Building a Better Ark:
Theoretical and Analytical Approaches for Managing Species at the Population Level

by

Logan A. Volkmann

B.Sc., Western Washington University, 2011

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Biological Sciences
Faculty of Science

SIMON FRASER UNIVERSITY

Summer 2014

Approval

Name:	**Logan A. Volkmann**
Degree:	**Master of Science (Biological Sciences)**
Title:	***Building a Better Ark: Theoretical and Analytical Approaches for Managing Species at the Population Level***
Examining Committee:	**Chair:** Sherryl Bisgrove Associate Professor

Arne Mooers
Senior Supervisor
Professor

Michael Hart
Supervisor
Professor

Mark Statham
Supervisor
Project Scientist
Veterinary Genetics Laboratory
University of California, Davis

Michael Russello
External Examiner
Associate Professor
Department of Biology
University of British Columbia
(Okanagan)

Date Defended/Approved: June 09, 2014

Partial Copyright Licence

Abstract

Biodiversity losses and limited resources may soon call for the preservation of key populations rather than entire species. However, successful population-level management requires both an understanding of where evolutionarily distinct taxa occur on the landscape and an efficient method for prioritizing taxa based on survey data. The present study addresses these needs. I begin by investigating the genetic identity and origins of red foxes (*Vulpes vulpes*) in North America's Intermountain West. I then demonstrate a new approach for prioritizing populations that extends the metrics for evolutionary isolation from phylogenetic trees to phylogenetic networks, using two example species. Patterns of genetic differentiation for red foxes are consistent with endemism or natural range expansion in the Intermountain West, making this population a potential conservation target. Heuristic networks generated for spotted owls (*Strix occidentalis*) and mountain pygmy-possums (*Burramys parvus*) show how the approach can highlight peripheral populations that may merit increased conservation attention.

Keywords: conservation genetics; extinction; peripheral populations; phylogeography; ranking; red fox

Dedication

To my parents, who have never stopped believing in me.

Acknowledgements

Many people made this thesis possible. Firstly I thank Arne Mooers for his support and collaboration on this project, especially in regards to the creation and publication of the phylogenetic networks study (Chapter 3). I am also grateful to co-authors Iain Martyn, Vincent Moulton, and Andreas Spillner for their feedback on the Chapter 3 manuscript as it bounced among journals.

I thank Ben Sacks and Mark Statham for their advice and encouragement during my time at the University of California, Davis and throughout the development of Chapter 2. I would also like to thank Keith Aubry for providing an early direction to my red fox study and Quentin Voyce (though we never met) for developing the initial sequencing primers. Jocelyn Akins and Mason Reid provided a few much-needed interludes from lab work by inviting me to help with camera trap maintenance and radio telemetry projects in the Washington Cascades.

This thesis has gone through innumerable (and at times substantial) revisions to reach its final state. I thank Mike Hart for his advice regarding genetic analyses (some of which, unfortunately, did not make it into this thesis), as well as Jeff Joy, Mike Steel, Felix Breden, Ben Sandkam, Iva Popovic, and the FAB Lab for their discussion and comments on the early results of both studies presented below.

Above all I thank Jenn and Joey for all the other adventures along the way.

It's done, guys. Let's go outside.

Table of Contents

List of Tables

List of Figures

Chapter 1.

Introduction

Biodiversity models for the 21st century generally predict widespread range shifts and species losses (Jenkins 2003; Tittensor *et al*. 2010). Extinctions are particularly likely in freshwater (Dudgeon *et al*. 2006), grassland (Thomas *et al*. 2004), and Mediterranean ecosystems (Sala *et al*. 2000), where the effects of land-use pressure and climate change will be most pronounced. Similarly, the loss of species from biodiversity hotspots such as mountain ranges (La Sorte and Jetz 2010) and tropical rainforests (Sodhi *et al*. 2004) may now be unavoidable.

In a world of urgent threats and limited resources we must choose to preserve some species at the expense of others (Vane-Wright *et al*. 1991). To this end, researchers have developed methods to prioritize species for conservation and maximize the biodiversity that *can* be saved (see, *e.g.*, May 1990; Faith 1992; Weitzman 1992). However, extinction often begins below the species level with the loss of isolated ("peripheral") populations (Hughes *et al*. 1997; Ceballos and Ehrlich 2002). These populations can be evolutionarily distinct, possessing lineages with unique biological histories, and their contribution to within-species diversity can be proportionately large (Hampe and Petit 2005). The presence of such diversity has sparked interest in conservation at finer taxonomic scales (*e.g.*, Ryder 1986; Moritz 1994; Manel *et al*. 2003; Segelbacher *et al*. 2010). Indeed, as habitats become fragmented and connectivity declines, the preservation of key populations may be a more feasible alternative to range-wide species conservation.

At present, population-level conservation faces two important challenges. The first is a practical problem: how to define meaningful conservation units on a landscape. Evolutionarily distinct lineages are often cryptic and range-wide genetic surveys can be difficult, leaving gaps in our understanding of a species' population genetic structure. Further, the human-mediated movement of organisms can create misleading patterns of

diversity. The North American red fox (*Vulpes vulpes*), for example, occurs as a patchwork of native (Aubry *et al*. 2009; Sacks *et al*. 2010) and nonnative animals (Statham *et al*. 2012) alongside populations of unknown origin. Management policies for native and nonnative red foxes differ substantially (California Department of Fish and Game 1999), making it imperative to determine the status of each population. In this case, targeted phylogeographic studies would help determine the origins of recently-established populations and provide important baseline knowledge for population-level conservation.

The second problem is more theoretical but perhaps more important: how to prioritize populations for conservation given the observed patterns of diversity. Species-level prioritization uses phylogenetic trees to represent relationships among taxa, but these methods do not always capture the more complex relationships that exist among populations (*e.g.*, Posada and Crandall 2001). In practice, tree-based analyses may overlook important variation and provide an inferior view of conservation needs. New systems of representation, such as phylogenetic networks (Bryant and Moulton 2004), may be better suited to the task of population-level conservation. However, an efficient approach to prioritize populations must first be developed.

The following studies examine both aspects of conservation below the species level: (1) identifying populations of interest and (2) determining which populations will contribute most to future biodiversity. First, I use population genetic methods to assess the identity and origins of red foxes in the Intermountain West region of North America. Then, I introduce a general approach for prioritizing populations based on pairwise distance data and the properties of phylogenetic networks. Results from the former study will be relevant to conservation of native red foxes, while the methods described in the latter study should be widely applicable to species in need of population-level conservation.

Chapter 2.

Genetic distinctiveness of red foxes in the North American Intermountain West as revealed through expanded sequencing of native mitochondrial haplotypes

Abstract

Western North America contains a mosaic of indigenous and introduced red fox (*Vulpes vulpes*) populations. Native red foxes occurred historically in high-elevation subalpine zones of the Rocky Mountains, Cascade Range, and Sierra Nevada, as well as in the desert-like habitat of California's Sacramento Valley. Red foxes have also been sporadically reported in the Great Basin of the Intermountain West since at least the early 1900s, but it was unclear whether these animals represented surviving relicts from the last ice age, occasional dispersers from adjacent montane populations, or early fur-farm escapees. A recent genetic study found Great Basin red foxes to be most closely related to Rocky Mountain red foxes but lacked sufficient resolution to determine whether this affinity reflected ancient or contemporary roots. Red foxes in the Intermountain West carried a native mitochondrial haplotype (A-19), which was also the most basal and widespread among western populations and thus equally consistent with ancient and contemporary connectivity. Here, to increase resolution, I sequenced 4,004 bp of the mitochondrial genome corresponding to A-19 and related haplotypes from historical and modern samples collected throughout western North America. The expanded sequences revealed significant phylogenetic diversity, including six A-19 variants falling into two divergent subclades. Although I found no variants unique to the Intermountain West, the two dominant variants were relatively rare elsewhere, especially among modern samples. An AMOVA consistently identified the Intermountain West population as distinct, ruling out the possibility of large-scale dispersal from the Rocky

Mountains. Together, my findings are most consistent with the long-term persistence of red foxes in the Intermountain West derived from a relictual population in this region or a small founder group from the Rocky Mountains. Human-mediated translocations from local sources may have also played a role, though I was unable to distinguish this process from natural, small-scale range expansion. Additional sampling and high-density nuclear genomic markers are needed to further clarify the origins and status of these red foxes.

Introduction

Geography, climate change, and human presence have influenced the distribution of numerous species in North America, particularly west of the Great Plains (see reviews by Weider and Hobæk 2000; Arbogast and Kenagy 2001; Brunsfeld *et al*. 2001). Regions such as the Alexander Archipelago, the Olympic Peninsula, and the Rocky Mountains have been found to contain ancient, isolated populations with distinct evolutionary histories (Soltis *et al*. 1997; Cook *et al*. 2001). Such populations are often cryptic, *i.e.*, difficult to distinguish phenotypically, but are nonetheless of conservation interest due to their genetic distinctiveness. The discovery of unique populations has, for example, influenced forest management policies (Cook *et al*. 2006) and improved our understanding of how future climate change will affect species distributions (Galbreath *et al*. 2009).

Red foxes (*Vulpes vulpes*) have the largest geographic range of all terrestrial carnivores (Lariviére and Pasitschniak-Arts 1996). Fossil evidence indicates that the species originated in Eurasia 0.5–1 million years ago (ya) (Kurtén 1968) and colonized North America via the Bering land bridge sometime prior to the Sangamon interglacial (>125,000 ya; Péwé and Hopkins 1967), a timeline supported by genetic data (Aubry *et al*. 2009; Kutschera *et al*. 2013). Mitochondrial analysis further indicates the arrival of a subsequent Eurasian matriline during the Wisconsin glaciation (100,000–10,000 ya; Aubry *et al*. 2009). Following the Last Glacial Maximum (LGM, 20,000 ya; Clark *et al*. 2009), lineages south of the ice sheets diverged into two genetically distinct groups (Appendix A). Although the Nearctic clade as a whole exhibited incomplete lineage sorting, the monophyletic Eastern subclade corresponded to red foxes in eastern North America that presumably followed receding glaciers northeastward into Canada, and the monophyletic Mountain subclade corresponded to red foxes in western North America

that presumably tracked boreal-montane habitats elevationally into mountain ranges (Aubry *et al*. 2009; but see also Sacks *et al*. 2010 for an exceptional population). By pre-Columbian times (1,000 ya) migration and vicariant events gave rise to several populations of red foxes in western North America that were closely associated with subalpine forest and parkland habitats in the Rocky Mountains, Cascade Range, and Sierra Nevada (Fig. 1).

These "montane red foxes" are ecologically and genetically unique (Aubry 1983; Perrine 2005; Aubry *et al*. 2009; Kutschera *et al*. 2013), making them a target for current and future conservation efforts. In California, for example, the precipitous decline of Sierra Nevada red foxes (*V. v. necator*) prompted state agencies to designate the subspecies as a protected furbearer in 1974 (Gould 1978). This was followed by a formal petition in 2011 (currently in review) for listing under the United States' Endangered Species Act (Sacramento Fish and Wildlife Office 2011). Similarly, the decline of red foxes in the Washington Cascades (*V. v. cascadensis*) has led to their listing as a state candidate species and a Natural Heritage Critically Imperiled Species (Washington Department of Fish and Wildlife 2008). The disappearance of montane red foxes from parts of the Rocky Mountains (*V. v. macroura*) could warrant similar concern (Aubry 1984). However, conservation decisions are hindered by the presence of nonnative red foxes in other parts of the western North America. While montane red foxes appear to have undergone range reductions over the past century (Zielinski *et al*. 2005), nonnative populations have thrived and expanded their range in many low-elevation areas. Red foxes now inhabit areas such as southern California (Jurek 1992; California Department of Fish and Game 1999; Perrine *et al*. 2007; Sacks *et al*. 2011) and the San Juan Islands of Washington (Aubry 1984) due to intentional releases and fur farm escapes (Statham *et al*. 2012). The westward expansion of red foxes across the Great Plains (reviewed by Kamler and Ballard 2002) could also threaten the genetic integrity of montane populations. Nonnative red foxes might be removed or managed to curb their negative impact on endemic species (Jurek 1992; California Department of Fish and Game 1999), but native populations such as the recently-identified Sacramento Valley red fox (*V. v. patwin*) might conversely merit protection as endemics themselves (Sacks *et al*. 2010). Discriminating between cryptic native and nonnative populations is therefore vital.

Red foxes of uncertain origin occur west of the Rocky Mountains and east of the Cascade Range and Sierra Nevada, (hereafter, the "Intermountain West") an ecologically diverse region that includes the highlands of eastern Oregon, the Great Basin of Nevada and Utah, and the Snake River Plain of southern Idaho (Fig. 1). Some of these areas fall well outside the currently-defined boundaries of native populations (Hall and Kelson 1959; Aubry 1983). Historically, montane red foxes are known to have ranged from mountainous areas of Washington (Taylor and Shaw 1927), Idaho (Merriam and Stejnrger 1891), and Montana (Bailey and Bailey 1918) as far south as the Piute Mountains of California (Grinnell *et al*. 1937) and the Mogollon Mountains of New Mexico (Bailey 1931), but their occurrence in the Great Basin has long been disputed (Merriam 1900; Barnes 1922; Hall 1946; Fichter and Williams 1967). Several authors have noted that the establishment of fur farms in the early 1900s coincided with the appearance of red foxes in valleys and coastal areas (Fichter and Williams 1967; Aubry 1984; Verts and Carraway 1998). These observations led Kamler and Ballard (2002) and others to conclude that that all low-elevation red foxes in western North America were of nonnative origin. Hall (1946), however, found no evidence to doubt that the earliest red fox records from Nevada represented native occurrences. He further suggested that these animals belonged to the same population as the Sierra Nevada red fox—a hypothesis reflected in some early range maps for the subspecies (Hall and Kelson 1959). The ecological requirements of montane red foxes are poorly known (Aubry 1983; Benson *et al*. 2005; Perrine 2005), but their historical distribution in the Intermountain West could have been similar to that of yellow-bellied marmots (*Marmota flaviventris*), which persist on high-elevation "sky islands" of suitable habitat (Floyd *et al*. 2005). Because historical records and morphological data remain inconclusive, the true origin of these foxes must be determined through genetic methods.

Sacks *et al*. (2010) detected native mitochondrial haplotypes in red foxes from the Intermountain West and used microsatellites to confirm their predominantly native ancestry relative to historical and modern populations in western North America. However, sample size was too small to resolve population structure further. Resolution of a 696-bp mitochondrial marker was also too low to determine whether these haplotypes represented geographically distinct matrilines or potentially originated from nearby mountain systems (*e.g.*, via recent range expansion). In particular, the Intermountain West population was dominated by a single native haplotype ("A-19") that

was basal to, and more numerous and widespread than, other haplotypes in the Mountain subclade (Appendix A; Sacks *et al*. 2010). Consequently, evidence was equivocal as to whether this haplotype arose from a recently introduced matriline or one that traced back to the initial expansion of the lineage at the end of the LGM (Statham *et al*. 2012). Expanded sequencing of the A-19 haplotype throughout its distribution in western North America could reveal locally restricted mutations that arose since the LGM, thereby enabling clearer inferences about the origins of red foxes in the Intermountain West.

In the present study I sequenced ~4,000 bp of the mitochondrial genome in a sample of modern red foxes carrying the A-19 haplotype to find additional substitutions (*i.e.*, single nucleotide polymorphisms, SNPs), and then screened historical specimens for these SNPs to better resolve the maternal genetic structure of red foxes in western North America. I used this information to evaluate two hypotheses for the origins of red foxes in the Intermountain West. *Range Expansion* (1) of animals from Sierra Nevada or Rocky Mountain populations would be supported if red foxes in the Intermountain West exhibited haplotypes originating from one of both of these subspecies. *Endemism* (2) would be indicated if red foxes carried distinct haplotypes with mutations distinct but derived from those found in the nearby mountain populations, similar to evidence used to support endemism of the Sacramento Valley red fox (Sacks *et al*. 2010). I conducted spatial analyses to assess the distinctiveness of the Intermountain West population relative to surrounding mountain subspecies based on high-resolution variants of the otherwise widespread A-19 haplotype.

For comparison, I also conducted expanded sequencing of an endemic Sacramento Valley haplotype ("D-19") and the primary haplotype of the Washington Cascades subspecies ("O-24"). Secondary objectives included (1) increasing resolution of the endemic Sacramento Valley red fox haplotype to better assess its phylogenetic relationship to other montane red fox populations, and (2) investigating the origin of A-19 haplotypes in a putative nonnative population in coastal Oregon.

Methods

Samples

The red fox specimens examined here (*n* = 126) were collected throughout western North America between 1891 and 2012. In addition to haplotype A-19, I selected

specimens that had previously sorted into two other common haplogroups in the Mountain subclade. Specifically, I sequenced representatives of the endemic D-19 haplotype from the Sacramento Valley red fox population and the O-24 haplotype from the Washington Cascades and Northern Rockies (both presumably native populations), along with O-24 samples from a putative nonnative population localized near Salt Lake City, Utah (Appendix A; GenBank accession numbers EF064207, EF064209, EF064219, FJ830784, and FJ830785; Perrine *et al*. 2007; Aubry *et al.* 2009). This sample distribution covered the established historical range of western red foxes (Hall and Kelson 1959; Kamler and Ballard 2002; Aubry *et al*. 2009) as well as areas where red foxes were known or suspected to be nonnative (Statham *et al.* 2012). I also included outgroup specimens from Kansas, Alaska, and Eurasia representing the Nearctic Eastern subclade and Holarctic clade (Appendix A; GenBank accession numbers EF064211, EF064212 and GQ374180; Perrine *et al*. 2007; Aubry *et al*. 2009).

I defined two time periods based on the observed temporal distribution of samples (Appendix B). Red foxes collected prior to 1940 (hereafter, "historical"), including those from the Sacramento Valley, were considered to represent populations that existed prior to the advent of fur farming in North America. These populations characterized the pre-Columbian distribution of western haplotypes and would have been the most likely sources of translocated animals. Historical samples consisted of maxilloturbinate bone extracts from specimens at the Museum of Vertebrate Zoology at the University of California, Berkeley (MVZ; n = 18), the National Museum of Natural History in Washington, D.C. (USNM; n = 36), and the Burke Museum at the University of Washington, Seattle (UW; n = 6) (Aubry *et al*. 2009). Specimens collected after 1940 (hereafter, "modern") were considered to originate from natural, extended, or introduced populations. Modern samples consisted of maxilloturbinate bone extracts and skin snips from the Burke Museum (UW; n = 5), the Oregon Game Commission Collection at Oregon State University, Corvallis (OSU; n = 4), and the Slater Museum of Natural History at the University of Puget Sound, Tacoma (UPS; n = 3), along with frozen tissue from the University of California, Davis (UCD, n = 53).

I divided historical and modern samples into geographic regions (hereafter, "populations") based on collection locality, elevation, and the presence of major geographic barriers such as intermountain basins (Fig. 1; Appendix B). Populations of natural origin were defined as Northern Rockies (NR: Utah, Wyoming, Montana, Idaho),

Southern Rockies (SR: New Mexico, Colorado), Northern Cascades (NC: Washington), Southern Cascades (SC: Oregon, California), Sierra Nevada (SN: California), and the Sacramento Valley (SV: California). Based on previously-observed patterns of differentiation among mountain red foxes (Sacks *et al*. 2010), I followed Gordon's (1966) hypothesis that the Columbia River Gorge may be a significant barrier to gene flow for animals in the Washington and Oregon Cascades. Similarly, I considered the Wyoming Basin to be a potential barrier for red fox populations in the Rocky Mountains (Péwé 1983). Populations of undetermined origin were defined as Intermountain West (IW: Nevada, Idaho, Oregon) and Coastal Oregon (CO). In Idaho and Oregon, where the geographic cutoffs between natural and disputed populations were less clear, I assigned samples based on elevation and the documented historical presence of red foxes.

Laboratory procedures

I conducted DNA extraction, polymerase chain reaction (PCR) amplification, and sequencing at the University of California, Davis. Some PCR amplification of modern samples also took place at Simon Fraser University under similar reaction conditions (see below). Project collaborators BN Sacks and MJ Statham extracted modern DNA from tissue samples using DNeasy® tissue kits (Qiagen) and from scats using the QiaAmp® stool kits (Qiagen). DNA extraction of historical samples from maxilloturbinate bones and skin snips followed the phenol-chloroform protocol described in Wisely *et al*. (2004).

PCR and sequencing efforts targeted the red fox mitochondrial genome. Because historical specimens typically yield only degraded DNA in low concentrations, in these samples I focused on amplifying short (100–200 bp) fragments surrounding polymorphic sites, rather than complete loci. To identify polymorphic sites of interest, I first examined a larger section of the mitochondrial genome in modern samples.

I amplified complete sequences (hereafter, "full-locus" sequences) of cytochrome *b* (1,338 bp), the control region (1,432 bp), and cytochrome *c* oxidase subunit I (1,793 bp) in modern samples using primer sets developed by Quentin Voyce at the University of California, Davis. For historical samples, I targeted two fragments (hereafter, "targeted" sequences) of cytochrome *b* (totaling 399 bp) and four fragments of the control region (totaling 601 bp) using primer sets developed with the Primer3 online tool (Rozen and Skaletsky 1998). New primers were tested on a selection of modern

samples before attempting to amplify DNA in historical samples. Appendix C lists the primer sequences used in this study.

To minimize the chance of contamination with modern red fox samples, PCR preparation of historical samples took place under clean-room conditions in a separate facility dedicated to the extraction of ancient DNA (see Aubry *et al*. 2009). I used a total reaction volume of 23 µL consisting of 2 µL DNA extract, 1x PCR buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.1 µg/µL bovine serum albumin (BSA), 0.5 mM of each primer, and 1 U/µL *Taq* polymerase. PCR conditions were set at 94° C for 3 min (initialization), followed by 35 cycles of 94° C for 30 sec, 50° C for 30 sec and 72° C for 45 sec, followed by 72° C for 10 min (final extension). I ran PCR products on 2% agarose gels to confirm that each sample that amplified successfully.

I sequenced PCR products in both directions. For full-locus sequencing of cytochrome *b*, the control region, and cytochrome *c* oxidase subunit I, I used internal primers to obtain overlapping sequence reads (Appendix C). Modern samples prepared at Simon Fraser University were sequenced off-site through Molecular Cloning Laboratories (MCLAB). For all other samples, I conducted dye-terminator sequencing reactions using Applied Biosystems reagents and ran the products on an ABI 3730 capillary sequencer (Applied Biosystems).

I aligned sequences visually against a published reference of the red fox mitochondrial genome (Zhong *et al*. 2010) using MEGA v5 (Tamura *et al*. 2011). Samples that yielded poor sequences were re-amplified and re-sequenced when possible. Polymorphisms with an ambiguous signal were treated conservatively as N's in the final sequence data. For the analyses below (unless otherwise indicated), I used a full-locus dataset to increase statistical power and inferred missing sequence data for historical (targeted) samples by matching them to full-locus haplotypes. Because my targeted primer sets were designed to recover all haplotypes observed in modern (full-locus) samples, this procedure gave me conservative estimates of genetic diversity and differentiation in historical red fox populations (*i.e.*, the complete range of mutations seen in modern samples was also assayed in historical samples, but some diversity unique to the historical samples could have gone undetected).

