Are hotspots of evolutionary potential adequately protected in southern California?

Amy G. Vandergast\textsuperscript{a,}, Andrew J. Bohonak\textsuperscript{b}, Stacie A. Hathaway\textsuperscript{a}, Joshua Boys\textsuperscript{b}, Robert N. Fisher\textsuperscript{a}

\textsuperscript{a}US Geological Survey, Western Ecological Research Center, San Diego Field Station. 4165 Spruance Road, Suite 200, San Diego, CA 92101, United States
\textsuperscript{b}San Diego State University, Department of Biology, 5500 Campanile Drive, San Diego, CA 92182-4614, United States

A R T I C L E I N F O

Article history:
Received 20 December 2007
Received in revised form 3 April 2008
Accepted 9 April 2008
Available online 2 June 2008

Keywords:
Biogeography
Conservation genetics
Comparative phylogeography
Geographic Information Systems
Reserve design

A B S T R A C T

Reserves are often designed to protect rare habitats, or “typical” exemplars of ecoregions and geomorphic provinces. This approach focuses on current patterns of organismal and ecosystem-level biodiversity, but typically ignores the evolutionary processes that control the gain and loss of biodiversity at these and other levels (e.g., genetic, ecological). In order to include evolutionary processes in conservation planning efforts, their spatial components must first be identified and mapped. We describe a GIS-based approach for explicitly mapping patterns of genetic divergence and diversity for multiple species (a “multi-species genetic landscape”). Using this approach, we analyzed mitochondrial DNA datasets from 21 vertebrate and invertebrate species in southern California to identify areas with common phylogeographic breaks and high intrapopulation diversity. The result is an evolutionary framework for southern California within which patterns of genetic diversity can be analyzed in the context of historical processes, future evolutionary potential and current reserve design. Our multi-species genetic landscapes pinpoint six hotspots where interpopulation genetic divergence is consistently high, five evolutionary hotspots within which genetic connectivity is high, and three hotspots where intrapopulation genetic diversity is high. These 14 hotspots can be grouped into eight geographic areas, of which five largely are unprotected at this time. The multi-species genetic landscape approach may provide an avenue to readily incorporate measures of evolutionary process into GIS-based systematic conservation assessment and land-use planning.

Published by Elsevier Ltd.

1. Introduction

Across a variety of biological subdisciplines, there is growing recognition that conservation should aim to protect not only key species and populations (i.e., patterns of species diversity), but also the evolutionary processes that create and sus-
tain these patterns (Frankel, 1974; Erwin, 1991; Crandall et al., 2000; Myers and Knoll, 2001; Moritz, 2002). In any given re-
gion, biodiversity conservation nearly always begins with land protection. Biological, social and political factors all weigh into the reserve selection process, and limits to funding and land availability restrict the number of reserves that can...
ultimately be established. Nonetheless, conservation efforts must look beyond the establishment of local parks and land trusts to integrative plans at the level of large landscapes (Noss, 1983; Ferrier, 2002). Although geographical patterns of species and ecosystem level diversity are often considered in conservation assessments (Kiester et al., 1996; Williams et al., 1996; Myers et al., 2000; Orme et al., 2005), rarely have the underlying processes per se been examined in this framework (but see Cowling et al., 2003). Further, criteria that maximize existing diversity may not adequately preserve the underlying processes responsible for these patterns (Moritz, 2002; Rissler et al., 2006). Spatial components of the evolutionary processes that generate and maintain biodiversity must be identified, mapped and presented in an easily understood framework prior to their incorporation into conservation plans (Prendergast et al., 1999; Cowling and Pressey, 2001).

As a first approximation, the potential for adaptive and vicariant evolution may be estimated by examining the spatial distribution of genetic variability within species. Maintenance of adequate genetic variability ultimately facilitates these evolutionary processes as the biotic and abiotic environments change. For inferences concerning multiple taxa in a given region, comparative phylogenetic and population genetic approaches can reveal the geologic and ecological processes that have been most important in shaping contemporary genetic and species diversity patterns, presumably providing insight for the future (Avise, 1992; Humphries et al., 1995; Moritz, 1995; Moritz and Faith, 1998; Riddle and Hafner, 2006).

