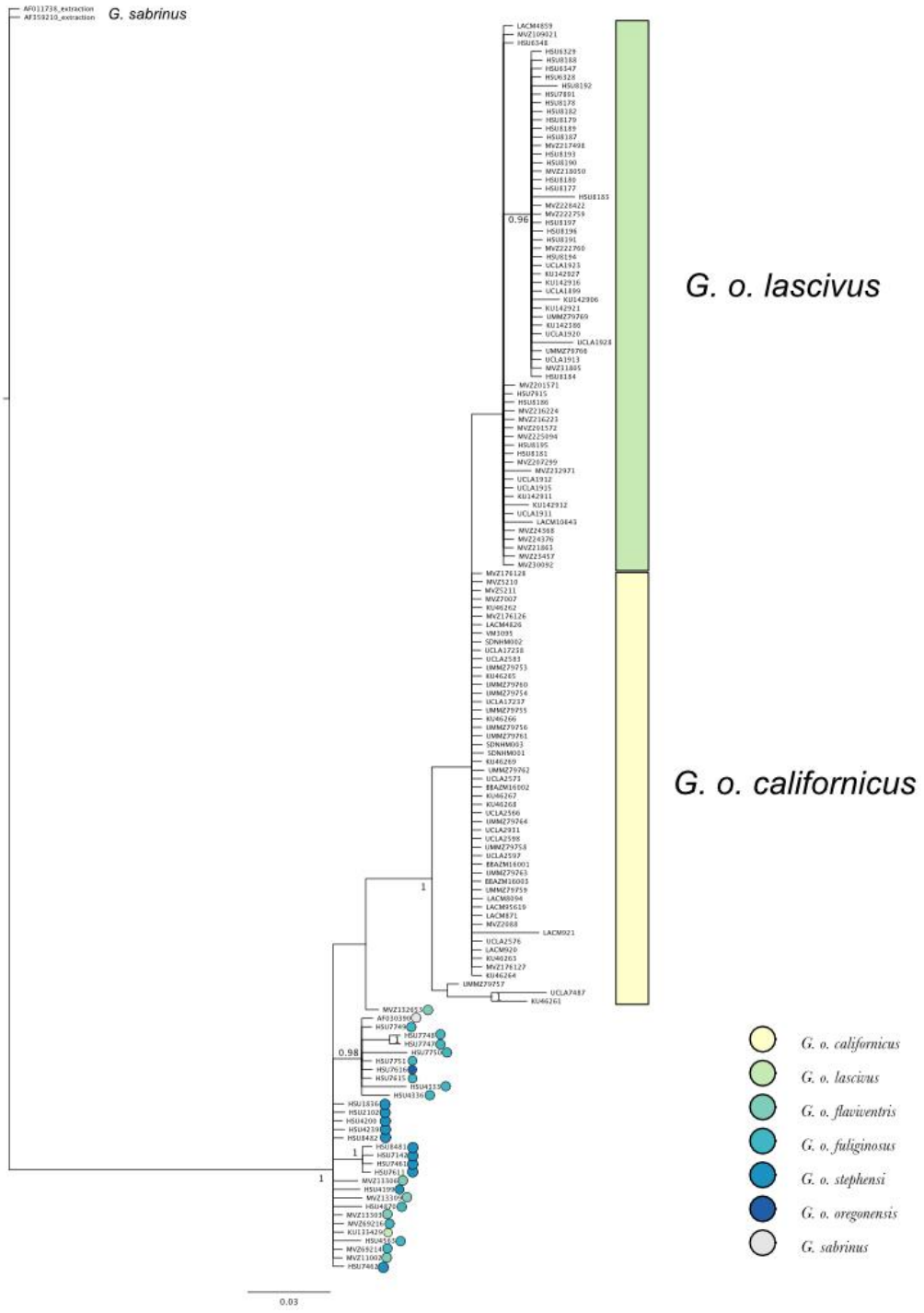
Catalog Number	Total raw reads	mtDNA reads	Microsatellite reads	Other/non-mapped reads
UMMZ-79755	23,618	589	11,398	11,631
UMMZ-79756	52,562	124	11,459	40,979
UMMZ-79757	31,094	21	21,163	9,910
UMMZ-79758	39,032	923	19,404	18,705
UMMZ-79759	20,972	260	12,253	8,459
UMMZ-79760	26,736	617	15,322	10,797
UMMZ-79761	38,590	465	20,127	17,998
UMMZ-79762	25,980	239	16,558	9,183
UMMZ-79763	40,320	441	21,616	18,263
UMMZ-79764	38,630	456	19,025	19,149
UMMZ-79766	51,518	572	21,771	29,175
UMMZ-79769	43,498	287	17,472	25,739
Total	**13,824,680**	**240,917**	**8,306,496**	**5,277,267**

Appendix C

Phylogenetic trees generated in maximum likelihood and Bayesian frameworks using a 300 bp alignment of 145 *G. oregonensis* and three *G. sabrinus* cyt-*b* sequences. The model of nucleotide substitution used was GTR + G as selected by PartitionFinder (Lanfear et al. 2016). The split between species was well-supported, but the branch support within *G. oregonensis* was not, due to the limited base pairs sequenced.