Data analyses

I constructed median-joining networks (Bandelt *et al.* 1999) under default parameters in Network v4.612 (Fluxus Technology) to describe the relationships among full-locus haplotypes. Polymorphisms on all mitochondrial regions were analyzed together and historical and modern samples were pooled. I also constructed a maximum likelihood (ML) tree of full-locus haplotypes in MEGA v5 (Tamura *et al*. 2011) using an HKY model of DNA substitution (Hasegawa *et al*. 1995), and calculated bootstrap support for the nodes based on 1,000 replicates.

To assess the extent to which my expanded sequences provided greater phylogenetic resolution in a geographic context, I investigated genetic differentiation among populations based on A-19 variants. First, I performed an analysis of molecular variance (AMOVA; Excoffier *et al*. 1992) in Arlequin v3.5 (Excoffier and Lischer 2010). I then generated matrices of pairwise genetic distance estimates based on haplotype frequencies (F_{ST}; Weir and Cockerham 1984) and on nucleotide differences between haplotypes (Φ_{ST}; Excoffier *et al*. 1992). Evaluation of statistical significance ($\alpha = 0.05$) for F_{ST} and Φ_{ST} values was based on 1,000 permutations; I used the sequential Bonferroni method to correct for multiple tests (Rice 1989).

I also examined the geographic structure of populations using SAMOVA v1.0 (Dupanloup *et al*. 2002), which is computationally similar to AMOVA but incorporates spatial orientation of samples explicitly. To investigate how populations were hierarchically structured, I tested a series of models where the assumed number of population groupings (*K*) ranged from 2 to 6. In general, populations that partition into separate groups at low increments of *K* will be more distinct from each other than those that partition only in higher-order models. Samples nearest the center of each geographic region (as described above) were used to define latitude and longitude for each population. I centered the Intermountain West population at the Independence Mountains (Elko County, Nevada), which is approximately equidistant from the marginal red fox records reported by Bailey (1936), Hall (1946), and Fichter and Williams (1967). All analyses were conducted with the same input files at 100 initial conditions. I followed recommendations in Dupanloup *et al*. (2002) to identify the *K* model that best described the dataset.

To assess population differentiation graphically, I created phylogenetic networks in SplitsTree v4.11 (Huson and Bryant 2006) using the pairwise genetic distance matrices generated in Arlequin. Unlike a phylogenetic tree, which attempts to fit the

underlying data to a single evolutionary scenario, a phylogenetic network is capable of representing many alternative scenarios in a single diagram (Bryant and Moulton 2004). In situations where genetic distance data are not tree-like, as is often the case when neighboring populations exchange migrants (Posada and Crandall 2001), a network can provide a more informative representation of relationships among populations.

Results

I obtained full-locus sequences (4,002–4,004 bp) from 42 higher-quality modern red fox specimens, from which I identified 7 distinct haplotypes (Fig. 2a). Five of these haplotypes (*w*1-*w*3, *w*6 and *w*7) corresponded to the widespread A-19 haplogroup while haplotypes *w*4 and *w*5 corresponded to D-19 and O-24, respectively. A-19 variants differed from each other by as many as 9 base changes and contained 17 variable sites (14 transitions, 1 transversion, and 2 indels).

I then obtained targeted sequences (755–757 bp) containing 10 variable sites, which were sufficient to discern all 7 haplotypes above, from an additional 42 historical red fox specimens and 13 lower-quality modern specimens. From these samples I identified one additional haplotype (*w*8) within the A-19 haplogroup (Fig. 2a). Voucher sequences were deposited in GenBank.

I observed a heterogeneous distribution of native haplotypes on the landscape (Fig. 3; Appendix D). Notably, the two most common and widespread haplotypes in the Intermountain West (*w*6 and *w*7) were not found in any modern samples from mountain populations and occurred infrequently in historical samples (Appendix B). Outside the Intermountain West, I detected haplotype *w*6 in a single historical sample from the Southern Rockies. Haplotype *w*7 occurred historically at similar latitudes on both sides of the Great Basin (Southern Cascades and Northern Rockies), whereas the closely-related haplotype *w*8 was found only in the Southern Cascades. I did not detect *w*8 in any modern (full-locus) samples, and my analyses using inferred sequence data treated this haplotype conservatively as *w*7 (Fig. 2b).

Although haplotypes *w*6 and *w*7 were moderately divergent from each other, they nonetheless formed a well-supported subclade on the ML tree separate from all other haplotypes (Fig. 2b). I also found strong support for a subclade containing the Sacramento Valley haplotype *w*4 and the widespread haplotype *w*3, which occurred in Coastal Oregon and several mountain populations (Fig. 3).

The AMOVA based on A-19 variants revealed significant differentiation for the Intermountain West and Lowland Oregon ($P << 0.001$; Fig. 4a). These populations occurred at the tips of distinct, long branches on the phylogenetic networks, reflecting relatively high levels of genetic distinctiveness (Fig. 4b). In contrast, most mountain populations were poorly differentiated from each other and tended to occupy nodes in more central positions.

The SAMOVA based on A-19 variants also supported the distinctiveness of the Intermountain West population. Φ_{CT} values were similar for all models tested and significant at $K = 3$ and $K = 5$ (Table 1). At all increments of K, the Intermountain West partitioned into an exclusive group. The most conservative model justified by the data was $K = 3$ ($P = 0.009$), which additionally assigned Coastal Oregon and the Northern Rockies together to a third group. The model with the highest overall support was $K = 5$ ($P = 0.008$), above which group structure began to dissolve and Φ_{CT} began to decline (Appendix E; Dupanloup *et al*. 2002). This model identified the Intermountain West, Coastal Oregon, and Northern Rockies to be separate from each other and from all other populations.

Discussion

Sequencing a greatly expanded portion of the mitochondrial genome (*cf*. Aubry *et al*. 2009; Statham *et al*. 2012) has helped clarify the origins and population genetic structure of red foxes in western North America, particularly in the Intermountain West. I uncovered a pattern suggestive of natural origins and historical isolation for red foxes in this region. The two widely-distributed haplotypes in the Intermountain West (*w*6 and *w*7) formed a well-supported subclade within my selection of Mountain subclade red foxes (Fig. 2b). Further, these haplotypes were historically rare in montane red foxes and absent from modern populations in the Cascades, Rocky Mountains, and Sierra Nevada (Fig. 3a). These results are consistent with the key prediction of the Endemism hypothesis (see Introduction), namely, that one should expect a prevalence of distinct haplotypes derived from nearby mountain populations.

I found little support for Hall and Kelson's (1959) hypothesis that red foxes in southern Nevada belong to the Sierra Nevada subspecies *V. v. necator* (Fig. 1). Structure analyses based on A-19 variants consistently identified the Intermountain West

as a distinct group and significantly differentiated from all other western regions (Fig. 4; Table 1).

The results above suggest two plausible origin scenarios. First, red foxes in the Intermountain West may have persisted locally on “sky islands” of suitable habitat until recent times. Historically, the distribution of red foxes in western North America (Hall and Kelson 1959; Kamler and Ballard 2002) appears to have been similar to that of golden-mantled ground squirrels (*Callospermophilus lateralis*; Bartels and Thompson 1993), yellow-pine chipmunks (*Neotamias amoenus*; Sutton 1992), and other species associated with montane forests. At the height of the Wisconsin glaciation, when the treeline was up to 1,000 m lower than it is today, conditions favorable to red foxes would have been more broadly distributed in the Intermountain West (Thompson and Mead 1982). Increased habitat connectivity would have facilitated gene flow among populations in the Rocky Mountains, Cascades, and Sierra Nevada, allowing red foxes to colonize intervening mountain ranges in the Great Basin. Following the LGM, however, xeric shrublands gradually replaced forests at low elevation, and forest-associated fauna would have experienced range fragmentations and local extinctions (Grayson 1987). American martens (*Martes americana*), for example, are not known to occur in the Great Basin today, but fossil evidence indicates they had a widespread presence during the Pleistocene (Heaton 1985; Mead and Mead 1989). At least one forest-associated carnivore species (the short-tailed weasel, *Mustela erminea*) does still persist in this region (Brown 1971), and some habitat islands may have been large enough to support low numbers of red foxes. Reports collected by Bailey (1936) and Hall (1946) hint at the existence of one or more source populations in southern Nevada and/or southeastern Oregon—a scenario supported by my observations.

Second, red foxes may have recently colonized the Intermountain West via downslope colonization from the northern Rocky Mountains or southern Cascades. I note that several of my Intermountain West specimens came from mid-elevation areas near these mountain systems (Fig. 3b), raising the possibility that further sampling would reveal the presence of *w*6 and *w*7 haplotypes in modern mountain populations. Downslope colonization followed by range expansion is consistent with the appearance of red foxes in valleys in Idaho and Montana beginning in the 1960s (Fichter and Williams 1967; Hoffmann *et al*. 1969). However, the prevalence of two haplotypes in the Intermountain West that are rare in modern mountain populations suggests that any

such expansion would likely have stemmed from a small number of colonizing individuals rather than from a wholesale population expansion followed by continuous connectivity. Indeed, expansion from a small and rare colonization event seems more plausible than expansion of the mountain population as a whole in light of findings elsewhere in western North America. For example, Swanson *et al*. (2005) found high-elevation red foxes in the Yellowstone region to be distinct from neighboring valley populations on the basis of microsatellite data. The high level of genetic differentiation observed between Sierra Nevada and Sacramento Valley red foxes, even in historical times, also supports the hypothesis that downslope dispersal is rare (Sacks *et al*. 2010). Additional surveys of the Rocky Mountains and the Oregon Cascades would help further determine the likelihood of dispersal from these regions into the Intermountain West.

Several factors potentially allowed red foxes in the Intermountain West to expand their range in recent times. Coyotes (*Canis latrans*) are known to restrict the abundance of red foxes where the two species co-occur (Voigt and Earle 1983; Sargeant *et al*. 1987), particularly in areas where apex predators have been removed (Ritchie and Johnson 2009). Competitive exclusion could have therefore prevented red foxes from inhabiting lowland areas until the mid-1900s, when the reduction of coyote populations by farmers allowed red foxes to gain a foothold (Kamler and Ballard 2002). Habitat conversion is also likely to have played a role. East of the Rocky Mountains, this process has long been hypothesized to explain the southward and westward expansion of red foxes that began in the early 1800s (Audubon and Bachman 1846; Newberry 1857; Hoffmann *et al*. 1969; Kamler and Ballard 2002; Statham *et al*. 2012). In western North America, the transformation of dense forests and xeric valleys into irrigated cropland may have created new migration corridors and edge habitats for native red foxes to utilize.

Although my findings lend credibility to a natural origin for red foxes in the Intermountain West, the impacts of fur farm activity and local translocations should not be underestimated. Despite the overwhelming presence of native microsatellites and mitochondrial haplotypes in this region, a few indisputably non-native haplotypes were detected in previous studies (Sacks *et al*. 2010; Statham *et al*. 2012). Moreover, the *w*5 (O-24) haplotype observed here in the Salt Lake Valley is known to co-occur with red foxes carrying the Holarctic “G-38” haplotype (Appendix A; BN Sacks unpublished data). Both O-24 and G-38 have been associated with fur farms (Statham *et al*. 2012; BN

Sacks unpublished data), which were historically abundant on the margins of the Great Salt Lake (Westwood 1989). Thus, it is possible that interbreeding between native mountain and nonnative fur farm-derived red foxes could have facilitated expansion across the Intermountain West.

Based on available records, an alternative hypothesis is that indigenous red foxes spread northward from the mountains of southern Nevada (Hall 1946; Hall and Kelson 1959). Further sampling from this region using additional markers (*e.g.*, microsatellites) is needed to clarify the origins and native-nonnative ancestry of contemporary red foxes in the Intermountain West.

My results for Lowland Oregon were more equivocal. The only haplotype detected in this population, *w*3, was relatively common historically in both the Southern Cascades and Northern Rockies (Fig. 3b). Definitive conclusions are hampered by the genetic homogeneity of Lowland Oregon red foxes and the limited number of samples available from the Oregon Cascades. Historically, Cooper and Suckley (1859) reported red foxes to be abundant in the vicinity of The Dalles (Fig. 1), and Newberry (1857) reported them from the Klamath Lakes and Deschutes Basin. Bailey's (1936) distribution map included parts of the Oregon Coast Range to the west of the Willamette Valley, based on specimens collected there in the form of skins traded by Native Americans in 1897 and 1930. Thus, additional markers will be necessary to resolve the origins of this population.

The close and robust grouping of Sacramento Valley haplotype *w*4 with mountain haplotype *w*3 observed on the median-joining network and ML tree may indicate that this low-elevation population is of more recent origin than the Intermountain West subclade (Fig 2b).

Although mitochondrial DNA is relatively easy to recover from degraded samples, making it a natural first choice in population genetic surveys, I note that using a single haploid locus has two important shortcomings. First, because mitochondria are maternally inherited, I cannot rule out the possibility of introgression from nonnative red foxes. Second, the observation of divergent haplotypes does not necessarily imply deep population divergence, as these polymorphisms may have been present long before the splitting event. Multiple loci are often needed to distinguish ancestral polymorphism from modern gene flow and accurately measure the extent of population differentiation (Marko

and Hart 2011). Until the data collected in this study can be compared to those of nuclear markers, I recommend cautious interpretation of my results.

Conservation Implications

Red foxes in the Intermountain West appear to represent a unique, native population, although it remains unclear whether this population originated from a Pleistocene relict endemic to the Great Basin or from a more recent colonization event from the Southern Cascades or Rocky Mountains. The contribution of nonnative ancestry to this population also remains to be characterized. Given the observation of apparently stable contact zones between native and nonnative populations elsewhere in western North America (Sacks *et al*. 2011), fine-scale sampling is needed to clarify population structure within the Intermountain West.

Figures and Tables

Figure 1. Selected historical records of red foxes (*Vulpes vulpes*) in western North America showing where the species may have occurred prior to the establishment of fur farms. Darkened areas denote the approximate pre-Columbian range of native red foxes (Hall and Kelson 1959; Aubry 1983). Large diamonds correspond to the type locality for each subspecies in the Mountain subclade: *V. v. cascadensis* of the Washington Cascades, *V. v. necator* of the Oregon Cascades and Sierra Nevada, *V. v. macroura* of the Rocky Mountains, and *V. v. patwin* of the Sacramento Valley (in green). Nonnative populations occur in the Puget Sound region of Washington (A) (Statham *et al*. 2012) and in southern California (C) (Lewis *et al*. 1999). Red foxes of uncertain origin (in yellow) occur in lowland areas of western Oregon (B) and the Intermountain West (D), and are of interest in the present study. Readers should note that the records presented here do not strictly correspond to samples used in my analyses.

Sources: 1 = Armstrong (1972), 2 = Bailey (1931), 3 = Bailey (1936), 4 = Bailey and Bailey (1918), 5 = Baird (1852), 6 = Barnes (1922), 7 = Grinnell and Storer (1924), 8 = Grinnell et al. (1930), 9 = Grinnell et al. (1937), 10 = Hall (1946), 11 = Merriam (1900), 12 = Merriam and Stejnrger (1891), 13 = Norris (1881), 14 = Sacks et al. (2010), 15 = Taylor and Shaw (1927).

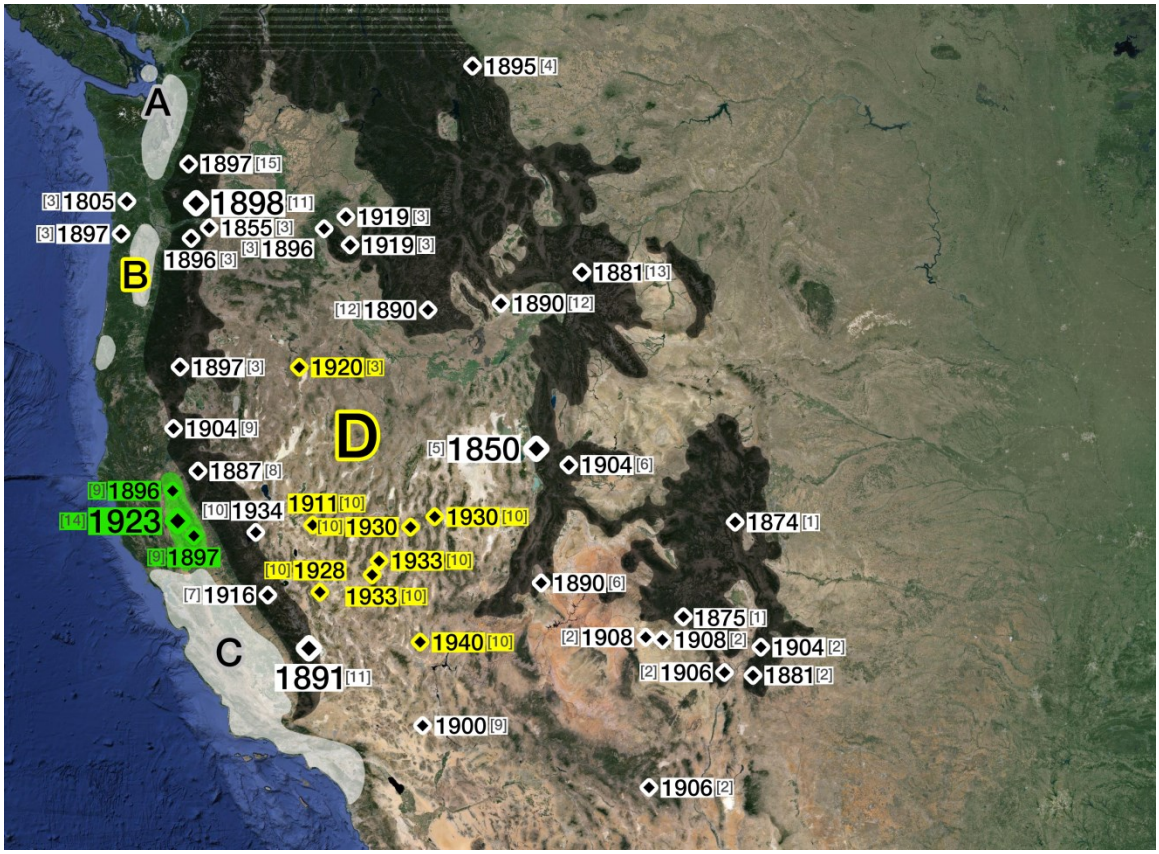

A
B
C
D
1895 [4]
1897 [15]
[3] 1805
1898 [11]
1855 [3]
1919 [3]
[3] 1897
[3] 1896
1919 [3]
1896 [3]
1881 [13]
[12] 1890
1890 [12]
1897 [3]
1920 [3]
1904 [9]
[5] 1850
1904 [6]
1887 [8]
[9] 1896
[10] 1934
1911 [10]
[14] 1923
[10] 1930
1930 [10]
1874 [1]
[9] 1897
[10] 1928
1933 [10]
1890 [6]
[7] 1916
1933 [10]
1875 [1]
[2] 1908
1908 [2]
1904 [2]
1940 [10]
1891 [11]
[2] 1906
1881 [2]
1900 [9]
1906 [2]

Figure 2. (A) Variable sites across 4,004 bp of cytochrome *b*, the control region, and cytochrome *c* oxidase subunit I for selected red fox haplotypes in western North America. Previous haplotype assignments come from Perrine *et al*. (2007) and Aubry *et al*. (2009). Loci are separated by dotted lines. Variable sites recovered from historical samples (targeted sequences) are highlighted. (B) Maximum-likelihood tree of full-locus haplotypes with outgroup specimens representing other major subclades in North America (see also Appendix A). Bootstrap values are shown beside each node.

A

Haplotype	Previous assignment	1 8 7	4 1 1	4 6 9	8 0 1	8 8 9	1 3 8	1 3 9	1 4 0	1 8 9	1 9 8	1 9 3	2 0 8	2 2 1	2 5 3	2 7 0	2 8 2	2 9 0	3 3 9	3 3 8	3 5 7	3 7 0	Δbp
w1	A-19	C	T	G	T	A	–	–	A	G	T	G	C	G	C	T	T	T	G	A	C	C	–
w2	A-19	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	G	·	·	2
w3	A-19	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	C	C	A	G	·	·	5
w4	D-19	·	C	A	·	·	·	·	·	·	·	·	·	·	·	·	·	C	A	G	·	·	5
w5	O-24	T	·	·	·	·	·	·	G	A	·	·	·	·	T	·	·	·	·	G	·	·	5
w6	A-19	·	·	·	C	G	·	·	·	·	A	A	T	·	T	·	·	·	·	G	T	T	9
w7	A-19	·	·	·	·	G	C	·	·	·	A	A	·	·	T	C	·	·	·	G	·	T	8
w8	A-19	?	·	·	·	G	C	C	?	·	A	?	·	·	?	?	?	?	?	?	?	?	≥4

B

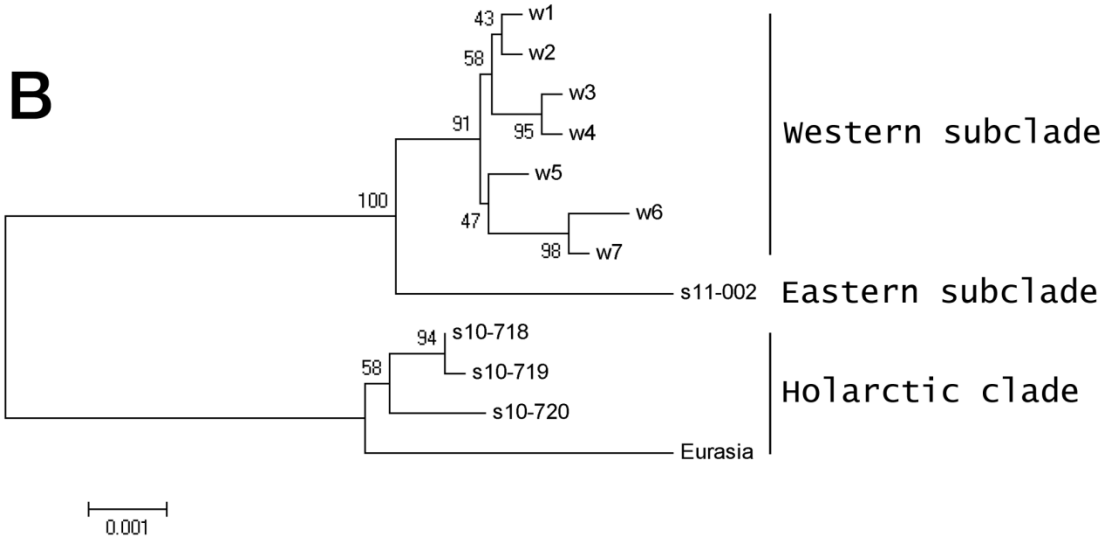


Figure 3. (A) Distribution of western red fox haplotypes identified in the present study. Polygons denote the approximate range of each red fox population of interest. Sample localities are colored according to the median-joining network in (B). Dark and light circles denote historical and modern samples, respectively (see also Appendices 2, 4). Enlarged circles denote locations sampled more than once. (B) Median-joining network of full-locus haplotypes. Circle size is proportional to the number of samples represented. The position of haplotype *w*8 is based on its close relationship to *w*7 in the targeted dataset.