California has been widely recognized as a biodiversity and endangerment “hotspot” (Wilson, 1992; Dobson et al., 1997; Myers et al., 2000), and numerous studies have focused on the evolutionary mechanisms driving species diversity patterns in this portion of North America (Raven and Axelrod, 1978; Yanev, 1980; Wake, 1997; Calsbeek et al., 2003; LaPointe and Rissler, 2005). However, evolutionary processes have not been explicitly incorporated into local or statewide reserve planning within California, despite their consideration in conservation plans elsewhere (Forest et al., 2007). The California State Parks system has recently initiated efforts to identify regions of high evolutionary potential in order to help prioritize future land acquisition (University of California, 2006; Davis et al., 2008). As in most other recreational and reserve systems, the California State Parks system has historically endeavored to acquire and protect areas that are exemplars of the ecoregions and geomorphic provinces of California, as well as particularly rare habitat types. As a conservation strategy, this approach implicitly seeks to minimize the loss of current species diversity due to extinction. However, at least some of the evolutionary processes responsible for creating and maintaining this diversity are more likely to occur in areas that can be thought of as hotspots of evolutionary potential (sensu Myers, 1990; Myers et al., 2000; Moritz, 2002).

With an emphasis here on spatial patterns of intraspecific genetic variation (rather than species richness), evolutionary potential can be characterized in terms of interpopulation genetic divergence and intrapopulation genetic variation. Geographic areas that display high divergence among populations are likely to harbor great evolutionary potential, as they typically reflect abiotic drivers of adaptive variation (e.g., an ecotone of steep environmental gradients) or areas in which divergent lineages now meet (suture zones). Less certain is whether areas with exceptionally low interpopulation divergence (or high genetic connectivity) may also reflect source populations of high evolutionary potential. However, one could argue that such areas may be poised for rapid evolutionary change if they reflect recent range expansions for multiple taxa, or a genetic “corridor” of high connectivity historically. In terms of intrapopulation genetic variation, it is widely appreciated that high levels of genetic variation provide a rich source for evolutionary change, whether this is due to large effective population sizes, the presence of an historical refuge, or a zone of mixing between previously isolated gene pools. In the absence of other mitigating factors, low evolutionary potential would be expected from areas that are depauperate of genetic variation across multiple taxa. Thus, we operationally define evolutionary hotspots as areas with high evolutionary potential because, across multiple taxa, they display exceptionally high levels of (1) genetic divergence, (2) genetic diversity, and to a lesser extent, (3) high genetic connectivity.

Some types of evolutionary hotspots are found where ecoregions or habitat types intergrade, and so they are not necessarily represented in the current set of lands that are preserved by governmental and private organizations. Limits to available lands and funding make it impossible to protect all areas that might fall into these categories. This is particularly true in southern California, where variable topography and climate create many potential evolutionary hotspots. Therefore, spatially explicit studies of genetic diversity in multiple lineages (i.e., comparative phylogeography) are needed to identify and prioritize areas that are the most important in shaping regional patterns of species diversity (Calsbeek et al., 2003; LaPointe and Rissler, 2005).

We have developed a method for mapping patterns of genetic divergence and diversity in a Geographic Information Systems (GIS) framework which we designate the multispecies genetic landscape. Our goals were to establish a spatially explicit evolutionary framework for southern California, and to evaluate whether current land reserves adequately protect areas with high evolutionary potential. Our focus was primarily on the southern California coastal ecoregion (Fig. 1), which is characterized by high floral and faunal diversity both above and below the species level (Stebbins and Major, 1965; Myers et al., 2000; Rissler et al., 2006). We compiled and standardized mitochondrial DNA datasets from 21 species that contained adequate sampling within the southern California coastal ecoregion to characterize geographic patterns of intraspecific genetic diversity. Joint analyses of the resulting genetic landscapes identified eight geographic areas where multiple species show atypical patterns of interpopulation divergence or intrapopulation diversity. These evolutionary hotspots were then spatially evaluated in terms of their current protection status (sensu Scott et al., 1993; Davis, 1994).

2. Methods

Twenty-one mtDNA sequence datasets for different animal taxa occurring in southern California were gathered from our previous work, published studies, and unpublished studies (Table 1). In selecting studies, we attempted to maxi-
mize taxonomic breadth while limiting inclusion to studies that sampled at least seven locations in southern California, covering most or all of the taxon’s range within southern California, and where data were available directly from study authors. The 21 species that met the criteria for this study are likely to provide a relatively unbiased sample of genetic diversity patterns in animals across the southern California coastal ecoregion (Table 1). They include invertebrates, amphibians, reptiles, birds, and mammals, and span a wide ecological breadth. For each species, authors provided sequences, population information, and the distribution of haplotypes among collection sites. Collection location coordinates were provided, gathered from the primary studies, or estimated using locality descriptions. The genetic data were extensively validated, standardized, and reformatted for Arlequin v. 3.1 (Excoffier et al., 2006).