Maximum likelihood tree built in PhyML (Guindon et al. 2010).



Bayesian phylogenetic tree built in MrBayes (Ronquist et al. 2012). Only supported nodes are shown. If no value is provided that bifurcation was not well supported.

Appendix D

Selection of CHIIMP output

Link to a webpage containing the full CHIIMP genotyping results. It includes information on allele alignments, PCR stutter, and samples with more than two prominent sequences:

{HYPERLINK "/Users/HawkFarm/Documents/glaucomys_chiimp/Glaucomys Thesis_5%_5_20_9 loci/report.html"}

Appendix E

Nuclear microsatellite summary statistics for *G. o. flaviventris*, *G. o. fuliginosus*, *G. o. oregonensis*, and *G. o. stephensi* at nine loci. The abbreviations listed are as follows: number of individuals successfully genotyped at that locus (n), total number of alleles (A), number of private alleles (A_p), observed heterozygosity (H_o), expected heterozygosity (H_e), fixation index (F), and Hardy-Weinberg Equilibrium deviation (HWE). Significant P-values are displayed in bold.

G. o. flaviventris

Locus	**n**	**A**	**A_p**	**H_o**	**H_e**	**F**	**HWE**
GS-2	6	8.000	1.000	1.000	0.847	-0.180	0.628
GS-4	5	6.000	1.000	1.000	0.780	-0.282	**0.050**
GS-8	6	5.000	0.000	0.667	0.681	0.020	0.384
GS-10	6	8.000	0.000	1.000	0.861	-0.161	0.363
GS-13	6	6.000	1.000	0.833	0.778	-0.071	0.065
GS-16	6	4.000	0.000	0.500	0.417	-0.200	0.995
GLSA-22	6	7.000	1.000	1.000	0.833	-0.200	0.400
GLSA-52	2	2.000	0.000	0.500	0.375	-0.333	0.637
GLSA-65	6	3.000	0.000	1.000	0.569	-0.756	0.112
Mean	5.444	5.444	-	0.833	0.682	-0.240	-
SE	0.444	0.709	-	0.073	0.062	0.073	-

G. o. fuliginosus

Locus	**n**	**A**	**A_p**	**H_o**	**H_e**	**F**	**HWE**
GS-2	13	8.000	1.000	0.769	0.846	0.091	**0.007**
GS-4	5	5.000	0.000	0.200	0.740	0.730	**0.029**
GS-8	13	5.000	0.000	0.615	0.716	0.140	0.307
GS-10	13	8.000	0.000	0.846	0.840	-0.007	**0.008**
GS-13	13	7.000	0.000	0.846	0.787	-0.075	**<0.001**
GS-16	13	4.000	0.000	0.308	0.512	0.399	**0.012**
GLSA-22	13	7.000	1.000	1.000	0.787	-0.271	0.229
GLSA-52	13	7.000	2.000	0.538	0.728	0.260	0.067
GLSA-65	13	3.000	0.000	0.462	0.447	-0.033	0.931
Mean	12.111	6.000	-	0.621	0.711	0.137	-
SE	0.889	0.601	-	0.089	0.047	0.098	-

G. o. oregonensis

Locus	**n**	**A**	**A_p**	**H_o**	**H_e**	**F**	**HWE**
GS-2	1	2.000	0	1.000	0.500	-1.000	0.317
GS-4	1	2.000	0	1.000	0.500	-1.000	0.317
GS-8	1	2.000	0	1.000	0.500	-1.000	0.317
GS-10	1	2.000	0	1.000	0.500	-1.000	0.317
GS-13	1	2.000	0	1.000	0.500	-1.000	0.317
GS-16	1	1.000	0	0.000	0.000	-	-
GLSA-22	1	2.000	0	1.000	0.500	-1.000	0.317
GLSA-52	1	1.000	0	0.000	0.000	-	-
GLSA-65	1	2.000	0	1.000	0.500	-1.000	0.317
Mean	1.000	1.778	-	0.778	0.389	-1.000	-
SE	0.000	0.147	-	0.147	0.073	0.000	-