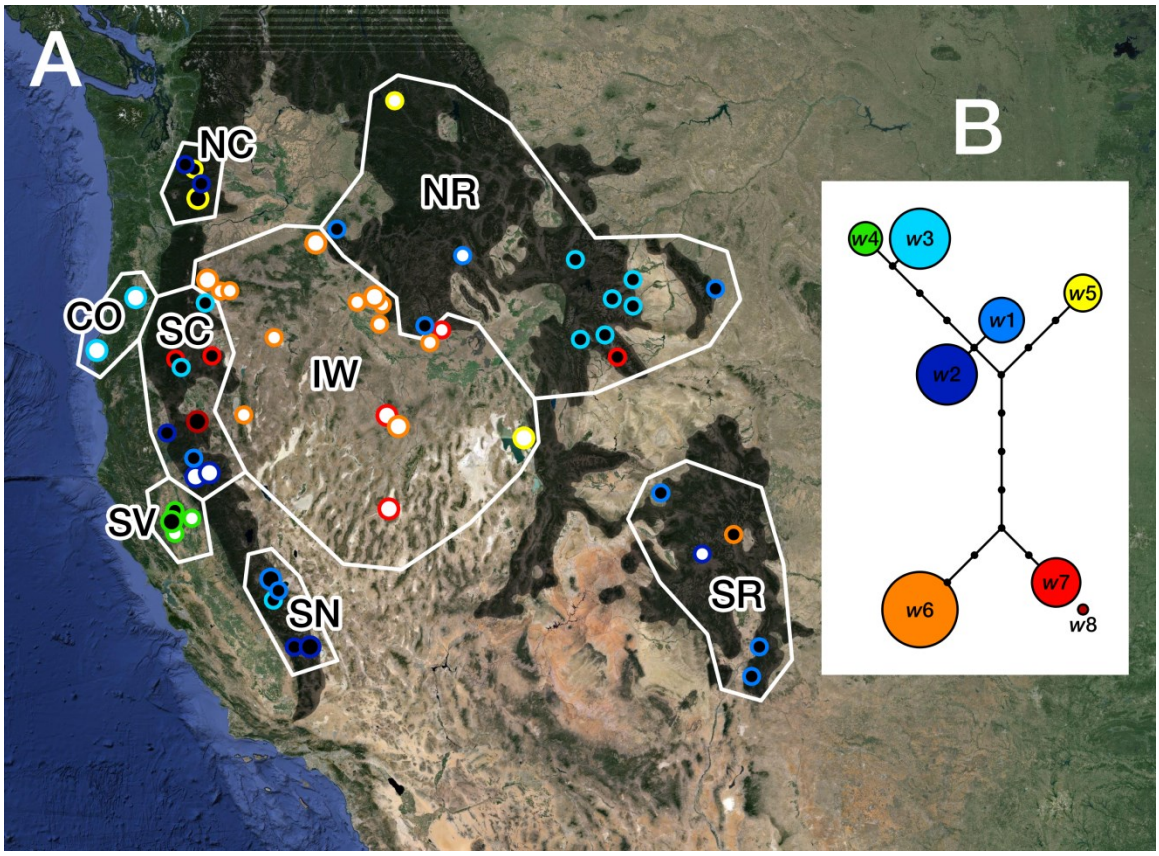
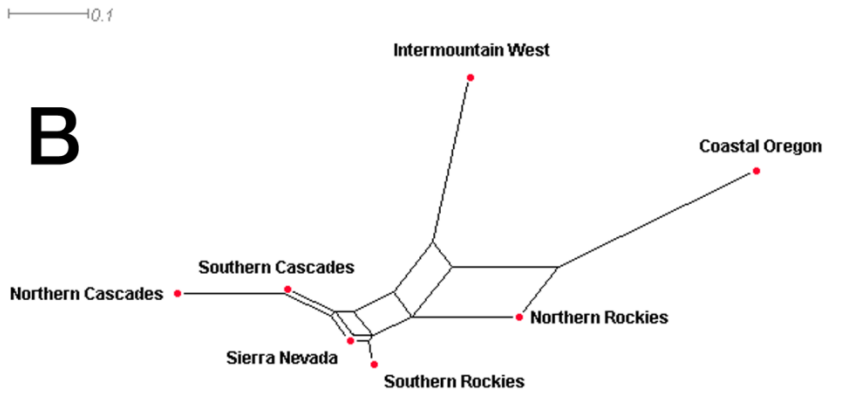


Figure 4. (A) Pairwise F_{ST} (below diagonal) and Φ_{ST} (above diagonal) estimates for 7 red fox populations in western North America based on A-19 variants. Bold entries are significant at $\alpha = 0.05$ after using a sequential Bonferroni correction for multiple tests (Rice 1989). (B) Phylogenetic networks of pairwise population genetic distances. Top: F_{ST}. Bottom: Φ_{ST}.

[a] NR = Northern Rockies, SR = Southern Rockies, NC = North Cascades, SC = South Cascades, SN = Sierra Nevada, IW = Intermountain West, CO = Coastal Oregon

A

	NR	SR	NC	SC	SN	IW	CO
NR[a]	—	0.182	0.267	0.187	0.131	**0.709**	0.204
SR	0.166	—	0.041	0.038	-0.013	**0.620**	**0.701**
NC	0.496	0.367	—	-0.126	-0.020	**0.758**	1
SC	**0.288**	0.219	-0.057	—	0.051	**0.589**	**0.512**
SN	0.136	-0.077	0.101	0.056	—	**0.747**	**0.714**
IW	**0.462**	**0.383**	**0.597**	**0.411**	**0.456**	—	**0.843**
CO	0.226	**0.681**	1	**0.519**	**0.557**	**0.656**	—



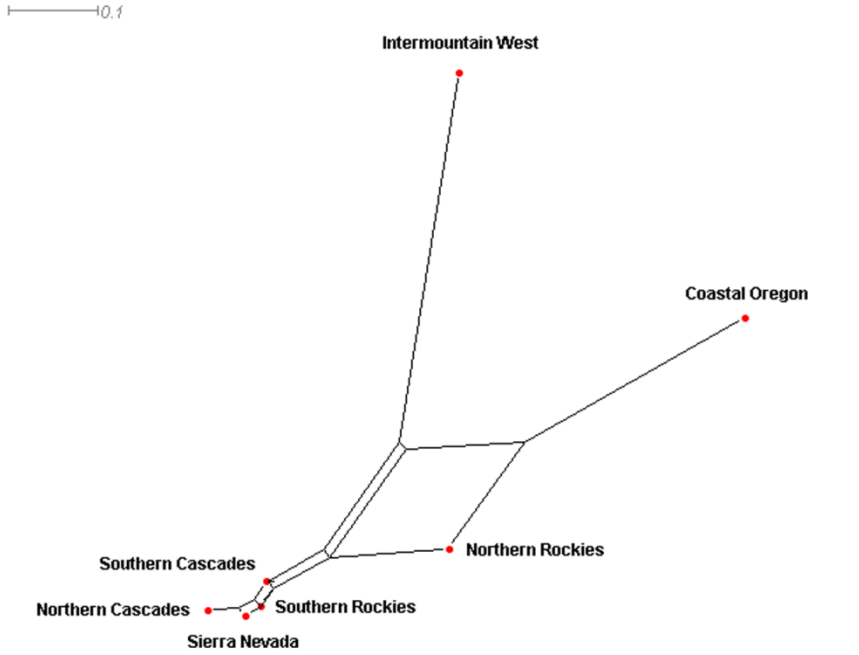


Table 1. Best population groupings found by SAMOVA under each *K* model. Letters indicate separate groups.

Population	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$
Intermountain West	A	A	A	A	A
Coastal Oregon	B	B	B	B	B
Northern Rockies	B	B	C	C	C
Northern Cascades	B	C	D	D	D
Southern Cascades	B	C	D	D	E
Southern Rockies	B	C	D	E	F
Sierra Nevada	B	C	D	E	F
ΦCT	0.541	0.584	0.593	0.602	0.592
P	0.152	0.009	0.059	0.008	0.107

Chapter 3.

Prioritizing populations for conservation using phylogenetic networks

Publication Note

The material in this chapter was previously published in *PLoS ONE*:

Volkmann L, Martyn I, Spillner A, Mooers AØ (2014). Prioritizing populations for conservation using phylogenetic networks. *PLoS ONE*, 9 (2): e88945.

Abstract

In the face of inevitable future losses to biodiversity, ranking species by conservation priority seems more than prudent. Setting conservation priorities within species (*i.e.*, at the population level) may be critical as species ranges become fragmented and connectivity declines. However, existing approaches to prioritization (*e.g.*, scoring organisms by their expected genetic contribution) are based on phylogenetic trees, which may be poor representations of differentiation below the species level. In this chapter I extend evolutionary isolation indices used in conservation planning from phylogenetic trees to phylogenetic networks. Such networks better represent population differentiation, and my extension allows populations to be ranked in order of their expected contribution to the set. I illustrate the approach using data from two imperiled species: the spotted owl *Strix occidentalis* in North America and the mountain pygmy-possum *Burramys parvus* in Australia. Using previously published mitochondrial and microsatellite data, respectively, I construct phylogenetic networks and score each population by its relative genetic distinctiveness. In both cases, my phylogenetic networks capture the geographic structure of each species: geographically peripheral populations harbor less-redundant genetic information, increasing their conservation rankings. I note that my approach can be used with all conservation-relevant distances (*e.g.*, those based on whole-genome, ecological, or adaptive variation) and suggest it be

added to the assortment of tools available to wildlife managers for allocating effort among threatened populations.

Introduction

Extinctions due to human impacts are now unavoidable: even optimistic scenarios predict significant changes in biodiversity by the year 2100 (Sala *et al*. 2000; Jenkins 2003), with most extinction starting with the loss of isolated populations (Hughes *et al*. 1997; Ceballos and Ehrlich 2002).

One prime conservation goal is to preserve genetic variation (Vane-Wright *et al*. 1991; Faith 1992): as a representation of past evolution and raw material for future evolution (Bowen and Roman 2005) and, potentially, as a surrogate for improved ecosystem function (Redding and Mooers 2010). However, not all genetic lineages are equally important, with more isolated lineages warranting additional interest because of their expected contribution to total variation (May 1990; Vane-Wright *et al*. 1991; Redding and Mooers 2010). Indices of evolutionary isolation have been developed to rank species on a phylogenetic tree based on unique and shared evolutionary history (*e.g.*, (Pavoine *et al*. 2005; Redding and Mooers 2006; Isaac *et al*. 2007; Haake *et al*. 2008). These metrics use rooted phylogenetic trees with edge lengths as input (Fig. 5), and rank tips with less shared history as requiring more urgent conservation attention. For example, the Zoological Society of London has made this approach operational in their “Edge of Existence” programme (www.edgeofexistence.org). In the United States, taxonomic distinctiveness is one of several explicit criteria for prioritizing conservation attention (Fay and Thomas 1983). The extension to populations within species would seem to be straightforward.

Below the species level, Ryder (1986) advocated the use of *evolutionarily significant units* (ESUs) to identify populations with genetic variation in need of long-term conservation; this was expanded by Moritz (1994) and Waples (1991) to the concepts of *management units* (MUs) and *distinct population segments* (DPSs), respectively, for species that had undergone more recent range fragmentation. More recently, Green (2005) defined *designatable units* (DUs) for conservation assessment using a broad range of survey methods. All these population-based approaches have enjoyed wide usage in population genetic studies (*e.g.*, Firestone *et al*. 1999; Roman *et al*. 1999; Gorbics and Bodkin 2001; Haig *et al*. 2001), and are the basis for identifying populations

worthy of protection in law (*e.g.*, Green 2005; Holtby and Ciruna 2007). Importantly, ESUs assume that the relationships among populations can be represented by a bifurcating tree. However, bifurcating trees often fail to capture the relationships among populations (Fitch 1997). DPSs and MUs can deal with populations that have more complex interrelationships (Fig. 5), but neither of these designations are designed to prioritize among populations. This would seem a major shortcoming if populations do need to be prioritized for conservation attention.

Previous authors have shown that the logic of measuring and maximizing phylogenetic diversity (Faith 1992), which forms the basis for tree-based prioritization schemes, can be generalized to phylogenetic networks (Weitzman 1992; Minh *et al*. 2009a, b). Here I show that the prioritization approaches for trees can also be adapted for populations within species by extending evolutionary isolation indices from trees to networks. I present efficient algorithms to compute these indices for NeighborNet networks (Bryant and Moulton 2004; Martyn *et al*. 2012), and illustrate their use with heuristic data from two imperiled species, the spotted owl (*Strix occidentalis* Xantus de Vesey 1860) and the mountain pygmy-possum (*Burramys parvus* Broom 1896). The new approach to assessing population differentiation might be of immediate practical use to those tasked with managing discrete populations of a threatened species, and may allow for new policy associated with conservation triage (Bottrill *et al*. 2008).

Methods

I present my approach for prioritizing populations in three steps. First, I briefly review the various approaches for measuring diversity and evolutionary isolation on bifurcating trees of taxa. I then review the properties of NeighborNet networks as a representation of pairwise evolutionary distances and describe how to prioritize taxa by their expected contribution to biodiversity. In Appendix F, I outline efficient algorithms for estimating evolutionary isolation on NeighborNet networks. Finally, I illustrate the new method of population prioritization using two small published datasets.

(i) Diversity measures on trees and networks. — The concept of evolutionary isolation can be understood in terms of a species' biological distinctiveness, which one might measure by comparing its adaptive or non-adaptive traits to those of related species. More generally, the goal is to measure a taxon's contribution to the current and/or future

"diversity" in a set of taxa. Several different approaches for quantifying such diversity have been proposed. One of the earliest, described by Weitzman (1992), is *expected diversity*. Rather than score taxa individually, this approach seeks to identify the set of taxa that will retain the most diversity on a future tree, given some measure of diversity and a probability of persistence for each potential combination of taxa. Although Weitzman's original diversity metric was rather general, he did consider an example of biological character-state differences that could be represented on a phylogenetic tree.

On such a tree, every taxon contributes an amount of unique evolutionary information denoted by the length of the branch (or *edge*) linking it to all other taxa (Fig. 5) (Faith 1992; Weitzman 1992). This length may be calibrated in units of time (*e.g.*, millions of years) or in raw or inferred genetic distances. Looking specifically at biological systems, Witting and Loeschcke (1995) and Faith and Walker (1996) combined Weitzman's (1992) expected diversity framework with Faith's (1992) concept of *phylogenetic diversity* (PD), the latter which specifically calculates the sum of all branch lengths on a tree (see next section). Like Weitzman (1992), this *expected PD* approach can be used to identify a set of taxa that maximizes the amount of total tree length retained, given a set of extinction probabilities for the tips.

The related *k of n problem* (Faith 1992) seeks to identify the most diverse subset of *k* taxa (*i.e.*, the one that maximizes PD) on a tree of size n. Faith (1996) and Weitzman (1998) explored the special case where $k = 1$, which Faith (2008) refers to as the *PD complementarity* of a given taxon.

An independently-derived approach based on Game Theory (Haake *et al*. 2008, first published 2005) explicitly considers the individual contribution of each taxon to future diversity. Like Weitzman's (1992) expected diversity framework, all possible subsets of taxa on a tree may persist. By calculating the amount of unique information each taxon contributes to future subsets (*i.e.*, the average length of the edge linking the taxon to all possible future trees), one can rank taxa in order of their relative impact on future diversity. This *Shapley metric* (SH) is almost identical to the ad-hoc *evolutionary distinctness* (ED) metric used by the Zoological Society of London in their Edge of Existence programme (www.edgeofexistence.org). The major difference between the two is that the ED metric is explicitly measured on a rooted tree, as opposed to the more general undirected graph that SH takes as input (Hartmann 2013).

The Shapley metric was further refined by Steel *et al*. (2007) and named HED (for *heightened evolutionary distinctiveness*). HED is the expected contribution of a given taxon to future subsets of taxa where the subsets are weighted by their probability of persistence. In this case, the focal taxon is assumed to persist (*i.e.*, its probability of extinction does not affect its HED score). On trees, HED is formally equivalent to a form of PD complementarity where the contribution of a taxon is measured with respect to all possible subsets, each weighted by their probability of persistence (Faith 2008). Weitzman (1998) also arrived at this formulation ten years earlier, which he termed the "distinctiveness" of a taxon, in the context of his "Noah's Ark Problem" of biodiversity preservation. Using Faith's (2008) terminology, HED, which combines the concepts of expected PD with PD complementarity, might be considered *expected PD complementarity*.

As a final antecedent, Minh *et al*. (2007, 2009a, b) extended PD to phylogenetic networks and presented algorithms for solving the *k* of *n* problem to maximize diversity for a given subset size. They referred to this metric as *split diversity* (SD).

In this context it should be possible to measure the PD contribution of individual taxa on a phylogenetic network. Critically for my purposes, the two metrics I use here (SH and HED) do not require a rooted phylogenetic tree, and so can be adapted to networks in the same way that PD indices can (Minh *et al*. 2007, 2009a, b; Martyn *et al*. 2012). SH and HED are formally defined in Appendix F and discussed further below. In short, if one does not have probabilities of extinction for taxa, they assume all future subsets of taxa are equally likely, and calculate SH. If one can estimate (even broadly) the probabilities of persistence of all taxa, they can weight future subsets by their probability, and use HED.

(ii) Interpreting phylogenetic networks, Shapley values, and HED. — NeighborNet (Bryant and Moulton 2004) is a method that permits the representation of pairwise distances between taxa in the form of a network. An important property of NeighborNet networks is that they permit the representation of relationships among the underlying taxa that cannot be depicted on any phylogenetic tree. For example, to the extent that populations exchange migrants, the between-population genetic distance data (F_{ST}) may yield many alternative trees, none of which accurately reflect the actual relationships among these populations (*e.g.*, Poczai *et al*. 2011). The NeighborNet framework, by

contrast, accommodates for such phylogenetic uncertainty and will always yield a single network with positive edge lengths, permitting calculation of SH and HED. If a pairwise distance matrix is tree-like (*i.e.*, yields only one possible phylogeny) the resulting NeighborNet output will resemble a phylogenetic tree. Where there is no tree-like history, a network representation should be more informative. Indeed, for many distance matrices (including Example A below, results not shown), the assumptions necessary to produce a tree are not met, and a neighbor-joining tree, for example, produces negative edge lengths. Here, a network representation would definitely be preferred (Bryant and Moulton 2004).

An example of a very simple matrix of pairwise distances and the resulting network is depicted in Figure 5. Each edge or set of parallel edges in the network corresponds to a partition of the underlying set of taxa into two non-overlapping subsets, called a *split* (S). The *edge length* reflects the *weight* of the split ($\lambda(S)$)—in other words, a component of the pairwise distance (F_{ST}, for example) separating any two taxa. Thus, just as a phylogenetic tree represents a collection of weighted splits (Σ) (Buneman 1971), where each branch of the tree denotes a split, a NeighborNet network represents a weighted collection of splits of the underlying set of taxa. As Figure 5 illustrates, the distance between two tips on a network (*i.e.*, the shortest path between two taxa) represents the observed distance in the distance matrix.

Whether represented on a tree or a network, every split system contains information on the overall diversity of its constituent taxa (Vane-Wright *et al*. 1991; Weitzman 1992). The common metric used for conservation planning, *phylogenetic diversity* (PD) (Faith 1992), can be calculated for split systems as

$$PD_{(\Sigma,\lambda)}(Y) = \sum_{\substack{A|B\in\Sigma \\ A\cap Y\neq 0, B\cap Y\neq 0}} \lambda(A \mid B)$$

where Y is a subset of taxa on the tree or network and $\lambda(A \mid B)$ is the weight of the split between two non-overlapping groups A and B of taxa. Note that the overall PD for both trees (Faith 1992; Steel 2005) and networks (Minh *et al*. 2007, 2009a, b) is simply the sum of all split weights (Fig. 5).

A very simple approach for measuring an individual taxon's PD contribution, illustrated in Figure 5, is to consider the change in PD when this taxon is removed from the tree or network (Faith 1994). This *PD complementarity* (PD_c) metric can be expressed as

$$PD\_c(x) = PD_{(\Sigma,\lambda)}(X) - PD_{(\Sigma,\lambda)}(\overline{X})$$

where X is the set of all taxa in the tree or network and $\overline{X}$ is the subset where a given taxon x has been removed from the underlying distance matrix.

One can also extend the metrics SH and HED from trees to NeighborNet networks using similar ideas for extending PD calculations from trees (*e.g.*, Faith 1992, Witting and Loeschcke 1995; Martyn *et al*. 2012) to networks (*e.g.*, Huson and Bryant 2006; Minh *et al*. 2007, 2009a, b). On a tree, the Shapley value (ψ_x^{sh}) for taxon x can be defined as the mean split weight of the set of splits defining $x|\overline{X}$, where $\overline{X}$ represents all unique possible subsets of the taxon set X that do not contain x. Importantly, Haake *et al*. (2008) present a formal proof that the Shapley value for x can also be calculated as a weighted sum of all the edge lengths on a tree, with the weights determined by the sizes of the sets containing x. This can be presented compactly using split notation as

$$\psi_x^{sh}(\Sigma,\lambda) = \sum_{S\in\Sigma} \frac{|\overline{S}(x)|}{|X||S(x)|}\lambda(S)$$

where (Σ,λ) is the set of splits defined by the network and their weights, $|X|$ is the total number of taxa, $|S(x)|$ is the size of a split set containing the taxon x, $|\overline{S}(x)|$ is the size of the complementary set that does not contain x, and $\lambda(S)$ (following the notation from Minh *et al*. (2009a, b) is the split weight, equal to the edge length separating $S(x)$ from $\overline{S}(x)$. To calculate the Shapley value for taxon $x1$ in the network in Figure 1, I take the first split $x2|x1,x3,x4,x5,x6$ to be composed of $|S(x1)|=5$ and $|\overline{S}(x1)|=1$ and $\lambda(S)=0.373$, the second split $x3|x1,x2,x4,x5,x6$ to be composed of $|S(x1)|=5$ and $|\overline{S}(x1)|=1$ and $\lambda(S)=0.111$ and so on. With a taxon set containing six elements, $|X|=6$ and the Shapley value for taxon $x1$ is 0.870 (Fig. 5).

As with a phylogenetic tree, the sum of Shapley values will always equal the sum of all parallel split weights in the network. Because the shape of a network reflects the relative distances among its taxa, one should expect outlying taxa (*i.e.*, those connected to the rest of the network by long edges, like taxon $x4$) to show higher values for ψ_x^{sh}. Thus, the Shapley values calculated for a network can reflect the relative degree of

isolation of each taxon based on molecular, morphological, or any other evolutionarily-relevant distance measure.