We exclusively utilized datasets consisting of mtDNA sequences. Although there are limitations inherent in the use of any single gene to study evolutionary and demographic history (Hare, 2001; Ballard and Whitlock, 2004), the utility of mtDNA data for comparative phylogeographic inference is well-established (Avise, 1992, 2000; Cook et al., 2006), the mitochondrial genome may be of particular importance for reproductive isolation (Ellison and Burton, 2008), and the overwhelming majority of published and available datasets consist of mitochondrial sequence data. Accordingly, we felt it would be most informative to analyze as many taxa as possible using a single marker, relying upon multiple taxa (rather than multiple genes) to cross-validate observed patterns. This focus on the taxon as the unit of replication is consistent with our primary interest in regional history rather than any single species history. As a greater number of studies of nuclear genetic variation become available, these could be incorporated into future multi-species genetic landscapes.

2.1. Estimating genetic divergence

We operationally treated the collection locations defined by the study author as gene pools (populations) for landscape-level visualizations of genetic divergence. In most cases, large distances separated these locations, and we did not perceive that any putative population represented multiple gene pools. In some cases, multiple collection locations in close proximity may correspond to a single gene pool. However, we did not combine any locations for landscape-level visualizations of genetic divergence, because (1) we were interested in mapping regions of high genetic similarity as well as high divergence, and pooling samples would obscure these patterns, and (2) divergence patterns will be more accurately detected from multiple point locations scattered throughout a landscape than from fewer pooled samples each representing larger geographic areas.

In Arlequin, the genetic divergence between each pair of collection locations was calculated as

\[ D_{ij} = \frac{\hat{p}_{ij} - \frac{\hat{p}_i + \hat{p}_j}{2}}{\frac{\hat{p}_i + \hat{p}_j}{2}} \]

where the average divergence (in base pairs) between individuals from populations \( i \) and \( j \) is corrected for the average divergence within \( i \) and within \( j \) (\( \hat{s} \) and \( \hat{s} \), respectively; Nei and Li, 1979). To provide some correction for differing amounts of evolution across data sets, we used a Kimura (1980) 2-parameter model of evolution for these calculations. While we recognize that other models of nucleotide evolution may provide a better fit to some datasets, model selection is less...
important for intraspecific comparisons where overall genetic variability is lower than it would be if we were resolving deeper nodes where saturation is more likely (Kkelchner and Thomas, 2007). In light of this, we chose to use a single model to simplify calculations for multiple datasets. These genetic distances were visualized as genetic landscapes in a Geographic Information System (ArcGIS 9.1), inspired by the methods of Miller (2005). Euclidean distances were calculated between all pairs of locations, and genetic distance was regressed against (Euclidean) geographic distance using reduced major axis regression with the software IBDWS (Jensen et al., 2005). The residuals from these regressions were used to interpolate a genetic landscape for each species in ArcGIS 9.1 as follows. A triangular irregular network was first constructed from the locations. The network connects all collection locations to their nearest neighbors with non-overlapping edges, forming irregularly distributed triangles. Residual $D_A$ values were mapped to the geographic midpoints between

<table>
<thead>
<tr>
<th>Species</th>
<th>Hotspots of high divergence</th>
<th>Hotspots of high connectivity</th>
<th>High variance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Invertebrates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apomastus schlingeri/</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kristeneae$^1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branchinecta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sandiegensis$^3$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemileuca electr$^a$</td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Stenopelmatus</td>
<td></td>
<td>D/C</td>
<td></td>
</tr>
<tr>
<td>&quot;mahogany&quot;$^a$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Amphibians</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batrachocep$^5$</td>
<td>+/-</td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td>nigrientris$^5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batrachoceps major$^6$</td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td><strong>Birds</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chamaea fasciata$^7$</td>
<td>D</td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td>Picoide albolavras$^8$</td>
<td>+/-</td>
<td>+/-</td>
<td>D</td>
</tr>
<tr>
<td>Toxostoma redivivum$^9$</td>
<td></td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td><strong>Mammals</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neotoma fuscipes$^{10}$</td>
<td>+/-</td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td>Neotoma lepida$^{11}$</td>
<td></td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td>Perognatus longimembris$^{12}$</td>
<td></td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td>Sorex ornatus$^{13}$</td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td><strong>Reptiles</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emys marmorata$^{14}$</td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Eumeces gilberti$^{15}$</td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Lampropelitis zonata$^{16}$</td>
<td></td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td>Lichinura trivirgata$^{17}$</td>
<td></td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td>Masticophis flagellum$^{18}$</td>
<td></td>
<td>D</td>
<td>D/C</td>
</tr>
<tr>
<td>Phrynosoma coronatum$^{19}$</td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Sceloporus occidentalis$^{20}$</td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Xantusia henshawi$^{21}$</td>
<td></td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>