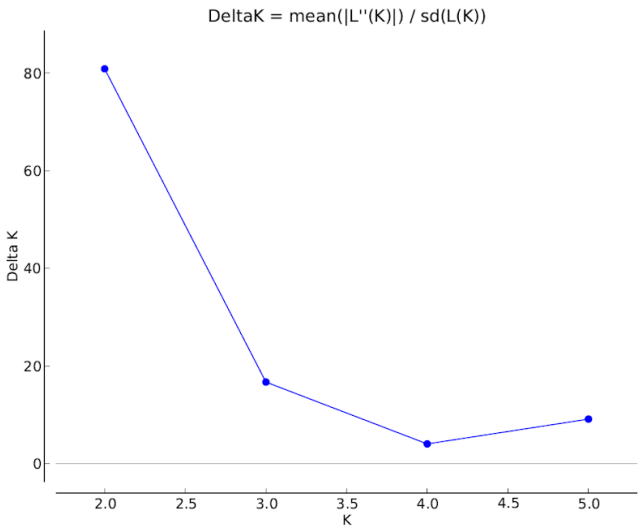
G. o. stephensi

Locus	**n**	**A**	**A_p**	**H_o**	**H_e**	**F**	**HWE**
GS-2	11	7.000	0.000	0.909	0.826	-0.100	0.569
GS-4	10	7.000	2.000	0.400	0.820	0.512	**0.023**
GS-8	11	5.000	0.000	0.818	0.731	-0.119	0.058
GS-10	11	10.000	1.000	0.909	0.888	-0.023	**0.028**
GS-13	11	7.000	0.000	0.818	0.789	-0.037	0.469
GS-16	11	5.000	0.000	0.818	0.669	-0.222	0.369
GLSA-22	11	4.000	0.000	0.909	0.678	-0.341	**0.015**
GLSA-52	11	5.000	1.000	0.455	0.653	0.304	0.418
GLSA-65	11	2.000	0.000	0.455	0.434	-0.048	0.875
Mean	10.889	5.778	-	0.721	0.721	-0.008	-
SE	0.111	0.760	-	0.073	0.045	0.087	-

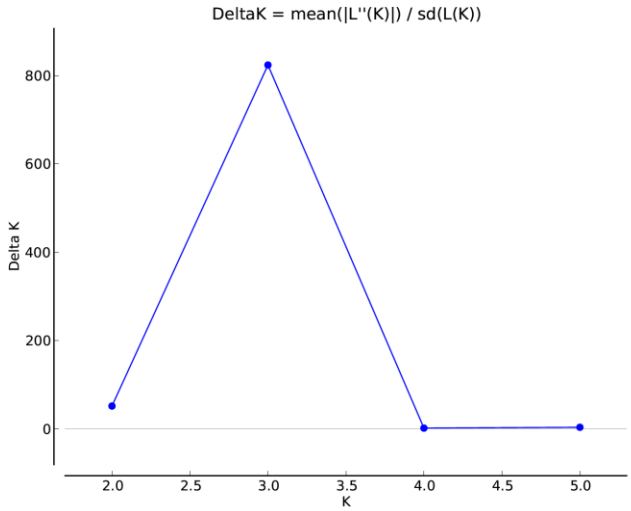
Appendix F

Additional STRUCTURE HARVESTER (Earl and vonHoldt 2012) Data

Evanno et al. (2005) correction of STRUCTURE output in the admixture model (K=2):



Evanno et al. (2005) correction of STRUCTURE output in the no admixture model (K=3):



Appendix G

Summary of *Glaucomys oregonensis* sample type and extraction protocol by subspecies.

Subspecies	**Museum**[a]	**Frozen Tissue**[b]	**Frozen Tissue (Whole Organism)**[b]	**Degraded Tissue (Whole Tail)**[a]	**Hair in RNALater**[a]	**Total (Subspecies)**
G. o. californicus	44	0	1	2	4	51
G. o. lascivus	31	34	0	0	0	65
G. o. flaviventris	6	0	0	0	0	6
G. o. fuliginosus	6	7	0	0	0	13
G. o. oregonensis	0	1	0	0	0	1
G. o. stephensi	10	1	0	0	0	11
Total (Sample Type)	97	43	1	2	4	147

[a]Extracted using Qiagen QIAamp DNA Mini Kits (Qiagen, Valencia, CA).

[b]Extracted using Qiagen DNeasy Blood and Tissue Kits (Qiagen, Valencia, CA).