Though conceptually similar, the calculation of HED (ψ_x^{hed}) is somewhat more complex, as it accounts for differences in the probability of extinction $p(y)$ for each taxon:

$$\psi_x^{hed}(\Sigma,\lambda) = \sum_{S\in\Sigma}\Big(\prod_{y\in(S(s)-\{x\})} p(y)\Big)\cdot\Big(1-\prod_{y\in\bar{S}(x)} p(y)\Big)\cdot\lambda(S)$$

Here, the first product operator considers $p(y)$ for every taxon in $S(x)$ but excludes $p(y)$ for taxon x itself (Steel *et al*. 2007; Martyn *et al*. 2012). The second product operator considers $p(y)$ for every taxon in $\bar{S}(x)$. Unlike SH, the sum of HED scores will not equal the sum of split weights in the split system. I also note, based on the properties of the equation above, that $\lambda(S)$ will influence HED more strongly than $p(y)$ for highly divergent taxa. Thus, the ranking order for outlying populations should be similar for SH and HED, regardless of which populations have a higher extinction probability.

A more detailed mathematical treatment of the SH and HED metrics and efficient algorithms for their computation are given in Appendix F. For the datasets in this chapter, I used the implementation of NeighborNet in the SplitsTree software package (Huson and Bryant 2006) to compute networks. For a given matrix of pairwise distances, this yields the network together with the corresponding collection of weighted splits. I used custom R scripts (Appendix G; R-Development Core Team 2009) to compute SH and HED on the outputs from SplitsTree.

(iii) Application. – I present SH and HED ranking for two datasets based on putatively neutral genetic markers. In the first example (*A*), the size of each population (and hence the probability of extinction for each population) is not known. In the second example (*B*), population sizes are known, allowing me to estimate separate probabilities of extinction for each population.

I selected my two examples based on the following criteria: (1) The species as a whole is of conservation interest (*i.e.*, vulnerable, endangered, or critically endangered), (2) its distribution is fragmented (*i.e.*, one can define multiple populations), (3) sampling efforts have covered its entire range, and (4) genetic analyses have been published or the raw sequence data made publicly available.

Readers should note that the primary goals of this article are to introduce and illustrate my network ranking approach, not to advocate new management decisions for the taxa described below.

Example A. Spotted owls (*Strix occidentalis*) are distributed throughout late-succession conifer forests in western North America (Gutiérrez *et al*. 1995). Four subspecies are currently recognized (Fig. 6a): *S. o. caurina* from southern British Columbia to northwest California, *S. o. occidentalis* in California and Nevada, *S. o. lucida* in Utah, Colorado, Arizona, New Mexico, and northern Mexico, and *S. o. juanaphillipsae* in central Mexico (Dickerman 1997; Barrowclough *et al*. 1999). Populations in the United States continue to decline due largely to poor timber harvesting practices, but also as a result of climate change and the westward expansion of barred owls (*S. varia* Barton 1799) (US Fish and Wildlife Service 2011). *S. o. caurina* (the northern spotted owl) and *S. o. lucida* (the Mexican spotted owl) are threatened subspecies under the United States' Endangered Species Act, and *S. o. occidentalis* (the California spotted owl) is a subspecies of special concern in the state of California (Gould 1985). Spotted owls in the American Southwest "sky islands" (mostly *S. o. lucida*) are particularly fragmented and perhaps most suitable for population-level conservation (Barrowclough *et al*. 2006). Although genetic data for the Mexican subspecies remain poor, I can construct a reasonably complete representative phylogenetic network for subspecies in the United States.

Spotted owl mitochondrial sequences were obtained from Genbank (accession numbers AY833608–AY833644, AY836774–AY836776, DQ230843–DQ230888) and aligned in Mega v. 5 (Tamura *et al*. 2011) using the MUSCLE method (Edgar 2004). These sequences comprise about 1105 bp of the control (D-loop) region and represent 86 haplotypes from 32 populations in the United States and Mexico (Fig. 6b; Table 2) (Barrowclough *et al*. 2005, 2006). I ran a standard analysis of molecular variance (AMOVA) (Excoffier *et al*. 1992) on all 298 aligned sequences in Arlequin v. 3.5 (Excoffier *et al*. 2005) using the Kimura 2-Parameter model (Kimura 1980) to compute distances among haplotypes (Φ_{ST}). This procedure generated a pairwise differentiation matrix for the 32 populations (Table S1). A NeighborNet based on this matrix (Fig. 6b) (Bryant and Moulton 2004) was then constructed in SplitsTree v. 4.11 (Huson and Bryant 2006) under default assumptions. Negative Φ_{ST} values were treated as being equal to zero. Because the size of each population is not known, for the purposes of

illustration, I gave each population an extinction probability $p_i = 0.5$ when calculating HED—an approach similar to the "PD50" metric used by FISHBASE (www.fishbase.org; Faith *et al*. 2004).

Example B. Mountain pygmy-possums (*Burramys parvus*) are alpine specialists restricted to three small regions of the Australian Alps (Fig. 7a). The species depends on block streams and block fields found above 1,400 meters—habitats less than 10 km^2 in total extent (Heinze *et al*. 2004). The areas where mountain pygmy-possums still occur are particularly sensitive to destruction and fragmentation. Surveys conducted in the 1990s estimated the adult population size to be 2,600 (Mansergh and Broome 1994). A decade later this number had decreased to below 2,000 (Heinze *et al*. 2004), with signs of continued decline (Mitrovski *et al*. 2007). At present, the IUCN lists mountain pygmy-possums as critically endangered (Menkhorst *et al*. 2008).

Because of its restricted distribution and high extinction risk, the species has been subject to extensive population genetic research (Osborne *et al*. 2000; Mitrovski *et al*. 2005, 2007, 2008). Unlike my example with spotted owls, direct estimates of population sizes are available, within-population sample sizes are uniformly large, and genetic data are available across the mountain pygmy-possums' entire range. This provides us with an opportunity to compare SH to HED and assess the effect of variable population sizes on conservation ranking.

I used a published matrix of genetic differentiation (F_{ST}) based on data from 8 microsatellite loci (Mitrovski *et al*. 2007) to construct a phylogenetic network for 13 mountain pygmy-possum populations (Fig. 7b). My methods for generating NeighborNet outputs, and for computing SH and HED, were the same as above.

I modeled the probabilities of extinction for individual populations (p_i) of a given size (n_i) as a negative exponential

$$p_i \propto e^{-cn_i}$$

where the constant of proportionality c is $-\ln(P)/N$, with P being the probability that the entire species goes extinct and N being the total census size of the species (the sum of n_i). I used a conservative 100-year extinction probability for the entire species, $P = 0.4$, to derive HED (see Mooers *et al*. 2008).

Results

Example A. As expected for a set of lineages with a recent history of gene flow, the network for spotted owls is quite non tree-like (Fig. 6b). However, populations with the greatest degrees of genetic differentiation, relative to all other populations, occupy nodes subtending the longest edges. Populations at relatively isolated nodes, such as those from Mount San Jacinto and the Huachuca Mountains, share few mutations with neighboring populations and subsequently exhibit higher pairwise Φ_{ST} values (Table 2; Appendix H). Conversely, the (uncorrected) pairwise Φ_{ST} values for closely-related populations are either negative (as great as -1) or close to zero, indicating higher levels of genetic differentiation within these populations than among them (Excoffier *et al*. 1992).

I observe strong geographic structure across the United States consistent with current subspecific designations (Fig. 6b). Populations of *S. o. lucida* exhibit a more star-like phylogenetic network that may reflect historical isolation in the "sky islands" of the American Southwest (Barrowclough *et al*. 2006). The intermediate position of the Lassen National Forest population, in contrast, may be due to its location near the point of contact between southern *S. o. caurina* and northern *S. o. occidentalis* (Barrowclough *et al*. 2005).

The results of my SH and HED ranking are shown in Table 2. As expected, populations at relatively isolated nodes score higher than those closer to the interior of the network (Fig. 6b, c). The rankings are highly consistent between the two metrics (Spearman rank correlation = 0.91), and the same populations receive top ranking for both SH and HED.

Example B. As with spotted owls, the most genetically differentiated populations of mountain pygmy-possums, namely those in the northern and southern areas of their range, occupy nodes that are separated from most other populations by long edges (Fig. 7b). Overall the structure of my network is in good agreement with the species' present distribution (Fig. 7a). Given the habitat requirements and limited dispersal ability of mountain pygmy-possums, it is not likely that Mount Buller and Kosciusko National Park still exchange migrants with the Bogong High Plains (Mitrovski *et al*. 2007). In contrast,

the close grouping of central populations in my phylogenetic network, and subsequently their low SH and HED, is consistent with a shared history and/or recent gene flow.

The ranking results are shown in Table 3. Again, the phylogenetic network for mountain pygmy-possums reflects geographic distribution. Although I did not make *a priori* group assignments based on sampling location, the 13 populations still partition into northern, central, and southern regions. Again, outlying populations on the network tend to receive higher SH and HED scores. Unsurprisingly, the small and isolated Mount Buller population consistently ranks highest. For HED, no bias towards small or large populations is apparent; populations with high extinction probabilities do not necessarily receive high scores (Steel *et al*. 2007). Again, although ranking order changes slightly between SH and HED, the two methods provide roughly equivalent rankings (Spearman rank correlation = 0.97, Fig. 7c). High-ranking populations are similar in both cases.

I note that SH and HED calculations on a network consider a taxon's distance from *all other* taxa. Thus, although the three northern populations are closely related to each other, they still receive high SH and HED scores because of the long branches separating them from the central and southern populations (Table 3; Fig. 7c).

Discussion

The premise of conservation below the species level is not novel. Faith's original (1992) discussion of prioritizing taxa also considered populations on a tree. Several economically-important taxa have received population-level management since the late 1980s, *e.g.*, Atlantic salmon (*Salmo salar* Linnaeus 1758; Fontaine *et al*. 1997), brown trout (*Salmo trutta* Linnaeus 1758; Ferguson 1989) and yellowfin tuna (*Thunnus albacares* Bonnaterre 1788; Ward 1995). Managing species at the population level implies at least an informal ranking scheme, one which would rely, for example, on estimates of habitat patch size or effective population size (Lande and Barrowclough 1987). Habitat degradation, climate change, and the demands of a growing human population have ensured the continued fragmentation of species' ranges over the next century (see, *e.g.*, pikas (*Ochotona princeps* Richardson 1828; Galbreath *et al*. 2009). In the midst of such rapid change, managing an imperiled species over its entire range may no longer be feasible, such that population rankings may be necessary.

Phylogenetic diversity measures have previously been adapted for non-treelike population genetic data (*e.g.*, Minh *et al*. 2007, 2009a, b). However, the PD

complementarity scores that can be obtained from these methods are contingent, *i.e.*, subject to change if extinction alters the shape of the network. The ranking scheme described here is the first to consider a taxon's contribution to all possible future networks (*sensu* Weitzman 1992), a potentially relevant framework for preserving future biodiversity. Given the stochastic nature of extinction, the general ranking systems offered by SH and HED may be more useful to wildlife managers than those that only consider the present structure of a phylogenetic network. Unlike previous approaches based on PD (*e.g.*, Minh *et al*. 2007, 2009a, b), SH and HED rankings allow one to lengthen or shorten the list of taxa to conserve in the event that resources become more or less available.

Molecular techniques are now inexpensive and robust enough to make population genetic sampling a standard component of conservation planning, and I argue that a phylogenetic network approach offers insight into a species' population structure complementary to the current statistical assessments of differentiation employed by MUs and DPSs (Waples 1991; Moritz 1994). I encourage researchers to employ such networks in future population genetic studies to provide conservation agencies with more informative analysis of datasets. Genotyping at multiple loci will provide more accurate estimates of population differentiation and allow for more sophisticated analyses of conservation-relevant processes such as recent demographic history and gene flow (Marko and Hart 2011).

I acknowledge that the mathematical shortcomings of Φ_{ST} and F_{ST} estimators (Jost 2008) may influence the magnitude and ranking of SH and HED scores, depending on the number of loci measured and the distribution of genetic diversity in a set of taxa. My intention here is not to solve these theoretical problems but to demonstrate my network-based prioritization method with existing data. Newly-developed metrics such as Jost's *D* can be used to calculate SH and HED just as readily as traditional Φ_{ST} and F_{ST} distances, and I encourage the use of such unbiased estimators whenever such data are available. Indeed, any conservation metric of difference (*e.g.*, ecological, genomic, adaptability) can be used.

Several properties of the networks described here invite further investigation. In both my heuristic datasets, geographically peripheral populations are more genetically isolated, meaning they would rank highly on SH and HED. However, this was based on only very few putatively neutral markers. Two related questions concern how processes

such as demographic history and current patterns of gene flow map onto genetic isolation as I measure it here, and also how phylogenetic networks map onto networks produced from ecological data (*e.g.*, niche use differences among populations).

I do not advocate relying solely on genetic isolation when deciding where and how resources should be allocated at the population level. Total population genetic diversity (*i.e.*, number of haplotypes) might also be considered. I note that in my examples, low-ranking populations tend to be geographically close to one another, meaning that their haplotypes are often shared. Important differences in ecology and adaptability (Bonin *et al*. 2007) and current and future connectivity (Geffen *et al*. 2007), must also be considered. However, my network approach and ranking system based on genetic differentiation can supplement existing systems of MUs and DPSs to improve the conservation of evolutionarily distinct populations in a world of increasing pressures and limited resources.

Figures and Tables

Figure 5. Using pairwise distances to rank species or populations. Consider a hypothetical group of taxa (A)—a set of closely-related species or populations of a single species—that is distributed across several islands in an archipelago (B). Differences among the taxa, labeled *x1* through *x6*, can be organized into a pairwise distance matrix (C). One can represent this matrix either as a phylogenetic tree or as a phylogenetic network (D), where a set of weighted *splits* describes the relationships among the taxa (E). Altogether, these splits represent the group's phylogenetic diversity (PD). By selecting subsets of splits that exclude a given taxon, one can calculate each taxon's contribution to the total PD of the tree or network (F). The *Shapley metric* (SH) and *expected PD complementarity* (PD_c) are different approaches for ranking taxa based on split data. Note that the highest-scoring taxa (highlighted values) can differ considerably depending on the type of metric used and whether the splits come from a tree or network. I discuss the reasons for these differences and methods for ranking taxa in Section (*ii*) of the main text.

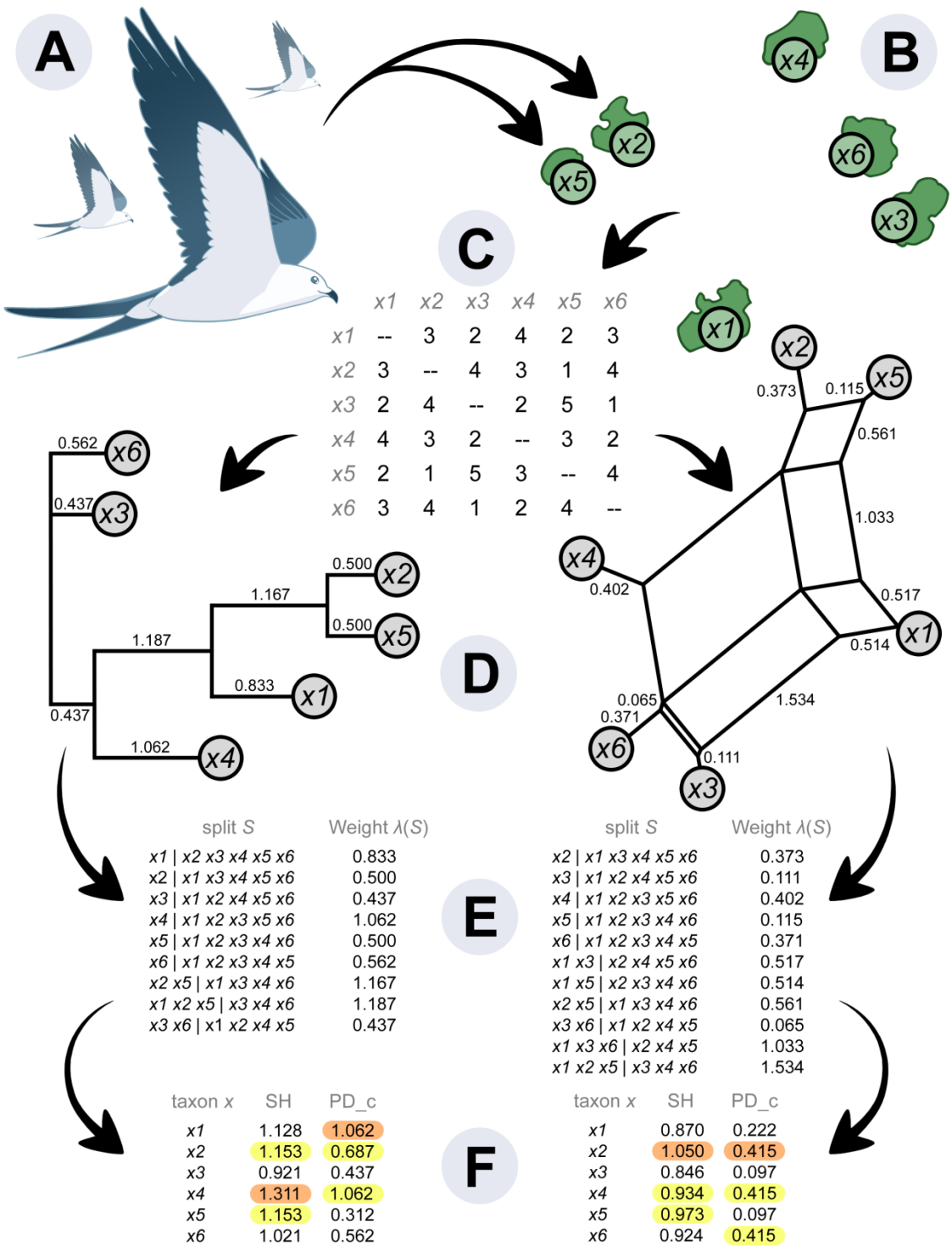

A
B
C
D
E
F
x1 x2 x3 x4 x5 x6
x1 -- 3 2 4 2 3
x2 3 -- 4 3 1 4
x3 2 4 -- 2 5 1
x4 4 3 2 -- 3 2
x5 2 1 5 3 -- 4
x6 3 4 1 2 4 --
0.562
0.437
0.500
1.167
0.500
1.187
0.833
0.437
1.062
0.373
0.115
0.561
1.033
0.402
0.517
0.514
0.065
0.371
1.534
0.111
split S
Weight λ(S)
x1 | x2 x3 x4 x5 x6 0.833
x2 | x1 x3 x4 x5 x6 0.500
x3 | x1 x2 x4 x5 x6 0.437
x4 | x1 x2 x3 x5 x6 1.062
x5 | x1 x2 x3 x4 x6 0.500
x6 | x1 x2 x3 x4 x5 0.562
x2 x5 | x1 x3 x4 x6 1.167
x1 x2 x5 | x3 x4 x6 1.187
x3 x6 | x1 x2 x4 x5 0.437
x2 | x1 x3 x4 x5 x6 0.373
x3 | x1 x2 x4 x5 x6 0.111
x4 | x1 x2 x3 x5 x6 0.402
x5 | x1 x2 x3 x4 x6 0.115
x6 | x1 x2 x3 x4 x5 0.371
x1 x3 | x2 x4 x5 x6 0.517
x1 x5 | x2 x3 x4 x6 0.514
x2 x5 | x1 x3 x4 x6 0.561
x3 x6 | x1 x2 x4 x5 0.065
x1 x3 x6 | x2 x4 x5 1.033
x1 x2 x5 | x3 x4 x6 1.534
taxon x
SH
PD_c
x1 1.128 1.062
x2 1.153 0.687
x3 0.921 0.437
x4 1.311 1.062
x5 1.153 0.312
x6 1.021 0.562
x1 0.870 0.222
x2 1.050 0.415
x3 0.846 0.097
x4 0.934 0.415
x5 0.973 0.097
x6 0.924 0.415

Figure 6. Conservation prioritization of spotted owl (*Strix occidentalis*) populations. (a) Distribution of spotted owls in the United States and the populations sampled by Barrowclough *et al.* (2005, 2006). Shaded areas denote suitable habitat based on forest cover data (US Geological Service 2013). Colors denote the subspecies *S. o. caurina* (blue), *S. o. occidentalis* (green), and *S. o. lucida* (orange). Populations 31 and 32 represent the *S. o. juanaphillipsae* subspecies in Mexico (range not shown). (b) NeighborNet of sampled populations based on mtDNA differentiation (pairwise Φ_{ST} values). (c) Histogram of SH values, highlighting the populations with the highest scores. See Table 2 for an explanation of abbreviations used.

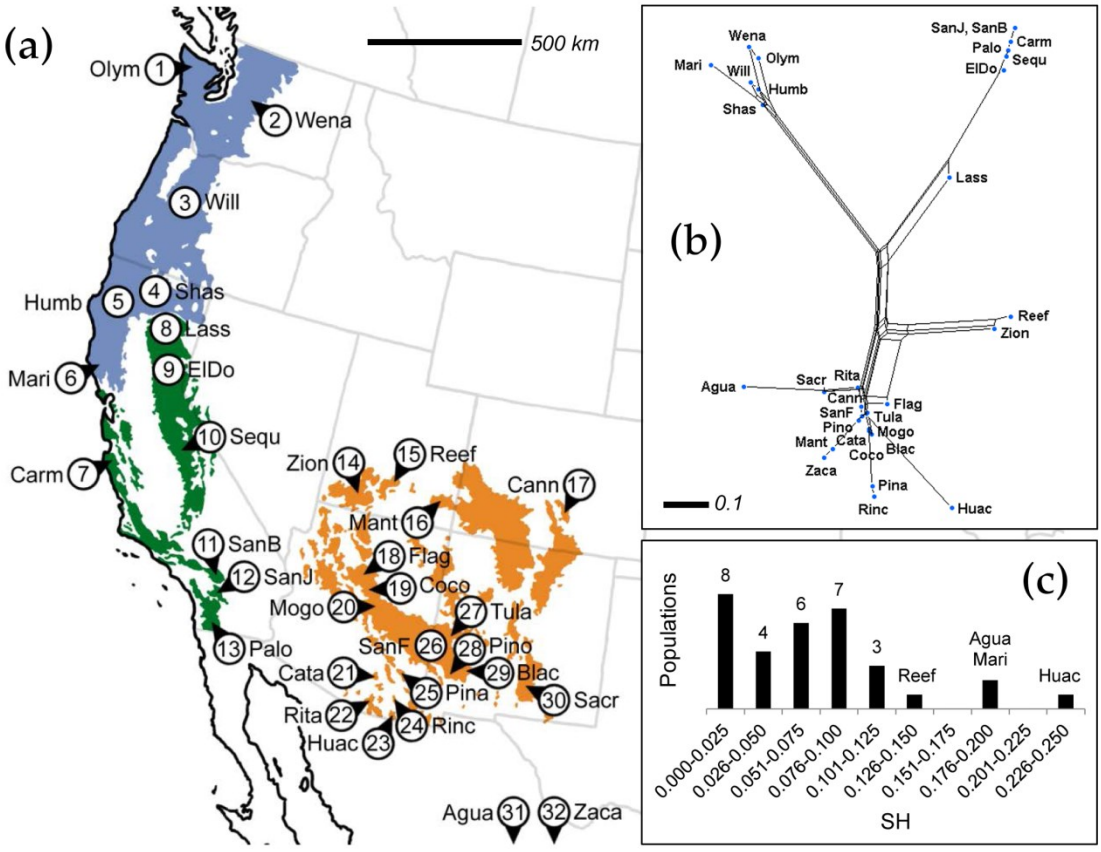


Figure 7. Conservation prioritization of mountain pygmy-possum (*Burramys parvus*) populations. (a) Distribution of mountain pygmy-possums in Australia (gray inset), showing populations sampled by Mitrovski *et al.* (2007). Shaded areas denote suitable habitat above 1,400 m. (b) NeighborNet of sampled populations based on microsatellite differentiation (pairwise F_{ST} values). (c) Histograms of SH and HED values, highlighting the populations with the highest scores. See Table 3 for an explanation of abbreviations used.