Areas with mixed results are labeled (D/C), and those without strong signal (< 1 SD away from mean) are indicated with (+/-). Underlined entries indicate species with complete surface coverage across the area of interest.
locations along the edges of the network. By using residuals in these analyses, we were able to focus on regions of unusually high genetic divergence, after correcting for the accumulation of genetic differences between increasingly distant locations. Finally, a surface was interpolated from the midpoint coverage using Inverse Distance Weighted interpolation (power = 2, variable search radius with 12 points, grid cell size 1 km$^2$). To avoid extrapolating beyond the original collection locations, individual species surfaces were clipped to the extent of the original triangular irregular network or to the boundaries of the southern California coastal ecoregion in cases where this encompassed a smaller spatial extent. Individual species surfaces are presented in Fig. S1 (Supplementary material).

The genetic landscapes for all taxa were averaged into a single multi-species genetic landscape to highlight areas of congruence. To assure that each species received equal weighting in the multi-species genetic landscape, we tested two different rescaling techniques. The “scaled divergence multi-species genetic landscape” was calculated by dividing each $D_{ab}$ within a dataset by the maximum $D_{ab}$ found in that data set. Thus, the maximum divergence for a species was scaled to 1.0 before interpolating the surface. The resulting scaled multi-species genetic landscape was an average of the 21 individual species surfaces. To determine whether dissimilar skew and kurtosis across data sets might bias our results, we also analyzed a “rank-scaled multi-species genetic landscape” by rescaling the raw $D_{ab}$ values using ranks: $D_{ab}(rank – scaled) = (rank(D_{ab}) – 1)/n – 1$, where $n$ is the number of $D_{ab}$ values within the data set. We present here only the scaled divergence multi-species genetic landscape, because although general conclusions from the two approaches were similar, variance in the rank-scaled multi-species genetic landscape was 20–70 times higher.

Because the spatial coverage of individual species’ genetic landscapes was not always equal, the number of species represented in each 1 km$^2$ grid cell of the multi-species genetic landscape also varied. To assess multi-species concordance, we clipped each multi-species genetic landscape extent to areas with coverage for three or more species. Based on the range of values present, we considered the most divergent areas in the multi-species genetic landscape to be greater than 1.5 standard deviations from the mean, and the least divergent areas to be less than 1.5 standard deviations from the mean. Finally, a coverage of the sample variance, which represents the dispersion of the individual species’ genetic landscape values from the average multi-species genetic landscape, was calculated using the standard formula $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$.

### 2.2. Estimating within-site genetic diversity

To estimate genetic diversity within each collection location, we calculated the average sequence divergence among individuals as $\pi$, under the Kimura (1980) 2-parameter model of evolution. This statistic is less biased by differences in sample size among populations than others such as the number of alleles or segregating sites. Fourteen of the original 21 data sets contained multiple collection locations with two or more individuals and these were used for this analysis. Because the genetic diversity calculations are more dependent on sample size than the divergence calculations, we pooled collection locations that were less than five km apart in species datasets where individual sample sizes were low (1–3 individuals/location). For the diversity calculations, seven species (Hemileuca electra, Lichinura trivirgata, Perognathus longimembris, Phrynosoma coronatum, Poides albolarvatus, Scoloporus occidentalis and Toxostoma redivivum) contained 2–10 pooled samples comprised 2–5 sampling locations each. The average distance among non-pooled sites in these species ranged from 114 to 219 km. Using Inverse Distance Weighted interpolation as described above, we interpolated the genetic diversity surface for each species (Supplementary material, Fig. S2) and calculated the average diversity multi-species genetic landscape and variance surface for all 14 datasets. In contrast to the genetic divergence multi-species genetic landscape, surface interpolation for the diversity multi-species genetic landscape was conducted around the actual collection locations, rather than the midpoints between them.

### 2.3. Protection of hotspots

To determine whether hotspots were adequately protected under current land conservation efforts, the average genetic landscapes were overlaid with a protected lands layer developed from a combination of sources. Protected open space data were provided by the Southern California Open Space Council and GreenInfo Network (updated in 2005 for Santa Barbara and Ventura Counties and in 2007 for Los Angeles, San Bernardino, Riverside, Orange, and San Diego Counties). To include the most current data available and to incorporate military bases that are considered here as important stewards of natural resources, these sources were combined with Public and Conservation Trust Lands for all counties in the southern California coastal ecoregion (California Resources Agency Legacy Project, data developed under contract by VESTRA Resources, Inc. 2005). These represent areas considered to have protection from conversion from natural land cover. By examining the multi-species genetic landscapes in relation to these protected lands we were able to perform a gap analysis (Scott et al., 1993) assessing how well areas of atypical genetic divergence and diversity are currently protected.

### 3. Results

#### 3.1. Divergence and diversity multi-species genetic landscapes

The scaled divergence multi-species genetic landscape encompassed 38,118 km$^2$ after clipping to the ecoregion boundary and to overlapping coverage for at least three species. The mean divergence value was 0.53 (SD = 0.05) across the 1 km$^2$ grid cells and ranged from 0.36 to 0.69 (Fig. 2a). Sample variance ranged from 0.00 to 0.08 with a mean of 0.03 (Fig. 2b).

The scaled diversity multi-species genetic landscape covered 35,807 km$^2$ when clipped using the same criteria as the divergence multi-species genetic landscape (Table 2 and Fig. 3). The mean diversity was 0.23 (SD = 0.07) and ranged from
Fig. 2 – (a) Multi-species genetic landscape for genetic divergence across the southern California coastal ecoregion. Genetic divergence ranges from its highest values in red to its lowest values in green. Regions of extremely high or low divergence are circled and abbreviated as follows: SP = Sierra Pelona Mountains; SG = San Gabriel Mountains; SBe = San Bernardino Mountains; WS = Warner Springs; NWT = northwestern Transverse Range; LAB = Los Angeles basin; PR = southern Peninsular Range. (b) The corresponding variance surface of the multi-species genetic landscape.
0.02 to 0.61 (Fig. 3a). The sample variance for the diversity multi-species genetic landscape ranged from 0.00 to 0.27 with a mean of 0.05 (Fig. 3b). Based on these interpolations, we identified areas of evolutionary potential where interpopulation divergence was unusually high or low, or where intrapopulation diversity was unusually high. The divergence multi-species genetic landscape highlighted six hotspots of high genetic divergence, five of high genetic connectivity and one region of unusually high variance (Fig. 2). Because this area of high variance in the southern Peninsular Range contained exceptionally high divergence for five species and exceptionally high genetic similarity for five others (Table 1), we considered it to be both a divergence and connectivity hotspot. The genetic diversity multi-species genetic landscape highlighted three hotspots of high genetic diversity within populations (Fig. 3). Two of these overlapped with areas highlighted in the divergence multi-species genetic landscape. The northwest portion of the diversity multi-species genetic landscape also had high variance, due to the fact that only three species had coverage there. Hotspots were grouped into eight regions for further discussion: the northwestern Transverse Range, the central Transverse Range, the Sierra Pelona, San Gabriel Mountains, San Bernardino Mountains, Los Angeles Basin, Warner Springs, and the southern Peninsular Range. Individual species contributing to each region are summarized in Tables 1 and 2.

<table>
<thead>
<tr>
<th>Species</th>
<th># of collection locations</th>
<th>Mean within pop. sample size</th>
<th>Hotspots of high diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invertebrates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apomastus schlingeri/kristenae</td>
<td>22</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Branchinecta sandiegonensis</td>
<td>73</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Hemileuca electrora</td>
<td>6</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Stenopelmatus &quot;mahogany&quot;</td>
<td>29</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Birds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chamaea fasciata</td>
<td>5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Picoideis albolarvatus</td>
<td>8</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Toxostoma redivivum</td>
<td>10</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Mammals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neotoma fuscipes</td>
<td>12</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Neotoma lepida</td>
<td>22</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Perognathus longimembris</td>
<td>13</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Sorex ornatus</td>
<td>7</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Reptiles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lichinura trivirgata</td>
<td>20</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Phrynosoma coronatum</td>
<td>14</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Sceloporus occidentalis</td>
<td>33</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
| H denotes high diversity, L denotes low diversity, other notation follows Table 1. Collection locations containing a single individual were excluded from the diversity analysis. Underlined entries indicate species with complete surface coverage across the area of interest.