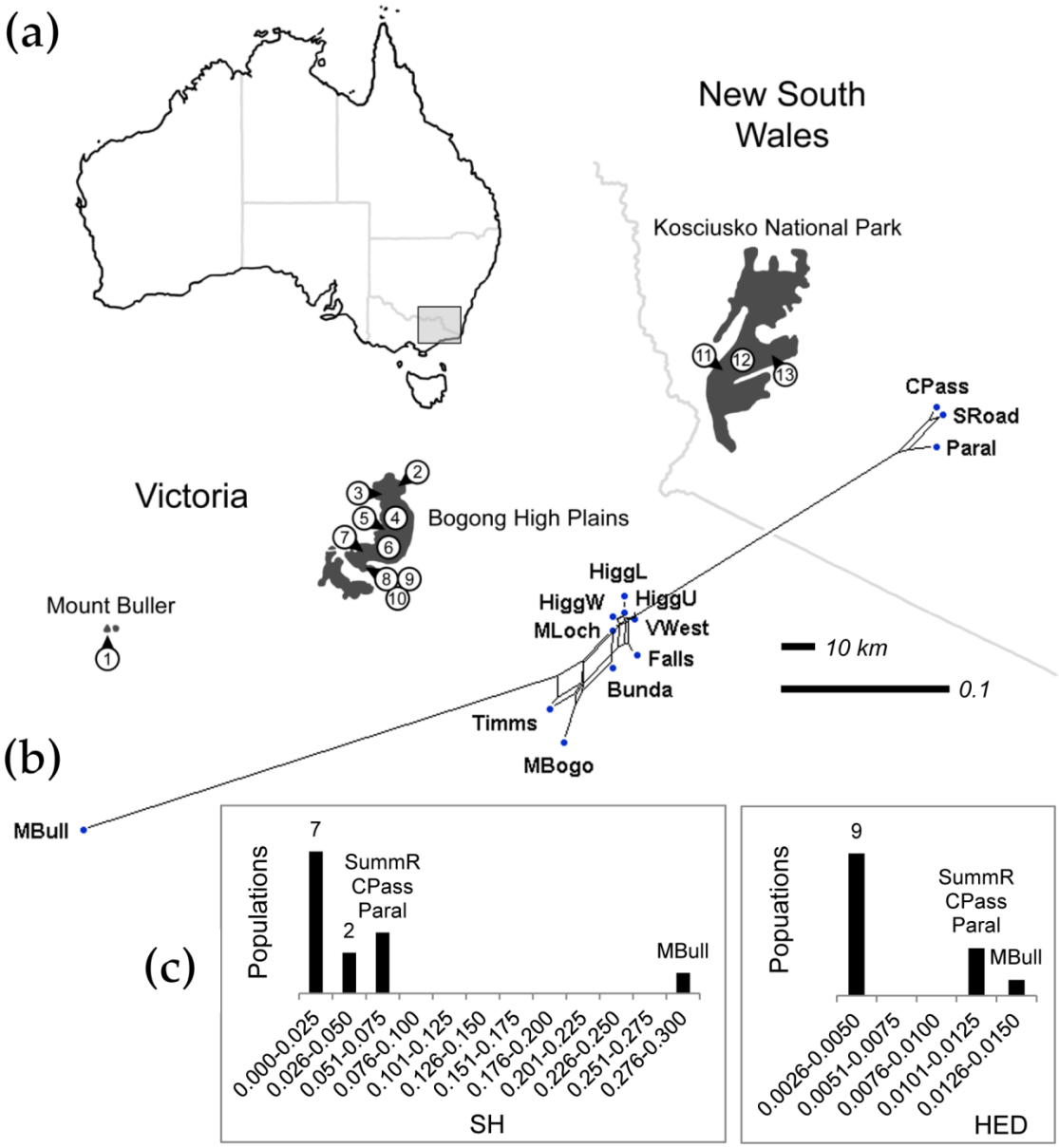


Table 2. Spotted owl populations sampled by Barrowclough *et al*. (2005, 2006) and ranked by Shapley value (SH) and heightened evolutionary distinctiveness (HED). Number of individuals (*n* ind.), number of haplotypes (*n* hap.), SH, and HED scores from the present study are reported.

Pop.	Code	Subspecies	State	Sampling Locality	*n* ind.	*n* hap.	SH	HED
23	Huac	*lucida*	AZ	Huachuca Mountains	5	2	0.242	7.43E-03
31	Agua	*juanaphillipsae*	—	Aguascalientes, Sierra Fria, Mexico	1	1	0.191	3.98E-03
6	Mari	*caurina*	CA	Marin County	8	3	0.177	4.07E-03
15	Reef	*lucida*	UT	Capitol Reef National Park	9	4	0.133	3.68E-03
2	Wena	*caurina*	WA	Cascade Range	10	8	0.111	1.78E-03
1	Olym	*caurina*	WA	Olympic Peninsula	10	4	0.106	1.73E-03
14	Zion	*lucida*	UT	Zion National Park	7	4	0.105	2.85E-03
11	SanB	*occidentalis*	CA	San Bernardino Mountains	15	1	0.091	9.42E-04
12	SanJ	*occidentalis*	CA	Mount San Jacinto	15	1	0.091	9.42E-04
24	Rinc	*lucida*	AZ	Rincon Mountains	8	4	0.089	2.54E-03
3	Will	*caurina*	OR	Willamette National Forest	15	8	0.081	7.62E-04
7	Carm	*occidentalis*	CA	Carmel Valley	10	1	0.079	5.19E-04
9	ElDo	*occidentalis*	CA	El Dorado National Forest	15	4	0.077	6.21E-04
32	Zaca	*juanaphillipsae*	—	Zacatecas, Sierra de Urica, Mexico	1	1	0.076	2.08E-03
13	Palo	*occidentalis*	CA	Mount Palomar	8	1	0.074	3.99E-04
5	Humb	*caurina*	CA	Humboldt and Siskiyou Counties	30	11	0.072	4.53E-04
10	Sequ	*occidentalis*	CA	Sierra National Forest	15	6	0.071	3.42E-04
25	Pina	*lucida*	AZ	Pinaleno Mountains Graham County	4	2	0.071	1.61E-03
4	Shas	*caurina*	CA	Klamath and Shasta National Forests	16	8	0.067	3.64E-04
16	Mant	*lucida*	UT	Manti-La Sal National Forest	2	2	0.054	1.28E-03
30	Sacr	*lucida*	NM	Sacramento Mountains	8	6	0.047	1.28E-03
8	Lass	*occidentalis*	CA	Lassen National Forest	11	6	0.041	1.01E-04
18	Flag	*lucida*	AZ	San Fransisco Peaks	4	4	0.027	2.14E-04
29	Blac	*lucida*	NM	Black Range	8	6	0.026	3.46E-04
19	Coco	*lucida*	AZ	Coconino Plateau	15	9	0.022	1.04E-04
21	Cata	*lucida*	AZ	Santa Catalina Mountains	5	3	0.021	6.36E-05
28	Pino	*lucida*	NM	Pinos Altos Mountains	5	4	0.02	2.83E-05
26	SanF	*lucida*	NM	San Fransisco Mountains	7	4	0.017	1.02E-06
17	Cann	*lucida*	CO	Near Cañon City	4	4	0.017	8.51E-07
22	Rita	*lucida*	AZ	Santa Rita Mountains	4	4	0.017	1.37E-05
20	Mogo	*lucida*	AZ	Mogollon Mesa	8	6	0.017	8.58E-07
27	Tula	*lucida*	NM	Tularosa Mountains	15	12	0.017	2.03E-06

Table 3. Mountain pygmy-possum populations sampled by Mitrovski *et al*. (2007) and ranked by Shapley value (SH) and heightened evolutionary distinctiveness (HED). Number of individuals (*n* ind.), number of alleles (*n* all.), allelic richness (*r*), and adult population sizes (*N*) are reported from previously-published data. Probabilities of extinction (*p*i, with *P* = 0.4), SH, and HED scores from the present study are also shown.

Pop.	Code	Region	Sampling Locality	*n* ind.	*n* all.	*r*	*N*	*p* i	SH	HED
1	MBull	Southern	Mount Buller	66	3.38	2.29	150	0.9072	0.292	1.50E-02
12	CPass	Northern	Charlottes Pass	44	6	5.21	45	0.9712	0.072	1.15E-02
13	Paral	Northern	Paralyser	40	6.63	5.77	22	0.9858	0.071	1.09E-02
11	SummR	Northern	Summit Road	43	6.13	5.16	25	0.9839	0.068	1.14E-02
2	MBogo	Central	Mount Bogong	42	6.5	5.77	100	0.9372	0.045	4.50E-03
3	Timms	Central	Timm Spur	120	6.88	5.32	120	0.9251	0.028	3.86E-03
4	Falls	Central	Falls Creek	35	6.63	5.73	30	0.9807	0.02	3.05E-03
8	HiggL	Central	Mount Higginbotham L	17	5.25	5.25	50	0.9681	0.019	2.85E-03
6	Bunda	Central	Bundara	78	7.25	5.45	120	0.9251	0.018	2.86E-03
10	HiggW	Central	Mount Higginbotham W	59	7.63	6.01	250	0.8502	0.013	2.66E-03
5	VWest	Central	Pretty Valley West	69	7	5.95	50	0.9681	0.01	2.54E-03
9	HiggU	Central	Mount Higginbotham U	56	7.25	5.76	50	0.9681	0.01	2.42E-03
7	Mloch	Central	Mount Loch	93	7	5.8	400	0.7714	0.01	2.65E-03

Chapter 4.

Conclusions

New analytical methods and the falling cost of genetic testing should help make population-level conservation an effective means of managing individual threatened species. However, several important challenges remain. Most obvious is the need for more systematic surveys to identify key populations. In many cases, (*e.g.*, in studies by Kyle and Strobeck 2001, 2002; Cegelski *et al*. 2006), population genetic data come from inconsistent sampling schemes or cover only a portion of the target species' range, leaving managers with an incomplete understanding of conservation needs. Now that methods exist to prioritize populations for conservation, future work should seek to characterize variation across a species' entire range rather than only in peripheral areas. Regional comparative studies might also aid population-level conservation efforts. The identification of common phylogeographic patterns in western North America, for example, would help focus conservation resources to protect key populations of multiple species (Soltis 1997; Brunsfeld 2001).

Perhaps the most important outstanding challenges to population-level conservation involve how to measure variation within a species and translate those data into conservation policy. First is the problem of *what* to sample. To date, the overwhelming majority of population genetic studies have focused on putatively neutral genetic markers such as mitochondrial loci and microsatellites, which are relatively easy to obtain and analyze. Although these markers can, for example, inform us about gene flow among populations and help us understand how species have responded to past climate change, neutral variation tells us very little about the evolutionary potential of populations we wish to conserve (Holderegger *et al*. 2006). This is problematic if we wish to preserve biodiversity in terms of adaptive variation, a typical goal of conservation at the species level and above (Bowen and Roman 2005).

Second is the problem of *how* to sample. Unlike species, populations do not necessarily exhibit discreet geographic ranges or phenotypic differences, making it more difficult to define meaningful conservation units and assess the success of new policy. Protected areas do not always coincide with areas of high within-species diversity. In addition, some key populations may be impossible to save due to climate change or other factors. How, then, should we designate conservation units within a species? What spatial scales are most appropriate when conducting population-level surveys?

Solving these problems is beyond the scope of the present studies, but we can consider two frameworks for conducting future surveys and determining conservation units below the species level. A *range-based* approach would define populations based on features of the species' geographic distribution. Sampling would focus on habitat patches of sustainable size, regardless of their current level of protection. Such an approach may be suitable for species with highly fragmented ranges and/or strict ecological requirements. The American pika (*Ochotona princeps*), for example, occupies numerous "sky islands" of alpine habitat, and dispersal between mountain systems is limited (Hafner and Sullivan 1995). Ecological niche modeling under climate-change scenarios (see, *e.g.*, Galbreath *et al*. 2009) would help refine sampling schemes and conservation priorities once phylogenetic data are collected.

Alternatively, a *refuge-based* approach would involve defining populations based on preexisting management areas. Sampling would focus on protected habitat patches as representative of the species' future geographic range. Such an approach may be suitable for species with more continuous ranges and/or generalized ecological requirements. The range of the Cascade golden-mantled ground squirrel (*Spermophilus saturatus*), for example, broadly coincides with national forests and national parks in Washington State (Trombulak 1988). Refuge-based prioritization of this species would allow managers to efficiently distribute conservation resources without the need to designate new protected areas.

Regardless of the approach taken, researchers should also be mindful to select sampling schemes and loci that measure adaptive variation (*e.g.*, Henry and Russello 2013) and that complement (at least to some degree) the datasets of previous population-level studies for their target species (*e.g.*, Mitrovski *et al*. 2008). The former will provide policymakers with more accurate picture of the species' conservation needs, and the latter will help track changes in population-level diversity over time.

In summary, the present studies have sought to advance two important aspects of conservation below the species level: (1) identification of key populations and (2) prioritization of these populations in a way that best preserves biodiversity. My genetic analyses of red foxes in the Intermountain West (Chapter 2) help fill an important gap in our understanding of how native and nonnative populations of this species are distributed on the landscape. The phylogenetic network-based methods illustrated in Chapter 3 should be immediately applicable to any species for which population genetic data are available. Future work, aimed at characterizing variation across a species' entire range, should be designed to help us collectively carry as much biodiversity as possible into the 22nd century and beyond.

References

Arbogast BS, Kenagy GJ (2001). Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, 28 (7): 819–825.

Armstrong DM (1972). *Distribution of Mammals in Colorado*. University of Kansas Printing Service, Lawrence, KS, USA. 415 pages.

Aubry KB (1983). The Cascade red fox: distribution, morphology, zoogeography and ecology. PhD dissertation, University of Washington, Seattle, WA, USA. 151 pages.

Aubry KB (1984). The recent history and present distribution of the red fox in Washington. *Northwest Science*, 58 (1): 69–79.

Aubry KB, Statham MJ, Sacks BN, Perrine JD, Wisely SM (2009). Phylogeography of the North American red fox: vicariance in Pleistocene forest refugia. *Molecular Ecology*, 18 (12): 2668–2686.

Audubon JJ, Bachman J (1846). *The Viviparous Quadrupeds of North America* (*Vol. 1)*. Privately published by JJ Audubon, New York, NY, USA. 389 pages.

Bailey V (1931). *Mammals of New Mexico*. North American Fauna No. 53, USDA Bureau of Biological Survey, Washington, DC, USA. 412 pages.

Bailey V (1936). *The Mammals and Life Zones of Oregon*. North American Fauna No. 55, USDA Bureau of Biological Survey, Washington, DC, USA. 416 pages.

Bailey V, Bailey FM (1918). *Wild Animals of Glacier National Park*. Government Printing Office, Washington, DC, USA. 210 pages.

Baird SF (1852). Mammals. Pages 309–313 in Stansbury H (editor): *Exploration and Survey of the Valley of the Great Salt Lake of Utah*. Lippincott Grambo & Co., Philadelphia, PA, USA. 487 pages.

Bandelt HJ, Forster P and Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16 (1): 37–48.

Barnes CT (1922). *Mammals of Utah*. Island Printing Company, Kaysville, UT, USA. 166 pages.

Barrowclough GF, Gutiérrez RJ, Groth JG (1999). Phylogeography of spotted owl (*Strix occidentalis*) populations based on mitochondrial DNA sequences: gene flow, genetic structure, and a novel biogeographic pattern. *Evolution*, 53 (3): 919–931.

Barrowclough GF, Groth JG, Mertz LA, Gutiérrez RJ (2005). Genetic structure, introgression, and a narrow hybrid zone between northern and California spotted owls (*Strix occidentalis*). *Molecular Ecology*, 14 (4): 1109–1120.

Barrowclough GF, Groth JG, Mertz LA, Gutiérrez RJ (2006). Genetic structure of Mexican spotted owl (*Strix occidentalis lucida*) populations in a fragmented landscape. *The Auk*, 123 (4): 1090–1102.

Bartels MA, Thompson DP (1993). *Spermophilus lateralis*. *Mammalian Species*, 440: 1–8.

Benson JF, Perrine JD, Golightly RT, Barrett RH (2005). Use of cover and response to cover type edges by female Sierra Nevada red foxes in winter. *Western North American Naturalist*, 65 (1): 127–130.

Bonin A, Nicole F, Pompanon F, Miaud C, Taberlet P (2007). Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology*, 21 (3): 697–708.

Bottrill MC, Joseph LN, Carwardine J, Bode M, Cook C, *et al*. (2008). Is conservation triage just smart decision making? *Trends in Ecology and Evolution*, 23 (12): 649–654.

Bowen BW, Roman J (2005). Gaia's handmaidens: the Orlog model for conservation biology. *Conservation Biology*, 19 (4): 1037–1043.

Brown JH (1971). Mammals on mountaintops: nonequilibrium insular biogeography. *The American Naturalist*, 105 (945): 467–478.

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001). Comparative phylogeography of northwestern North America: a synthesis. Pages 319–339 in Silvertown J, Antonovics J (editors): *Integrating Ecological and Evolutionary Processes in a Spatial Context*. Blackwell Science, Oxford, UK. 423 pages.

Bryant D, Moulton V (2004). Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, 21 (2): 255–265.

Buneman P (1971). The recovery of trees from measures of dissimilarity. Pages 387–395 in Hodson FR, Kendall DG, Tautu P (editors): *Mathematics in the Archaeological and Historical Sciences*. Edinburgh University Press, Edinburgh, UK. 565 pages.

California Department of Fish and Game (1999). Managing non-native species in California: the red fox. Sacramento, California, USA. 8 pages.

Ceballos G, Ehrlich PR (2002). Mammal population losses and the extinction crisis. *Science*, 296 (5569): 904–907.

Cegelski CC, Waits LP, Anderson NJ, Flagstad O, Strobeck C, *et al*. (2006). Genetic diversity and population structure of wolverine (*Gulo gulo*) populations at the southern edge of their current distribution in North America with implications for genetic viability. *Conservation Genetics*, 7 (2): 197–211.

Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, *et al*. (2009). The Last Glacial Maximum. *Science*, 325 (5941): 710–714.

Cook JA, Bidlack AL, Conroy CJ, Demboski JR, Fleming MA, *et al*. (2001). A phylogenetic perspective on endemism in the Alexander Archipelago of southeast Alaska. *Biological Conservation*, 97 (2): 215–227.

Cook JA, Dawson NG, MacDonald SO (2006). Conservation of highly fragmented systems: the north temperate Alexander Archipelago. *Biological Conservation*, 133 (1): 1–15.

Cooper JG, Suckley G (1859). *The Natural History of Washington Territory*. Baillière Brothers, New York, NY, USA. 399 pages.

Dickerman RW (1997). Geographic variation in southwestern US and Mexican spotted owls, with the description of a new subspecies. Pages 45–48 in Dickerman RW (editor): *The Era of Allen R. Phillips: a Festschrift*. Horizon Communications, Albuquerque, NM, USA. 246 pages.

Dudgeon D, Arthington AH, Gessner MO, Kawabata ZI, Knowler DJ, *et al*. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, 81 (2): 163–182.

Dupanloup I, Schneider S, Excoffier L (2002). A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, 11 (12): 2571–2581.

Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32 (5): 1792–1797.

Excoffier L, Lischer HE (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10 (3): 564–567.

Excoffier L, Laval G, Schneider S (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1: 47–50.

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

Faith DP (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61 (1): 1–10.

Faith DP (1994). Phylogenetic pattern and the quantification of organismal biodiversity. *Philosophical Transactions of the Royal Society B*, 345 (1311): 45–58.

Faith DP (1996). Conservation priorities and phylogenetic pattern. *Conservation Biology*, 10 (4): 1286–1289.

Faith DP (2008). Threatened species and the potential loss of phylogenetic diversity: conservation scenarios based on estimated extinction probabilities and phylogenetic risk analysis. *Conservation Biology*, 22 (6): 1461–1470.

Faith DP, Walker PA (1996). DIVERSITY – TD. Pages 63–74 in Faith DP, Nicholls AO (editors): *BioRap: Rapid Assessment of Biodiversity. Volume 3: Tools for Assessing Biodiversity Priority Areas*. The Australian BioRap Consortium, Canberra, Australia. 121 pages.

Faith DP, Reid CAM, Hunter J (2004). Integrating phylogenetic diversity, complementarity, and endemism for conservation assessment. *Conservation Biology*, 18 (1): 255–261.

Fay JJ, Thomas WL (1983). Endangered species listing and recovery priority guidelines. *Federal Register*, 48 (184): 43098–43105.

Ferguson A (1989). Genetic differences among brown trout, *Salmo trutta*, stocks and their importance for the conservation and management of the species. *Freshwater Biology*, 21 (1): 35–46.

Fichter E and Williams R (1967). Distribution and status of the red fox in Idaho. *Journal of Mammalogy*, 48 (2): 219–230.

Firestone KB, Elphinstone MS, Sherwin WB, Houlden BA (1999). Phylogeographical population structure of tiger quolls *Dasyurus maculatus* (Dasyuridae: Marsupialia), an endangered carnivorous marsupial. *Molecular Ecology*, 8 (10): 1613–1625.

Fitch WM (1997). Networks and viral evolution. *Journal of Molecular Evolution*, 44 (1): S65–S75.

Floyd CH, Van Vuren DH, May B (2005). Marmots on Great Basin mountaintops: using genetics to test a biogeographic paradigm. *Ecology*, 86 (8): 2145–2153.

Fontaine PM, Dodson JJ, Bernatchez L, Slettan A (1997). A genetic test of metapopulation structure in Atlantic salmon (*Salmo salar*) using microsatellites. Canadian Journal of Fisheries and Aquatic Sciences, 54 (10): 2434–2442.

Galbreath KE, Hafner DJ, Zamudio KR (2009). When cold is better: climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American pika, *Ochotona princeps*). *Evolution*, 63 (11): 2848–2863.

Geffen E, Waidyaratne S, Dalén L, Angerbjörn A, Vila C, *et al*. (2007). Sea ice occurrence predicts genetic isolation in the artic fox. *Molecular Ecology*, 16 (20): 4241–4255.

Gorbics CS, Bodkin JL (2001). Stock structure of sea otters (*Enhydra lutris kenyoni*) in Alaska. *Marine Mammal Science*, 17 (3): 632–647.

Gordon K (1966). Mammals and the influence of the Columbia River Gorge on their distribution. *Northwest Science*, 40 (4): 142–146.

Gould GI (1978). Status of the red fox in California. Progress Report, Job I-8, California Department of Fish and Game, Sacramento, CA, USA. 3 pages.

Gould GI (1985). Management of spotted owls by the California Department of Fish and Game. Pages 21–26 in Gutiérrez RJ, Carey AB (editors): *Ecology and Management of the Spotted Owl in the Pacific Northwest*. USDA Forest Service, Portland, OR, USA. 119 pages.

Grayson DK (1987). The biogeographic history of small mammals in the Great Basin: observations on the last 20,000 years. *Journal of Mammalogy*, 68 (2): 359–375.

Green DM (2005). Designatable units for status assessment of endangered species. *Conservation Biology*, 19 (6): 1813–1820.

Grinnell J, Storer TI (1924). *Animal Life in the Yosemite*. University of California Press, Berkeley, CA, USA. 752 pages.

Grinnell J, Dixon J, Linsdale JM (1930). *Vertebrate Natural History of a Section of Northern California through the Lassen Peak Region*. University of California Press, Berkeley, CA, USA. 594 pages.

Grinnell J, Dixon J, Linsdale J (1937). *Fur-bearing Mammals of California*. University of California Press, Berkeley, CA, USA. 777 pages.

Gutiérrez RJ, Franklin AB, Lahaye WS (1995). Spotted owl (*Strix occidentalis*). In Poole A (editor): *The Birds of North America Online*. Cornell Lab of Ornithology, Ithaca, NY, USA. Available: http://bna.birds.cornell.edu/bna/species/179. Accessed 28 November 2011.

Haake CJ, Kashiwada A, Su FE (2008). The Shapley value of phylogenetic trees. *Journal of Mathematical Biology*, 56 (4): 479–497.

Hafner DJ, Sullivan RM (1995). Historical and ecological biogeography of Nearctic pikas (Lagomorpha: Ochotonidae). *Journal of Mammalogy*, 76 (2): 302–321.

Haig SM, Wagner RS, Forsman ED, Mullins TD (2001). Geographic variation and genetic structure in spotted owls. *Conservation Genetics*, 2 (1): 25–40.

Hall ER (1946). *Mammals of Nevada*. University of California Press, Berkeley, CA, USA. 710 pages.

Hall ER, Kelson KR (1959). *The Mammals of North America*. The Ronald Press Company, New York, NY, USA. 1083 pages.

Hampe A, Petit RJ (2005). Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, 8 (5), 461–467.

Hartmann K (2013). The equivalence of two phylogenetic biodiversity measures: the Shapley value and the Fair Proportion index. *Journal of Mathematical Biology*, 67 (5): 1163–1170.

Hasegawa M, Kishino H, Yano T (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22 (2): 160–174.

Heaton TH (1985). Quaternary paleontology and paleoecology of Crystal Ball Cave, Millard County, Utah: with emphasis on mammals and description of a new species of fossil skunk. *The Great Basin Naturalist*, 45 (3): 337–390.

Heinze D, Broome LS, Mansergh I (2004). A review of the ecology and conservation of the mountain pygmy-possum *Burramys parvus* (Broom 1896). Pages 254–267 in Goldingay R, Jackson S (editors): *The Biology of Australian Possums and Gliders*. Surrey Beatty and Sons, Chipping Norton, Australia. 574 pages.

Henry P, Russello MA (2013). Adaptive divergence along environmental gradients in a climate-change-sensitive mammal. *Ecology and Evolution*, 3 (11): 3906–3917.

Hoffmann RS, Wright PL, Newby FE (1969). The distribution of some mammals in Montana I. Mammals other than bats. *Journal of Mammalogy*, 50 (3): 579–604.

Holderegger R, Kamm U, Gugerli F (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, 21 (6): 797–807.

Holtby LB, Ciruna KA (2007). Conservation units for Pacific salmon under the Wild Salmon Policy. Research Document 2007/070, Canadian Science Advisory Secretariat, Ottowa, ON, Canada. 350 pages.

Hughes JB, Daily GC, Ehrlich PR (1997). Population diversity: its extent and extinction. *Science*, 278 (5338): 689–692.

Huson DH, Bryant D (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23 (2): 254–267.

Isaac NBJ, Turvey ST, Collen B, Waterman C, Baillie JEM (2007). Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS One*, 2: e296.

Jenkins M (2003). Prospects for biodiversity. *Science*, 302 (5648): 1175–1177.

Jost L (2008). GST and its relatives do not measure genetic differentiation. *Molecular Ecology*, 17 (18): 4015–4026.

Jurek RM (1992). Nonnative red foxes in California. Nongame Bird and Mammal Report 92-04, California Department of Fish and Game, Sacramento, CA, USA. 16 pages.

Kamler JF, Ballard WB (2002). A review of native and nonnative red foxes in North America. *Wildlife Society Bulletin*, 30 (2): 370–379.

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16 (2): 111–120.

Kurtén B (1968). *Pleistocene Mammals of Europe*. Aldine Publishing Company, Chicago, IL, USA. 317 pages.

Kutschera VE, Lecomte N, Janke A, Selva N, Sokolov AA, *et al*. (2013). A range-wide synthesis and timeline for phylogeographic events in the red fox (*Vulpes vulpes*). *BMC Evolutionary Biology*, 13: 114.

Kyle CJ, Strobeck C (2001). Genetic structure of North American wolverine (*Gulo gulo*) populations. *Molecular Ecology*, 10 (2): 337–347.

Kyle CJ, Strobeck C (2002). Connectivity of peripheral and core populations of North American wolverines. *Journal of Mammalogy*, 83 (4): 1141–1150.

La Sorte FA, Jetz W (2010). Projected range contractions of montane biodiversity under global warming. *Proceedings of the Royal Society B*, 277 (1699): 3401–3410.

Lande R, Barrowclough GF (1987). Effective population size, genetic variation, and their use in population management. Pages 87–123 in Soulé ME (editor): *Viable Populations for Conservation*. Cambridge University Press, Cambridge, UK. 189 pages.

Lariviére S, Pasitschniak-Arts M (1996). *Vulpes vulpes*. *Mammalian Species*, 537: 1–11.

Lewis JC, Sallee KL, Golightly RT (1999). Introduction and range expansion of nonnative red foxes (Vulpes vulpes) in California. *The American Midland Naturalist*, 142 (2): 372–381.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, 18 (4): 189–197.

Mansergh I, Broome L (1994). *The Mountain Pygmy-possum of the Australian Alps*. New South Wales University Press, Sydney, Australia. 114 pages.

Marko PB, Hart MW (2011). The complex analytical landscape of gene flow inference. *Trends in Ecology and Evolution*, 26 (9): 448–456.

Martyn I, Kuhn TS, Mooers AO, Moulton V, Spillner A (2012). Computing evolutionary distinctiveness indices in large scale analysis. *Algorithms for Molecular Biology*, 7: 6.

May RM (1990). Taxonomy as destiny. *Nature*, 347 (6289): 129–130.

Mead EM, Mead JI (1989). Snake Creek burial cave and a review of the Quaternary mustelids of the Great Basin. *The Great Basin Naturalist*, 49 (2): 143–154.

Menkhorst P, Broome L, Driessen M (2008). *Burramys parvus*. IUCN Red List of Threatened Species. Available: http://www.iucnredlist.org/details/full/3339/0. Accessed 28 November 2011.

Merriam CH (1900). Preliminary revision of the North American red foxes. *Proceedings of the Washington Academy of Sciences*, 2: 661–676.

Merriam CH, Stejnrger L (1891). *Results of a Biological Reconnoissance of South-Central Idaho*. North American Fauna No. 5, Government Printing Office, Washington, DC, USA. 133 pages.

Minh BQ, Klaere S, von Haeseler A (2007). Phylogenetic diversity on split networks. Technical Report NI07090-PLG. Cambridge: Isaac Newton Institute. 19 p.

Minh BQ, Pardi F, Klaere S, von Haeseler A (2009a). Budgeted phylogenetic diversity on circular split systems. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 6 (1): 22–29.

Minh BQ, Klaere S, von Haeseler A (2009b). Taxon selection under split diversity. *Systematic Biology*, 58 (6): 586–594.

Mitrovski P, Heinze DA, Guthridge K, Weeks AR (2005). Isolation and characterization of microsatellite loci from the Australian endemic mountain pygmy-possum, *Burramys parvus* Broom. *Molecular Ecology Notes*, 5 (2): 395–397.

Mitrovski P, Heinze DA, Broome L, Hoffmann AA, Weeks AR (2007). High levels of variation despite genetic fragmentation in populations of the endangered mountain pygmy-possum, *Burramys parvus*, in alpine Australia. *Molecular Ecology*, 16 (1): 75–87.

Mitrovski P, Hoffmann AA, Heinze DA, Weeks AR (2008). Rapid loss of genetic variation in an endangered possum. *Biology Letters*, 4 (1): 134–138.

Mooers AØ, Faith DP, Maddison WP (2008). Converting endangered species categories to probabilities of extinction for phylogenetic conservation prioritization. *PLoS One*, 3: e3700.

Moritz C (1994). Defining 'evolutionarily significant units' for conservation. *Trends in Ecology and Evolution*, 9 (10): 373–375.

Newberry JS (1857). Report upon the zoology of the route [from the Sacramento Valley to the Columbia River]. Pages 35–110, plates 26–29 in Abbot HL (editor): *Explorations and Surveys for a Railroad Route from the Mississippi River to the Pacific Ocean (Vol. 6)*. Beverley Tucker, Washington, DC, USA. 620 pages.

Norris PW (1881). *Fifth Annual Report of the Superintendent of the Yellowstone National Park*. Government Printing Office, Washington, DC, USA. 81 pages.

Osborne MJ, Norman JA, Christidis L, Murray ND (2000). Genetic distinctness of isolated populations of an endangered marsupial, the mountain pygmy-possum, *Burramys parvus*. *Molecular Ecology*, 9 (5): 609–613.

Pavoine S, Ollier S, Dufour AB (2005). Is the originality of a species measurable? *Ecology Letters*, 8 (6): 579–586.

Perrine JD (2005). Ecology of red fox (*Vulpes vulpes*) in the Lassen Peak region of California, USA. PhD dissertation, University of California, Berkeley, CA, USA. 236 pages.

Perrine JD, Pollinger JP, Sacks BN, Barrett RH, Wayne RK (2007). Genetic evidence for the persistence of the critically endangered Sierra Nevada red fox in northern California. *Conservation Genetics*, 8 (5): 1083–1095.

Péwé TL (1983). The periglacial environment in North America during Wisconsin time. Pages 157–189 in Wright HE, Porter SC (editors): *Late Quaternary Environments of the United States*. University of Minnesota Press, Minneapolis, MN, USA. 433 pages.

Péwé TL and Hopkins DM (1967). Mammal remains of pre-Wisconsin age in Alaska. Pages 266–287 in Hopkins DM (editor): *The Bering Land Bridge*. Stanford University Press, Palo Alto, CA, USA. 495 pages.

Poczai P, Mátyás KK, Szabó I, Varga I, Hyvönen J, *et al*. (2011). Genetic variability of thermal *Nymphaea* (Nymphaeaceae) populations based on ISSR markers: implications on relationships, hybridization, and conservation. *Plant Molecular Biology Reporter*, 29 (4): 906–918.

Posada D, Crandall KA (2001). Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution*, 16 (1): 37–45.

R-Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. 1706 p.

Redding DR, Mooers AØ (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology*, 20 (6): 1670–1678.

Redding DW, Mooers AØ (2010). Can systematists help decide the relative worth of bits of biodiversity? *The Systematist*, 32: 4–8.

Rice WR (1989). Analyzing tables of statistical tests. *Evolution*, 43 (1): 223–225.

Ritchie EG, Johnson CN (2009). Predator interactions, mesopredator release and biodiversity conservation. *Ecology Letters*, 12 (9): 982–998.

Roman J, Santhuff SD, Moler PE, Bowen BW (1999). Population structure and cryptic evolutionary units in the alligator snapping turtle. *Conservation Biology*, 13 (1): 135–142.

Rozen S, Skaletsky HJ (1998). Primer3. Available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html.

Ryder OA (1986). Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, 1 (1): 9–10.

Sacks BN, Statham MJ, Perrine JD, Wisely SM, Aubry KB (2010). North American montane red foxes: expansion, fragmentation, and the origin of the Sacramento Valley red fox. *Conservation Genetics*, 11 (4): 1523–1539.

Sacks BN, Moore M, Statham MJ, Wittmer HU (2011). A restricted hybrid zone between native and introduced red fox (*Vulpes vulpes*) populations suggests reproductive barriers and competitive exclusion. *Molecular Ecology*, 20 (2): 326–341.

Sacramento Fish and Wildlife Office (2011). Endangered and threatened wildlife and plants; 90-day finding on a petition to list Sierra Nevada red fox as endangered or threatened. *Federal Register*, 77 (1): 45–52.

Sala OE, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, *et al*. (2000). Global biodiversity scenarios for the year 2100. *Science*, 287 (5459): 1770–1774.

Sargeant AB, Allen SH, Hastings JO (1987). Spatial relationships between sympatric coyotes and red foxes in North Dakota. *The Journal of Wildlife Management*, 51 (2): 285–293.

Segelbacher G, Cushman SA, Epperson BK, Fortin MJ, Francois O, *et al*. (2010). Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*, 11 (2): 375–385.

Sodhi NS, Koh LP, Brook BW, Ng PKL (2004). Southeast Asian biodiversity: an impending disaster. *Trends in Ecology and Evolution*, 19 (12): 654–660.

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997). Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, 206 (1–4): 353–373.

Statham MJ, Sacks BN, Aubry KB, Perrine JD, Wisely SM (2012). The origin of recently established red fox populations in the United States: translocations or natural range expansions? *Journal of Mammalogy*, 93 (1): 52–65.

Steel M (2005). Phylogenetic diversity and the greedy algorithm. *Systematic Biology*, 54 (4): 527–529.

Steel M, Mimoto A, Mooers AØ (2007). Hedging our bets: the expected contribution of species to future phylogenetic diversity. *Evolutionary Bioinformatics Online*, 3: 237–244.

Sutton DA (1992). *Tamias amoenus*. *Mammalian Species*, 390: 1–8.

Swanson BJ, Fuhrmann RT, Crabtree RL (2005). Elevational isolation of red fox populations in the Greater Yellowstone Ecosystem. *Conservation Genetics*, 6 (1): 123–131.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, *et al*. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28 (10): 2731–2739.

Taylor WP, Shaw WT (1927). *Mammals and Birds of Mount Rainier National Park*. Government Printing Office, Washington, DC, USA. 249 pages.

Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, *et al*. (2004). Extinction risk from climate change. *Nature*, 427 (6970): 145–148.

Thompson RS, Mead JI (1982). Late Quaternary environments and biogeography in the Great Basin. *Quaternary Research*, 17 (1): 39–55.

Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, *et al*. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, 466 (7310), 1098–1101.

Trombulak SC (1988). *Spermophilus saturatus*. *Mammalian Species*, 322: 1–4.

US Fish and Wildlife Service (2011). Revised recovery plan for the northern spotted owl (*Strix occidentalis caurina*). US Fish and Wildlife Service, Portland, OR, USA. 277 pages.

US Geological Service (2013). *The National Atlas of the United States of America*. US Department of Interior, Geological Service, Washington, DC, USA. Available: http://www.nationalatlas.gov/biology.html. Accessed 28 November 2011.

Vane-Wright RI, Humphries CJ, Williams PH (1991). What to protect?—Systematics and the agony of choice. *Biological Conservation*, 55 (3): 235–254.

Verts BJ, Carraway LN (1998). *Land Mammals of Oregon*. University of California Press, Berkeley, CA, USA. 673 pages.

Voigt DR, Earle BD (1983). Avoidance of coyotes by red fox families. *The Journal of Wildlife Management*, 47 (3): 852–857.

Waples RS (1991). Pacific salmon, *Oncorhynchus* spp., and the definition of 'species' under the endangered species act. *Marine Fisheries Review*, 53 (3): 11–22.

Ward RD (1995). Population genetics of tunas. *Journal of Fish Biology*, 47: 259–280.

Washington Department of Fish and Wildlife (2008). *Priority Habitats and Species List*. Olympia, WA, USA. 177 pages.

Weider LJ, Hobæk A (2000). Phylogeography and arctic biodiversity: a review. *Annales Zoologici Fennici*, 37 (4): 217–231.

Weir BS, Cockerham CC (1984). Estimating f-statistics for the analysis of population structure. *Evolution*, 38 (6): 1358–1370.

Weitzman ML (1992). On diversity. *The Quarterly Journal of Economics*, 107 (2): 363–405.

Weitzman ML (1998). The Noah's Ark Problem. *Econometrica*, 66 (6): 1279–1298.

Westwood RE (1989). Early fur farming in Utah. *Utah Historical Quarterly*, 57: 320–339.

Wisely SM, Maldonado JE, Fleischer RC (2004). A technique for sampling ancient DNA that minimizes damage to museum specimens. *Conservation Genetics*, 5 (1):105–107.

Witting L, Loeschcke V (1995). The optimization of biodiversity conservation. *Biological Conservation*, 71 (2): 205–207.

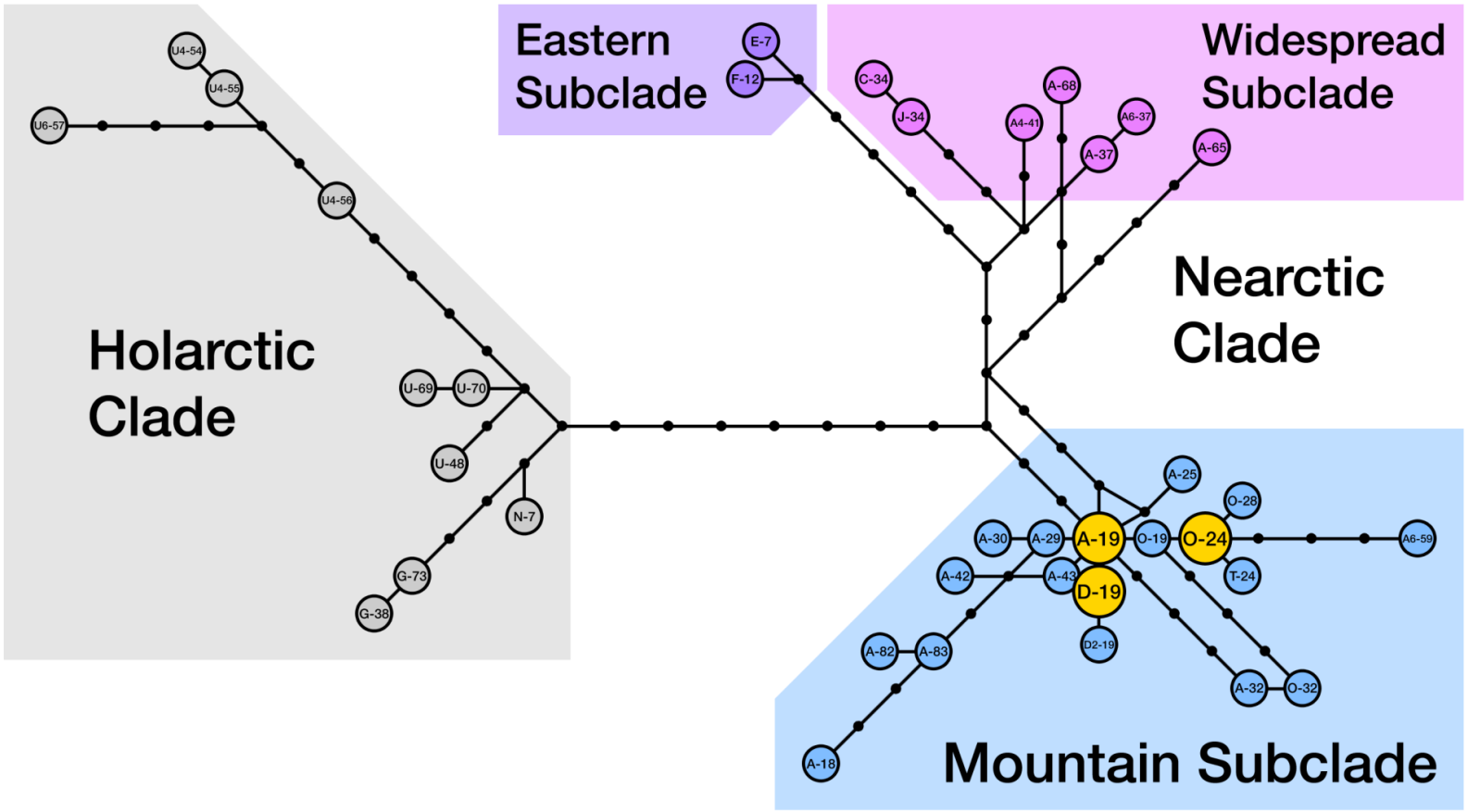
Zhong HM, Zhang HH, Sha WL, Zhang CD, Chen YC (2010). Complete mitochondrial genome of the red fox (*Vulpes vulpes*) and phylogenetic analysis with other canid species. *Dongwuxue Yanjiu*, 31 (2): 122–130.

Zielinski WJ, Truex RL, Schlexer FV, Campbell LA, Carroll C (2005). Historical and contemporary distributions of carnivores in forests of the Sierra Nevada, California, USA. *Journal of Biogeography*, 32 (8): 1385–1407.

Appendix A.

Median-joining network of North American red fox haplotypes reported in previous studies

Median-joining network of North American red fox haplotypes reported in previous studies (Perrine *et al*. 2007; Aubry *et al*. 2009; Sacks *et al*. 2010), representing 354 bp of the cytochrome *b* gene and 342 bp of the D-loop region of the mitochondrial genome. Sequences were obtained from GenBank (accession numbers EF064207–EF064220, FJ830756–FJ830829, FJ840491, GU004541, GQ911200–GQ911203, GU224186–GU224187) and aligned visually using MEGA v5 (Tamura *et al*. 2011). I constructed the median-joining network in Network v4.611 (Fluxus Technology), treating polymorphisms in cytochrome b as twice the weight of those in the D-loop. The haplotypes of interest in the present study are indicated in yellow.


Eastern Subclade
Widespread Subclade
Nearctic Clade
Holarctic Clade
Mountain Subclade
E-7
F-12
C-34
J-34
A4-41
A-68
A6-37
A-37
A-65
U4-54
U4-55
U6-57
U4-56
U-69
U-70
U-48
N-7
G-73
G-38
A-25
O-28
A-30
A-29
A-19
O-19
O-24
A6-59
T-24
A-42
A-43
D-19
D2-19
A-82
A-83
A-32
O-32
A-18

Appendix B.

List of red fox specimens

List of red fox specimens used in this study (Chapter 2).

No.	Sample ID	State or province	Latitude	Longitude	Elevation in ft.	Year collected	Time period	Population	Sequence type	Haplotype
1	82	Washington	46.755	-121.717	5942	1921	historical	NC	targeted	w2
2	83	Washington	46.468	-121.436	5739	1932	historical	NC	targeted	w2
3	84	Washington	46.758	-121.598	4660	1900	historical	NC	targeted	w5
4	85	Washington	46.101	-121.506	3787	1895	historical	NC	targeted	—[1]
5	86	Washington	46.151	-121.491	6580	1900	historical	NC	targeted	w5
6	279	Washington	46.161	-121.506	6475	1901	historical	NC	targeted	w5
7	280	Washington	46.143	-121.489	6080	1900	historical	NC	targeted	w5
8	281	Washington	46.151	-121.491	6580	1900	historical	NC	targeted	w5
9	s12-246	Washington	46.750	-121.814	2760	2012	modern	NC	full-locus	w6 (?)[2]
10	UW-31868	Washington	46.900	-121.554	3600	1900	historical	NC	targeted	—[1]
11	UW-32524	Washington	46.332	-121.493	4820	1900	historical	NC	targeted	—[1]
12	UW-32525	Washington	46.332	-121.493	4820	1900	historical	NC	targeted	—[1]
16	UW-32532	Washington	47.637	-122.837	20	1978	modern	—[3]	targeted	w5
17	UW-32534	Washington	47.679	-122.662	220	1978	modern	—[3]	targeted	w5
18	UW-32541	Washington	47.595	-122.643	260	1980	modern	—[3]	targeted	w5
13	UW-32554	Washington	46.943	-121.474	4240	1900	historical	NC	targeted	—[1]
14	UW-32558	Washington	46.969	-121.491	3800	1900	historical	NC	targeted	—[1]
15	UW-32559	Washington	46.969	-121.491	3800	1900	historical	NC	targeted	—[1]
19	UW-32950	Washington	46.608	-120.469	1140	1980	modern	—[3]	targeted	w5
20	UW-34306	Washington	47.444	-122.353	40	1984	modern	—[3]	targeted	w6
21	78	Oregon	42.183	-120.229	6530	1924	historical	SC	targeted	—[1]
22	79	Oregon	42.703	-121.993	4179	1897	historical	SC	targeted	w3
23	80	Oregon	44.056	-121.313	3628	1932	historical	SC	targeted	w3
24	195	Oregon	42.942	-121.116	4981	1931	historical	SC	targeted	w7
25	286	Oregon	45.569	-117.526	2929	1924	historical	NR	targeted	w1
26	MVZ-89553	Oregon	42.894	-122.135	6420	1939	historical	SC	targeted	w7
27	OSU-1175	Oregon	44.545	-123.238	220	1970	modern	CO	targeted	—[1]
28	OSU-5543	Oregon	44.315	-123.404	360	1951	modern	CO	targeted	—[1]
29	OSU-8778	Oregon	43.327	-119.320	4280	1983	modern	IW	targeted	w6
30	OSU-8779	Oregon	43.351	-118.936	4100	1988	modern	IW	targeted	—[1]
31	s10-357	Oregon	43.055	-124.439	20	2010	modern	CO	full-locus	w3
32	s10-358	Oregon	42.970	-124.469	20	2010	modern	CO	full-locus	w3
33	s10-359	Oregon	42.949	-124.480	20	2010	modern	CO	full-locus	w3
34	s10-360	Oregon	42.949	-124.480	20	2010	modern	CO	full-locus	w3
35	s10-361	Oregon	42.987	-124.462	20	2010	modern	CO	full-locus	—[1]
36	s11-425	Oregon	44.306	-120.706	3000	2011	modern	IW	full-locus	w6
37	s11-426	Oregon	45.303	-118.069	2800	2011	modern	IW	full-locus	w6
38	s11-427	Oregon	45.299	-118.112	3600	2011	modern	IW	full-locus	w6
39	s11-428	Oregon	45.308	-118.073	2751	2011	modern	IW	full-locus	w6
40	s11-429	Oregon	44.505	-121.229	2540	2011	modern	IW	full-locus	w6
41	s11-430	Oregon	44.505	-121.229	2540	2011	modern	IW	full-locus	w6
42	s11-431	Oregon	44.299	-120.871	3390	2011	modern	IW	full-locus	w6
43	UPS-13382	Oregon	44.156	-123.311	380	1970	modern	CO	targeted	w3
44	UPS-13383	Oregon	44.156	-123.311	380	1970	modern	CO	targeted	w3
45	UPS-14885	Oregon	44.189	-122.800	440	1970	modern	CO	targeted	—[1]
46	101	California	36.568	-118.320	14030	1891	historical	SN	targeted	w2
47	103	California	37.700	-119.510	8050	1893	historical	SN	targeted	w3
48	104	California	40.801	-121.636	7060	1900	historical	SC	targeted	w1
49	217	California	41.301	-122.356	4417	1893	historical	SC	targeted	w2
50	219	California	36.568	-118.320	14030	1891	historical	SN	targeted	w2
51	291	California	36.562	-118.773	6444	1896	historical	SN	targeted	w2
52	292	California	39.727	-122.096	200	1906	historical	SV	targeted	w4[5]
53	F01	California	40.490	-121.510	10160	1998	modern	SC	full-locus	w2
54	F02	California	40.550	-121.510	10160	1998	modern	SC	full-locus	w2
55	F03	California	40.590	-121.510	10160	2000	modern	SC	full-locus	w2
56	F05	California	40.550	-121.460	6320	2000	modern	SC	full-locus	w2

57	M01	California	40.550	-121.560	6500	1998	modern	SC	full-locus	w2
58	MVZ-16252	California	36.800	-118.500	9800	1911	historical	SN	targeted	w9 (?)[4]
59	MVZ-16374	California	36.210	-118.260	8760	1911	historical	SN	targeted	—[1]
60	MVZ-23696	California	37.750	-119.480	7700	1916	historical	SN	targeted	—[1]
61	MVZ-32809	California	38.046	-119.260	9760	1922	historical	SN	targeted	w1
62	MVZ-33382	California	37.996	-119.190	10000	1922	historical	SN	targeted	w1
63	MVZ-33472	California	37.824	-119.191	11000	1923	historical	SN	targeted	w1
64	MVZ-33550	California	39.280	-122.180	100	1923	historical	SV	targeted	w4
65	MVZ-34984	California	40.530	-120.830	5500	1925	historical	SC	targeted	—[1]
66	MVZ-36492	California	39.223	-122.010	150	1926	historical	SV	targeted	w4
67	MVZ-36494	California	39.300	-122.020	200	1926	historical	SV	targeted	—[1]
68	MVZ-36495	California	39.223	-122.010	150	1926	historical	SV	targeted	—[1]
69	MVZ-36497	California	39.420	-122.000	100	1926	historical	SV	targeted	—[1]
70	MVZ-44095	California	39.223	-122.010	150	1929	historical	SV	targeted	w4
71	MVZ-68857	California	41.580	-121.600	6680	1934	historical	SC	targeted	w8
72	MVZ-68858	California	41.580	-121.500	7020	1934	historical	SC	targeted	w8
73	MVZ-70285	California	39.280	-122.180	100	1935	historical	SV	targeted	w4
74	MVZ-95401	California	39.390	-120.180	5800	1941	modern	SN	targeted	—[1]
75	MVZ-115439	California	39.223	-122.010	150	1948	modern	SV	targeted	w4
76	s10-128	California	40.507	-121.192	6960	2009	modern	SC	targeted	w2
77	s10-129	California	40.507	-121.192	6960	2009	modern	SC	targeted	w2
78	s10-140	California	40.502	-121.179	6600	2009	modern	SC	targeted	w2
79	s10-374	California	41.680	-120.218	4560	2010	modern	IW	full-locus	w6
80	X1	California	39.120	-122.180	120	2001	modern	SV	full-locus	w4
81	X2	California	39.360	-121.700	100	2001	modern	SV	full-locus	w4
82	X3	California	39.400	-121.700	100	2001	modern	SV	full-locus	—[1]
83	206	Idaho	43.588	-115.044	6080	1916	historical	NR	targeted	w1
84	DK1	Idaho	45.020	-113.966	5480	2006	modern	NR	full-locus	w1
85	DK10	Idaho	43.509	-114.825	6450	2006	modern	IW	full-locus	w7
86	DK11	Idaho	44.032	-116.359	3980	2006	modern	IW	full-locus	w6
87	DK15	Idaho	43.350	-115.384	4880	2006	modern	IW	full-locus	w6
88	VVM6	Idaho	44.060	-116.410	4360	2006	modern	IW	full-locus	w6
89	VVM7	Idaho	44.060	-116.410	4360	2006	modern	IW	full-locus	w6
90	VVM16	Idaho	43.626	-116.320	2520	2006	modern	IW	full-locus	w6
91	VVM17	Idaho	44.050	-116.930	2165	2006	modern	IW	full-locus	w6
92	s12-820	Montana	48.065	-115.884	3200	2012	modern	NR	full-locus	w5
93	s12-821	Montana	48.065	-115.884	3200	2012	modern	NR	full-locus	—[1]
94	s12-822	Montana	48.065	-115.884	3200	2012	modern	NR	full-locus	—[1]
95	s12-823	Montana	48.065	-115.884	3200	2012	modern	NR	full-locus	—[1]
96	97	Wyoming	44.126	-109.629	6480	1915	historical	NR	targeted	w3
97	99	Wyoming	42.921	-109.507	9850	1893	historical	NR	targeted	w7
98	100	Wyoming	43.381	-109.893	9600	1915	historical	NR	targeted	w3
99	214	Wyoming	44.967	-110.710	6622	1902	historical	NR	targeted	w3
100	216	Wyoming	43.288	-110.558	6290	1917	historical	NR	targeted	w3
101	288	Wyoming	43.957	-109.029	7620	1915	historical	NR	targeted	w3
102	290	Wyoming	44.525	-109.055	4993	1916	historical	NR	targeted	w3
103	315	Wyoming	44.347	-106.698	4639	1915	historical	NR	targeted	w1
104	94	Utah	38.083	-110.772	5015	1916	historical	SR	targeted	w10 (?)[4]
105	MAM-3752	Utah	41.203	-112.163	4320	2007	modern	IW	full-locus	w5
106	MAM-3757	Utah	41.203	-112.163	4320	2007	modern	IW	full-locus	w5
107	MAM-3758	Utah	41.203	-112.163	4320	2007	modern	IW	full-locus	w5
108	MAM-3759	Utah	41.203	-112.163	4320	2007	modern	IW	full-locus	—[1]
109	BB1	Nevada	39.656	-116.063	6000	2003	modern	IW	full-locus	w7
110	BB2	Nevada	39.656	-116.063	6000	2003	modern	IW	full-locus	w7
111	BB4	Nevada	39.656	-116.063	6000	2003	modern	IW	full-locus	w7
112	BB5	Nevada	39.656	-116.063	6000	2003	modern	IW	full-locus	w7
113	BB6	Nevada	39.656	-116.063	6000	2003	modern	IW	full-locus	w7
114	BB7	Nevada	41.631	-115.970	7920	2002	modern	IW	full-locus	w6
115	BB10	Nevada	41.876	-115.431	6200	2002	modern	IW	full-locus	—[1]

116	BB11	Nevada	41.766	-116.066	7600	2002	modern	IW	full-locus	w7
117	BB12	Nevada	41.628	-115.963	8040	2002	modern	IW	full-locus	w6
118	s10-238	Nevada	41.678	-116.098	6680	2010	modern	IW	full-locus	w7
119	s10-239	Nevada	41.678	-116.098	6680	2010	modern	IW	full-locus	w7
120	S10-240	Nevada	41.678	-116.098	6880	2010	modern	IW	full-locus	w6
121	88	Colorado	39.129	-106.178	11843	1923	historical	SR	targeted	w6
122	89	Colorado	40.088	-108.244	5700	1915	historical	SR	targeted	w1
123	90	Colorado	38.680	-107.076	10294	1952	modern	SR	targeted	w2
124	210	New mexico	35.906	-105.666	12454	1903	historical	SR	targeted	w1
125	211	New mexico	36.593	-105.450	9424	1903	historical	SR	targeted	w1
126	241	Alberta	53.473	-113.451	2242	1896	historical	—[3]	targeted	w1

[1]Sequencing failed to produce an unambiguous haplotype.

[2]Results contradict previous haplotype assignment due to suspected contamination; sample excluded from analyses.

[3]Sampled population out of range; sample excluded from analyses.

[4]Sequencing produced a novel partial haplotype; sample excluded from analyses.

[5]Previously identified as an "A-19" haplotype by Sacks *et al*. (2010), but revised here to "D-19".

Appendix C.

PCR and sequencing primers

PCR and sequencing primers used in this study. Primers used for complete sequence reads are listed first. Note (*) that only nonrepetitive sections of the control region (CR-I, 345 bp; CR-II, 711 bp) were included in my analyses.

Primer name	Locus	Primer length (bases)	Primer sequence (5' to 3')	Product length (bases)	Analyzed length (bases)
VVmc-cF	cytochrome b	20	CAA AGC TCA TCA CAA AAA CA	1338	1298
VVmc-cR		20	TTG AGT CTT GGG GAG AAT TA		
VVmd-cF	control region	20	TAA TTC TCC CCA AGA CTC AA	1430-1432	1054-1056*
VVmd-cR		20	GGA CCA AAC CTT TGT GTT TA		
VVx1-cF	cytochrome c oxidase I	20	AAG CTG CTT CTT TGA ATT TG	1793	1753
VVx1-cR		20	GGG TTT CAC CTA TGA TTT GA		
VVmc-367F	cytochrome b	20	ACA TAC ATG CAA ACG GAG CA	198	158
VVmc-564R		20	TAA TTA CGG TTG CCC CTC AG		
VVmc-780F		21	CCG TTT CAC CCC TAC TAC ACA	201	160
VVmc-980R		20	AGG GAT GGA CCG GAG AAT AG		
VVmd-7F	control region	20	TCC CCA AGA CTC AAG GAA GA	157-159	117-119
VVmd-164R		20	GGC ATA GCA GTA ATG CAC GA		
VVmd-553F		20	CAC TCA GCT ATG ACC GCA AC	168	127
VVmd-720R		21	TGT CCT GAA ACC ATT GAC TGA		
VVmd-1104F		20	CCC CCG TAA ACT CAT GCT AC	163	123
VVmd-1266R		20	AAG ATC TGG CGG GTG TAA GA		
VVmd-1318F		20	ATA ATT CGC ACC GAC AAT GC	113	70
VVmd-1430R		23	GAC CAA ACC TTT GTG TTT ATG GA		

(B) Internal primers.

Primer name	Locus	Primer length (bases)	Primer sequence (5' to 3')	Product length (bases)	Analyzed length (bases)
VV_CytB_Fb	cytochrome b	20	TTC TAT CTG CTA TCC CCT AT	768	not analyzed
VV_Dloop_Fb	control region	20	GGG AAC TTG CTA TCA CTC AG	892	not analyzed
VV_Cox1_Fb	cytochrome c oxidase I	20	CCG GAG TCT CTT CAA TTT TA	1216	not analyzed
VVmc-498R	cytochrome b	20	CCA TGG TTG CGA ACA ATA AG	132	not analyzed
VVmc-897R		22	GGA TGT AAT TGT CTG GGT CTC C	118	not analyzed

Appendix D.

Frequency of selected native western haplotypes

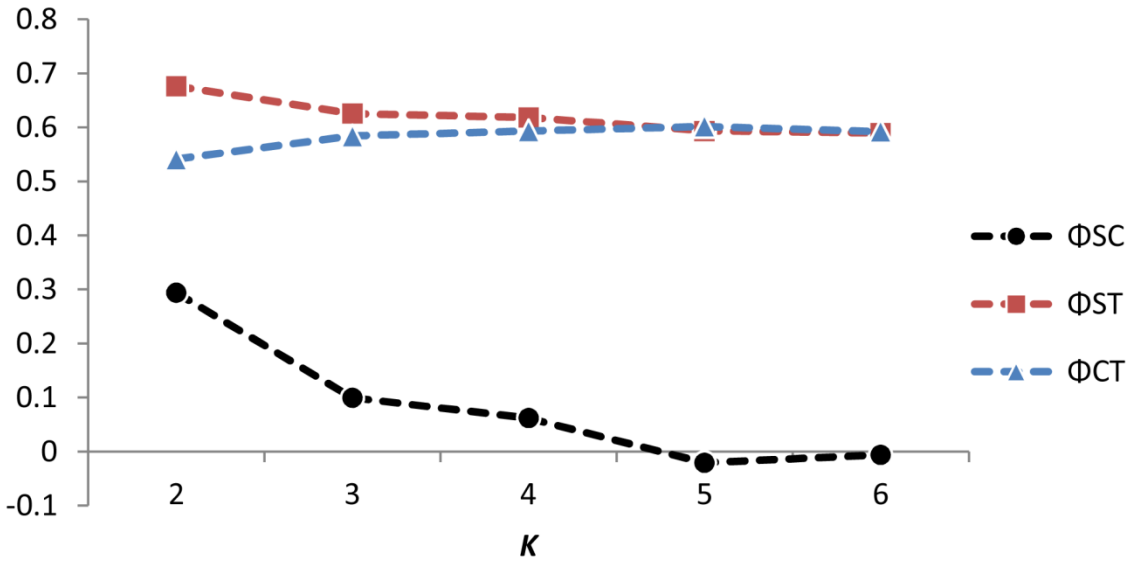
Frequency of selected native western haplotypes among 8 red fox populations sampled in this study.

Population	*w* 1	*w* 2	*w* 3	*w* 4	*w* 5	*w* 6	*w* 7	*w* 8
NR	4	—	6	—	1	—	1	—
SC	3	1	—	—	—	1	—	—
NC	—	2	—	—	5	—	—	—
SC	1	9	2	—	—	—	2	2
SN	3	3	1	—	—	—	—	—
SV	—	—	—	8	—	—	—	—
IW	—	—	—	—	3	18	9	—
CO	—	—	6	—	—	—	—	—
Total	11	15	15	8	9	19	12	2

Appendix E.

Spatial analysis of molecular variance

Spatial analysis of molecular variance (SAMOVA) under 5 different *K* models, showing trends in Φ_{SC}, Φ_{ST}, and Φ_{CT}.



Appendix F.

Mathematical treatment of SH and HED and associated algorithms

Publication Note

The material in this appendix was previously published in *PLoS ONE* (see Chapter 3 for full citation).

Supplementary material for: Prioritizing populations for conservation using phylogenetic networks

Logan Volkmann Iain Martyn Vincent Moulton
Andreas Spillner Arne Mooers

August 29, 2013

1 Preliminaries

Let X be a finite set of taxa and Σ a set of bipartitions or *splits* of X. Such a set Σ is called a *split system* on X. A split $S = \{A, B\}$ of X is denoted by $S = A|B$. For each split S of X and any $x \in X$ we denote by $S(x)$ the set in S that contains x and by $\overline{S}(x)$ the other set in S. A split system Σ is called *compatible* if its elements correspond precisely to the edges of some (necessarily unique) phylogenetic tree (cf. [7]). A *weighted* split system on X is a pair (Σ, λ) consisting of a split system Σ on X together with a map $\lambda : \Sigma \to \mathbb{R}_{\geq 0}$ that assigns a non-negative weight to each split in Σ.

The *phylogenetic diversity* of a subset $Y \subseteq X$ [8] relative to a weighted split system (Σ, λ) is given by:

$$PD_{(\Sigma,\lambda)}(Y) = \sum_{\substack{A|B \in \Sigma \\ A \cap Y \neq \emptyset, B \cap Y \neq \emptyset}} \lambda(A|B).$$

This quantity is also known as the *split diversity* [6]. It was originally defined for compatible split systems, i.e. for phylogenetic trees (see e.g. [9]). In particular, the phylogenetic diversity with respect to some tree is just the split diversity relative to the compatible split system corresponding to the tree, where split weights correspond to branch lengths.

1.1 The Shapley value (SH)

The Shapley value of a taxon x with respect to a phylogenetic tree on X, as defined in [3], can be extended to any weighted split system (Σ, λ) on X as follows:

$$\psi_x^{sh}(\Sigma, \lambda) = \frac{1}{|X|!} \sum_{Z \subseteq X, x \in Z} (|Z|-1)!(|X|-|Z|)!\big(PD_{(\Sigma,\lambda)}(Z) - PD_{(\Sigma,\lambda)}(Z - \{x\})\big).$$

Shapley values of elements relative to a tree correspond to values in this last formula relative to the corresponding compatible split system. As with phylogenetic trees, it is not hard to show that the Shapley value can also be written as the following linear combination of the weights of the splits in Σ:

$$\psi_x^{sh}(\Sigma, \lambda) = \sum_{S \in \Sigma} \frac{|\overline{S}(x)|}{|X||S(x)|} \lambda(S). \tag{1}$$

Note that, since the coefficients in this linear combination can be computed in polynomial time, the Shapley value of x with respect to (Σ, λ) can also be computed in polynomial time.

In case the split system Σ has a special structure, there can be efficient algorithms for computing the Shapley values $\psi_x^{sh}(\Sigma, \lambda)$ for all $x \in X$. For example, in case Σ is compatible there is an algorithm with run time in $O(n)$, $n = |X|$ [5]. Here we show that for so-called circular split systems there is an algorithm with run time in $O(n + m)$, $m = |\Sigma|$, for simultaneously computing Shapley values for all x in X.

A split system Σ is *circular* if there exists an ordering $x_1, x_2, \ldots, x_n$ of the taxa in X such that for every split $A|B \in \Sigma$ there are $1 \leq i \leq j < n$ with $A = \{x_i, x_{i+1}, \ldots, x_j\}$ or $B = \{x_i, x_{i+1}, \ldots, x_j\}$. We shall say that such an ordering $x_1, x_2, \ldots, x_n$ is *suitable* for the circular split system Σ. Note that split systems corresponding to NeighborNet networks are always circular [1] (a fact that also implies that NeighborNet networks can always be drawn in the plane without crossing edges).

Now, assume that we are given an ordering $x_1, x_2, \ldots, x_n$ of the taxa in X that is suitable for Σ. Moreover, assume that, for each split $A|B \in \Sigma$, the two indices $1 \leq i \leq j < n$ with $\{x_i, x_{i+1}, \ldots, x_j\} = A$ or $\{x_i, x_{i+1}, \ldots, x_j\} = B$ are stored. Then, for any given $S \in \Sigma$ and any given $x \in X$, one can compute $|S(x)|$ in constant time and, therefore, by evaluating (1), we obtain $\psi_{x_1}^{sh}(\Sigma, \lambda)$ in $O(m)$ time.

Next, we perform a preprocessing of the splits in Σ. More specifically, we form subsets Σ_k, $1 \leq k < n$, by putting each split of the form $S = \{x_i, x_{i+1}, \ldots, x_j\}|X - \{x_i, x_{i+1}, \ldots, x_j\} \in \Sigma$ into Σ_j and also, in case $i > 1$, into Σ_{i-1}. Intuitively, the splits in Σ_k are those whose associated interval $\{x_i, x_{i+1}, \ldots, x_j\}$ starts directly after taxon x_k or ends with taxon x_k.

Our algorithm for computing Shapley values processes the taxa $x_2, x_3, \ldots, x_n$ in the given ordering. Assuming that the Shapley value $\psi_{x_{k-1}}^{sh}(\Sigma, \lambda)$ has already been computed, we now explain how to compute $\psi_{x_k}^{sh}(\Sigma, \lambda)$, $2 \leq k \leq n$. It is not hard to see that only the splits in Σ_{k-1} might change their contribution to Shapley values when going from x_{k-1} to x_k. So, first we subtract the values $\frac{|\overline{S}(x_{k-1})|}{|X||S(x_{k-1})|}\lambda(S)$ from $\psi_{x_{k-1}}^{sh}(\Sigma, \lambda)$, and then we add $\frac{|\overline{S}(x_k)|}{|X||S(x_k)|}\lambda(S)$ for each $S \in \Sigma_{k-1}$. This yields $\psi_{x_k}^{sh}(\Sigma, \lambda)$.

The $O(n+m)$ bound on the run time for this algorithm follows by observing that $\psi_{x_k}^{sh}(\Sigma, \lambda)$ can be computed in $O(|\Sigma_{k-1}|)$ time and noting that every split in Σ is contained in at most two of the sets Σ_h, $1 \leq h < n$.

Remark 1.1 To achieve the $O(n+m)$ run time the circular split system must be stored in a succinct way such as the one described above. Note that this is not the case for commonly used file formats, such as, for example, the *nexus*-format employed in the SplitsTree software package [4], which uses $O(n \cdot m)$ space to store an arbitrary split system Σ with m splits by listing, for each split in $S \in \Sigma$, the elements in one of the sets in S.

1.2 Heightened evolutionary distinctiveness (HED)

The HED value of a taxon x in X is computed in terms of probabilities $p(y)$ that are given for each element y in X. As with Shapley values, HED was originally defined for trees [10] but extends naturally to weighted split systems (Σ, λ) on X:

$$\psi_x^{hed}(\Sigma,\lambda) = \sum_{S\in\Sigma} \Big(\prod_{y\in(S(x)-\{x\})} p(y)\Big) \cdot \Big(1 - \prod_{y\in\overline{S}(x)} p(y)\Big) \cdot \lambda(S). \qquad (2)$$

Again, HED values of elements relative to a tree correspond to values in this last formula relative to the corresponding compatible split system. The approach taken for computing the Shapley values for a weighted circular split system outlined above can be adapted for HED as follows. Note that for the special case of phylogenetic trees this was already shown in [5]. Here we briefly outline how $\psi_x^{hed}(\Sigma,\lambda)$ can be computed for all $x \in X$ in $O(n+m)$ time for any weighted circular split system (Σ, λ).

Let $x_1, x_2, \ldots, x_n$ be an ordering of the taxa in X that is suitable for Σ. As for trees [5], we first perform a preprocessing step to compute, for every split $A|B \in \Sigma$, the values $\prod_{x\in A} p(x)$ and $\prod_{x\in B} p(x)$. To do this, we first perform a range minimum query preprocessing in $O(n)$ time so that afterwards we can decide in $O(1)$ time whether a given interval $x_i, x_{i+1}, \ldots, x_j$ contains an element y with $p(y) = 0$ using a range minimum query [2]. This yields all the values $\prod_{x\in A} p(x)$ and $\prod_{x\in B} p(x)$, respectively, that are equal to 0 in $O(n+m)$ time. The remaining values can also be computed in $O(n+m)$ time by restricting the ordering $x_1, x_2, \ldots, x_n$ to those $y \in X$ with $p(y) > 0$ and then computing the product over every prefix of the restricted ordering in $O(n)$ time. From these prefix-products the values $\prod_{x\in A} p(x)$ and $\prod_{x\in B} p(x)$, respectively, that are not equal to 0 can then be computed in constant time each, yielding $O(n+m)$ time in total, as claimed.

After finishing the preprocessing, we just use the algorithm for computing the Shapley values to compute $p(x)\psi_x^{hed}(\Sigma,\lambda)$ for all $x \in X$. This yields $\psi_x^{hed}(\Sigma,\lambda)$ for all $x \in X$ with $p(x) > 0$. For those $x \in X$ with $p(x) = 0$ we run a second phase just as we did for phylogenetic trees (cf. [5]).

References

[1] D. Bryant and V. Moulton. NeighborNet: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*,

21:255–265, 2004.

[2] J. Fischer and V. Heun. A new succinct representation of RMQ-information and improvements in the enhanced suffix array. In *Proc. International Symposium on Combinatorics, Algorithms, Probabilistic and Experimental Methodologies*, volume 4614 of *LNCS*, pages 459–470. Springer, 2007.

[3] C.J. Haake, A. Kashiwada, and F.E. Su. The Shapley value of phylogenetic trees. *Journal of Mathematical Biology*, 56:479–497, 2008.

[4] D.H. Huson and D. Bryant. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23:254–267, 2006.

[5] I. Martyn, T. Kuhn, A. Mooers, V. Moulton, and A. Spillner. Computing evolutionary distinctiveness indices in large scale analysis. *Algorithms for Molecular Biology*, 7, 2012.

[6] B. Minh, S. Klaere, and A. von Haeseler. Taxon selection under split diversity. *Systematic Biology*, 57:586–594, 2009.

[7] C. Semple and M. Steel. *Phylogenetics*. Cambridge University Press, 2003.

[8] A. Spillner, B. T. Nguyen, and V. Moulton. Computing phylogenetic diversity for split systems. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 5:235–244, 2008.

[9] M. Steel. Phylogenetic diversity and the greedy algorithm. *Systematic Biology*, 54:527–529, 2005.

[10] M. Steel, A. Mimoto, and A. Mooers. Hedging our bets: the expected contribution of species to future phylogenetic diversity. *Evolutionary Bioinformatics*, 3:237–244, 2007.

Appendix G.

Annotated R code for calculating SH and HED

Publication Note

The material in this appendix was previously published in *PLoS ONE* (see Chapter 3 for full citation).

```
#Scripts used for: "Prioritizing populations for conservation using phylogenetic
networks"
#L. Volkmann, I. Martyn, V. Moulton, A. Spillner, and A.Ø. Mooers

#Scripts prepared by I. Martyn, 2011
#Annotations by L. Volkmann, 2014

#Last Updated January 2014

#-----------------------------------------------------------------------------
#-CONTENTS------------------------------------------------------------------
#-----------------------------------------------------------------------------

#-1: Calculating SH on a Phylogenetic Network-----------------------------

#-2: Calculating HED on a Phylogenetic Network----------------------------

#-----------------------------------------------------------------------------

#Open the following libraries:

library(gdata)
library(motmot)

#-1-----------------------------------------------------------------------------
#-CALCULATING SH ON A PHYLOGENETIC NETWORK-----------------------------------
#-----------------------------------------------------------------------------

#This section generates Shapley (SH) rankings from a phylogenetic network (made in
SplitsTree).
#Load the NetworksSH function first (in a separate R script), then load the data file,
which contains output from SplitsTree.

file="PATH NAME"
file

#-----------------------------------------------------------------------------
#NetworksSH function:

NetworksSH<-function(file)
{

#Read in the complete Nexus file but keep only the splits information:

CompleteNexus=scan(file, what = "", sep = "\n", quiet = TRUE,skip = 0)

#Identify beginning of splits block and allocate taxa:

st=grep("CYCLE",CompleteNexus)
taxset=unlist(strsplit(strsplit(CompleteNexus[st],"CYCLE")[[1]][2]," "))
taxset=unlist(strsplit(strsplit(strsplit(CompleteNexus[st],"CYCLE")[[1]][2],";")[[1]][
1]," "))
taxset=as.numeric(taxset[taxset>0])
taxset=taxset[!is.na(taxset)]

#Identify end of splits block:

edn=grep("END",CompleteNexus)
edn=edn[which(((edn-st)>0))[1]]
SplitsNexus=CompleteNexus[(st+2):(edn-2)]

#Identify all taxa names:
```

```
st=grep("TAXLABELS",CompleteNexus)
edn=grep("END",CompleteNexus)
edn=edn[which(((edn-st)>0))[1]]
NamesNexus=CompleteNexus[(st+1):(edn-2)]

#Verify that the split system is circular:

Names=matrix(0,length(NamesNexus),2)
for (i in 1:length(NamesNexus))
{number=strsplit(strsplit(strsplit(NamesNexus[i],"
")[[1]][1],"\\[")[[1]][2],"\\]")[[1]][1]
name=strsplit(NamesNexus[i]," ")[[1]][2]
name=strsplit(name,"'")[[1]][2]
Names[i,1]=name
Names[i,2]=number
}
TaxaNotInCycle=setdiff(as.numeric(Names[,2]),taxset)
if (!length(TaxaNotInCycle)==0)
{taxset=c(taxset,TaxaNotInCycle)
}

#Transform data from the splits block to manipulable, array form:

SplitsNexus=strsplit(SplitsNexus,"\t")
SplitsArray=matrix(0,length(SplitsNexus),3)
M=matrix(0,length(SplitsNexus),length(taxset))
for (i in 1:length(SplitsNexus))
{tmp=SplitsNexus[[i]]
sa1=strsplit(tmp[1],"[=]")[[1]][2]
sa1=strsplit(sa1,"]")[[1]][1]
sa2=as.numeric(tmp[2])
sa3=unlist(strsplit(tmp[3],"[ ,]"))
sa3=as.numeric(sa3[sa3>0])
SplitsArray[i,1]=as.numeric(sa1)
SplitsArray[i,2]=sa2
tmp2=intersect(taxset,sa3)
SplitsArray[i,3]=length(tmp2)
M[i,tmp2]=1
}

#Compute Shapley values.
#First taxa:

SH=as.vector(matrix(0,1,length(taxset)))
n=length(SH)
for (i in 1:(dim(M)[1]))
{
if (M[i,taxset[1]]==1)
{SH[1]=SH[1]+((n-SplitsArray[i,3])/(SplitsArray[i,3])*SplitsArray[i,2])
} else {
SH[1]=SH[1]+(SplitsArray[i,3]/(n-SplitsArray[i,3])*SplitsArray[i,2])
}
}

#All other taxa:

for (i in 2:n)
{SH[i]=SH[i-1]
enter=which((M[,taxset[(i-1)]]-M[,taxset[i]])==1)
leave=which((M[,taxset[(i-1)]]-M[,taxset[i]])==-1)
if (length(enter)==0)
```

```
{
} else {
for(j in 1:length(enter))
{a1=SplitsArray[enter[j],3]/(n-SplitsArray[enter[j],3])*SplitsArray[enter[j],2]
s1=((n-SplitsArray[enter[j],3])/(SplitsArray[enter[j],3])*SplitsArray[enter[j],2])
SH[i]=SH[i]+a1-s1
}
}
if (length(leave)==0)
{
} else {
for(k in 1:length(leave))
{a1=SplitsArray[leave[k],3]/(n-SplitsArray[leave[k],3])*SplitsArray[leave[k],2]
s1=((n-SplitsArray[leave[k],3])/(SplitsArray[leave[k],3])*SplitsArray[leave[k],2])
SH[i]=SH[i]-a1+s1
}
}
}

#Normalize and post-process:

SH=SH/length(taxset)

SH=as.matrix(SH[order(taxset)])
rownames(SH)=Names[,1]

tmp1=names(SH[order(SH,decreasing=TRUE),])
SH=matrix(SH[order(SH,decreasing=TRUE)],n,1)
rownames(SH)=tmp1

SH
}

#---------------------------------------------------------------------------
#Run the script on the data file:

NetworksSH(file)

#-2-----------------------------------------------------------------------
#-CALCULATING HED ON A PHYLOGENETIC NETWORK-----------------------------
#---------------------------------------------------------------------------

#This section generates Heightened Evolutionary Distinctiveness (HED) rankings from a
phylogenetic network (made in SplitsTree).
#Load the NetworksHED function first (in a separate R script), then load the data
file, which contains output from Splitstree.

file="PATH NAME"
file

#You will also need to load data file containing the probability of extinction for
each taxon.

ProbExtinctionFile="PATH NAME"
ProbExtinctionFile

#---------------------------------------------------------------------------
#NetworksHED function:

NetworksHED<-function(file,ProbExtinctionFile)
{
```

```
#Read in the complete Nexus file but keep only the splits information:

CompleteNexus=scan(file, what = "", sep = "\n", quiet = TRUE,skip = 0)

#Identify beginning of splits block and allocate taxa:

st=grep("CYCLE",CompleteNexus)
taxset=unlist(strsplit(strsplit(CompleteNexus[st],"CYCLE")[[1]][2]," "))
taxset=unlist(strsplit(strsplit(strsplit(CompleteNexus[st],"CYCLE")[[1]][2],";")[[1]][
1]," "))
taxset=as.numeric(taxset[taxset>0])
taxset=taxset[!is.na(taxset)]

#Identify end of splits block:

edn=grep("END",CompleteNexus)
edn=edn[which(((edn-st)>0))[1]]
SplitsNexus=CompleteNexus[(st+2):(edn-2)]

#Identify all taxa names:

st=grep("TAXLABELS",CompleteNexus)
edn=grep("END",CompleteNexus)
edn=edn[which(((edn-st)>0))[1]]
NamesNexus=CompleteNexus[(st+1):(edn-2)]

#Verify that the split system is circular:

Names=matrix(0,length(NamesNexus),2)
for (i in 1:length(NamesNexus))
{number=strsplit(strsplit(strsplit(NamesNexus[i],"
")[[1]][1],"\\[")[[1]][2],"\\]")[[1]][1]
name=strsplit(NamesNexus[i]," ")[[1]][2]
name=strsplit(name,"'")[[1]][2]
Names[i,1]=name
Names[i,2]=number
}
TaxaNotInCycle=setdiff(as.numeric(Names[,2]),taxset)
if (!length(TaxaNotInCycle)==0)
{taxset=c(taxset,TaxaNotInCycle)
}

#Transform data from the splits block to manipulable, array form:

SplitsNexus=strsplit(SplitsNexus,"\t")
SplitsArray=matrix(0,length(SplitsNexus),3)
M=matrix(0,length(SplitsNexus),length(taxset))
for (i in 1:length(SplitsNexus))
{tmp=SplitsNexus[[i]]
sa1=strsplit(tmp[1],"[=]")[[1]][2]
sa1=strsplit(sa1,"]")[[1]][1]
sa2=as.numeric(tmp[2])
sa3=unlist(strsplit(tmp[3],"[ ,]"))
sa3=as.numeric(sa3[sa3>0])
SplitsArray[i,1]=as.numeric(sa1)
SplitsArray[i,2]=sa2
tmp2=intersect(taxset,sa3)
SplitsArray[i,3]=length(tmp2)
M[i,tmp2]=1
}

#Read in and allocate probability of extinction values:
```

```
Probext=read.table(ProbExtinctionFile)
HED=as.vector(matrix(0,1,length(taxset)))
n=length(HED)
p=rep(0,n)
for (i in 1:n)
{tmp=Names[i,1]
p[i]=as.numeric(Probext[grep(tmp,Probext[,1]),2])
}

#Compute HED values.
#The first taxa:

HED=as.vector(matrix(0,1,length(taxset)))
for (i in 1:(dim(M)[1]))
{
if (M[i,taxset[1]]==1)
{tmp=which(M[i,]==1)
S=prod(p[tmp])/p[taxset[1]]
S_=prod(p[-tmp])
HED[1]=HED[1]+S*(1-S_)*SplitsArray[i,2]
} else {
tmp=which(M[i,]==1)
S_=prod(p[tmp])
S=prod(p[-tmp])/p[taxset[1]]
HED[1]=HED[1]+S*(1-S_)*SplitsArray[i,2]
}
}

#All other taxa:

for (i in 2:n)
{factmult=(p[taxset[(i-1)]]/p[taxset[i]])
HED[i]=HED[(i-1)]*factmult
enter=which((M[,taxset[(i-1)]]-M[,taxset[i]])==1)
leave=which((M[,taxset[(i-1)]]-M[,taxset[i]])==-1)

if (length(enter)==0)
{
} else {
for(j in 1:length(enter))
{tmp=which(M[enter[j],]==1)
S=prod(p[-tmp])/p[taxset[i]]
S_=prod(p[tmp])
a=S*(1-S_)*SplitsArray[enter[j],2]
S_=prod(p[-tmp])
S=prod(p[tmp])/p[taxset[(i-1)]]
b=S*(1-S_)*SplitsArray[enter[j],2]
HED[i]=HED[i]+a-b*factmult
}
}

if (length(leave)==0)
{
} else {
for(k in 1:length(leave))
{tmp=which(M[leave[k],]==1)
S=prod(p[-tmp])/p[taxset[(i-1)]]
S_=prod(p[tmp])
a=S*(1-S_)*SplitsArray[leave[k],2]
S_=prod(p[-tmp])
S=prod(p[tmp])/p[taxset[i]]
```

```
b=S*(1-S_)*SplitsArray[leave[k],2]
HED[i]=HED[i]+b-a*factmult
}
}
}

#Normalize and post-process:

HED=HED/length(taxset)

HED=as.matrix(HED[order(taxset)])
rownames(HED)=Names[,1]

tmp1=names(HED[order(HED,decreasing=TRUE),])
HED=matrix(HED[order(HED,decreasing=TRUE)],n,1)
rownames(HED)=tmp1

HED
}

#-------------------------------------------------------------------------
#Run the script on the data file:
```

Appendix H.

Pairwise genetic distances for spotted owl populations

Pairwise genetic distances (Φ_{ST}) for spotted owl (*Strix occidentalis*) populations based on data from Barrowclough *et al*. (2005, 2006), with negative values revised to zero.

Publication Note

The material in this appendix was previously published in *PLoS ONE* (see Chapter 3 for full citation).

	Zion	Reef	Mant	Cann	Flag	Coco	Mogo	SanF	Tula	Pino	Blac	Sacr	Cata	Rinc	Rita	Huac
Zion	—															
Reef	0	—														
Mant	0.5496	0.6018	—													
Cann	0.1869	0.2808	0	—												
Flag	0.0503	0.1529	0.0737	0	—											
Coco	0.4154	0.4511	0.0027	0.0118	0.0109	—										
Mogo	0.2280	0.2851	0.0557	0	0	0	—									
SanF	0.3219	0.3821	0.0509	0	0	0.0498	0.0041	—								
Tula	0.3453	0.3658	0.0122	0	0	0	0	0	—							
Pino	0.3460	0.3919	0.0597	0	0	0	0	0	0	—						
Blac	0.4880	0.5250	0.0792	0.0186	0.1237	0.0292	0	0.0966	0	0.0529	—					
Sacr	0.4622	0.4995	0.1921	0.0309	0.1870	0.2264	0.1721	0.1077	0.1656	0.0707	0.2152	—				
Cata	0.3697	0.4436	0	0	0	0	0	0	0	0	0	0.1594	—			
Rinc	0.6521	0.6769	0.1817	0.2033	0.3044	0.0203	0.1387	0.1687	0.0346	0.1419	0.0629	0.3069	0.0262	—		
Rita	0.3145	0.3684	0.0137	0	0	0.0618	0	0	0	0	0.0333	0	0	0.2006	—	
Huac	0.6929	0.7193	0.5522	0.2705	0.4368	0.2273	0.2468	0.3099	0.1880	0.3750	0.1862	0.3233	0.2529	0.4382	0.1873	—
Pina	0.6171	0.6540	0.2113	0.0500	0.2222	0	0.0295	0.0829	0	0.0687	0	0.2233	0	0	0.0877	0.4764
Zaca	0.5591	0.6298	0	0	0	0	0	0	0	0.0455	0	0.0496	0	0.4000	0	0.6000
Agua	0.5367	0.6047	0.4286	0	0.1429	0.3697	0.1830	0.1159	0.2214	0.2222	0.3885	0.0091	0.3125	0.7273	0	0.8788
Olym	0.8192	0.8320	0.8127	0.7510	0.7826	0.8108	0.7711	0.7738	0.7737	0.7966	0.7990	0.7749	0.8021	0.8565	0.7648	0.8588
Wena	0.8426	0.8531	0.8486	0.7809	0.8128	0.8269	0.7935	0.7981	0.7901	0.8257	0.8221	0.7969	0.8301	0.8805	0.7935	0.8850
Will	0.7966	0.8080	0.7807	0.7263	0.7572	0.7819	0.7453	0.7474	0.7466	0.7657	0.7673	0.7468	0.7713	0.8176	0.7395	0.8217
Shas	0.7532	0.7670	0.7203	0.6728	0.7042	0.7454	0.7012	0.7007	0.7096	0.7126	0.7217	0.7016	0.7197	0.7689	0.6865	0.7700
Humb	0.7755	0.7828	0.7600	0.7203	0.7423	0.7646	0.7356	0.7365	0.7366	0.7448	0.7500	0.7374	0.7524	0.7830	0.7303	0.7880
Mari	0.8570	0.8670	0.8734	0.7848	0.8265	0.8279	0.7956	0.8014	0.7872	0.8412	0.8283	0.7972	0.8433	0.8995	0.8012	0.9121
Lass	0.5061	0.5309	0.5104	0.3973	0.4153	0.5495	0.4553	0.4577	0.4902	0.4523	0.5394	0.5095	0.4906	0.6145	0.4127	0.6196
ElDo	0.7634	0.7680	0.8615	0.7089	0.7387	0.7278	0.6955	0.7118	0.6691	0.7633	0.7775	0.7382	0.7832	0.8626	0.7237	0.8913
Sequ	0.7829	0.7859	0.8864	0.7360	0.7659	0.7451	0.7168	0.7355	0.6874	0.7909	0.7971	0.7585	0.8078	0.8824	0.7509	0.9115
Carm	0.8205	0.8220	0.9611	0.7590	0.8122	0.7528	0.7273	0.7492	0.6864	0.8388	0.8169	0.7650	0.8494	0.9322	0.7799	0.9765
SanB	0.8601	0.8576	0.9738	0.8198	0.8614	0.7892	0.7803	0.8016	0.7297	0.8792	0.8552	0.8122	0.8875	0.9476	0.8363	0.9830
SanJ	0.8601	0.8576	0.9738	0.8198	0.8614	0.7892	0.7803	0.8016	0.7297	0.8792	0.8552	0.8122	0.8875	0.9476	0.8363	0.9830
Palo	0.7972	0.8019	0.9517	0.7208	0.7808	0.7341	0.6976	0.7191	0.6643	0.8137	0.7950	0.7384	0.8257	0.9231	0.7443	0.9724

Pina	Zaca	Agua	Olym	Wena	Will	Shas	Humb	Mari	Lass	ElDo	Sequ	Carm	SanB	SanJ	Palo
—															
0.5000	—														
0.8182	1	—													
0.8406	0.7965	0.8020	—												
0.8719	0.8395	0.8454	0	—											
0.8015	0.7589	0.7663	0.1332	0.0732	—										
0.7450	0.6919	0.6939	0.1546	0.1199	0.0279	—									
0.7705	0.7359	0.7431	0.1323	0.0493	0	0.0091	—								
0.8990	0.8811	0.8796	0.2966	0.2795	0.1383	0.1300	0.1186	—							
0.5609	0.4611	0.4252	0.6739	0.6936	0.6511	0.6063	0.6463	0.6942	—						
0.8727	0.8844	0.8730	0.8749	0.8934	0.8421	0.7991	0.7983	0.9115	0.0661	—					
0.8967	0.9098	0.9011	0.8852	0.9035	0.8523	0.8098	0.8070	0.9225	0.0970	0.0209	—				
0.9732	1	1	0.8933	0.9165	0.8549	0.8056	0.8042	0.9449	0.0857	0.0742	0.0019	—			
0.9810	1	1	0.9148	0.9337	0.8781	0.8345	0.8237	0.9576	0.1381	0.1169	0.0357	0	—		
0.9810	1	1	0.9148	0.9337	0.8781	0.8345	0.8237	0.9576	0.1381	0.1169	0.0357	0	0	—	
0.9680	1	1	0.8812	0.9069	0.8427	0.7908	0.7948	0.9374	0.0578	0.0507	0	0	0	0	—