3.2. Protected coverage of multi-species genetic landscapes

Approximately fifty percent of the land in the divergence multi-species genetic landscape coverage area is currently protected in some manner from future land conversion. Although less than 50% of the middle portion of the divergence distribution is protected, the tails of the distribution are somewhat overrepresented (Fig. 4). For example, 70% of the most divergent 1% of grid cells in the genetic divergence map is protected, and 84% of the most divergent 0.1% is protected. Likewise, 65% of the least divergent 1%, and 87% of the least divergent 0.1% is protected. These areas generally correspond to state and federal lands in mountainous regions. Patterns of protection were markedly different for the genetic diversity layer. The likelihood of protection increases almost linearly as a function of average genetic diversity, from 20% protection in the least diverse 1% of grid cells, to over 84% in the most genetically diverse 1% (Fig. 5). Our examination of the spatial overlay of protected lands reveals at least five regions where evolutionary hotspots fall mainly outside of protected lands: the Sierra Pelona, the San Bernardino Mountains, the Los Angeles Basin, Warner Springs and the southern Peninsular Range (Fig. 4a and Fig. 5a).

4. Discussion

4.1. Interpreting evolutionary hotspots

As described below, we interpret hotspots of high genetic divergence, high genetic connectivity and high genetic diversity as areas with high evolutionary potential.

4.1.1. Divergence hotspots

Because all genetic divergence hotspots are located in regions with steep environmental gradients (e.g., steep ecotones between mountains and basins), high mtDNA divergence is likely to correlate with high divergence for traits under natural selection. In many cases, these areas also represent secondary contact zones where previously allopatric lineages
of multiple species hybridize, termed “suture zones” (Stebbins and Major, 1965; Remington, 1968). Hybridization between divergent lineages creates novel gene combinations that can facilitate speciation and adaptive evolution in some cases (Barton et al., 1983; Barton and Hewitt, 1985; Turgeon et al., 1999; Ebert et al., 2002; Tallmon et al., 2004; Lavergne

Fig. 3 – (a) Multi-species genetic landscape for within-site genetic diversity. Genetic diversity ranges from its highest values in red to its lowest values in green. Three areas with extremely diversity are circled and labeled as follows: CT = central Transverse Range; SBe = San Bernardino Mountains; PR = southern Peninsular Range. (b) The corresponding variance surface of the multi-species genetic landscape.
Some species in our data set may be rare or even absent at the exact location that hotspots are mapped, since divergence is mapped to midpoints between sampling locations, rather than the sampling locations themselves. However, these areas may still have evolutionary and conservation potential for two reasons. First, gaps in a species’ distribution may be due to partial or complete barriers to establishment across climatic gradients that encompass the extremes of suitable habitat space. The isolation of peripheral populations in situations such as these has been widely hypothesized to promote evolutionary diversification through various combinations of genetic drift, natural selection and other microevolutionary processes (Mayr, 1963; Raven, 1964; Carson and Templeton, 1984; Templeton, 1996; Gavrilets and Boake 1998; Mooers et al., 1999; Dawson and Hamner, 2005). Second, the adaptive potential stored in these evolutionary hotspot areas is likely to become increasingly important as climatic conditions change in the future, and
these environmental barriers intensify, weaken or shift spatially (Etterson, 2004; Hellmann and Pineda-Krch, 2007).

4.1.2. Connectivity hotspots

In contrast to hotspots of divergence, areas of high genetic similarity across the landscape often represent relatively recent and rapid range expansions, where there has not yet been time for random drift to accumulate differences among populations. One might expect a high potential for current and future evolutionary change when multiple species have recently colonized a particular area. Alternatively, connectivity hotspots may reflect high ongoing rates of gene flow because of few ecological and topographic barriers to movement. Although the ecological importance of such corridors has long been appreciated, their relevance for genetic connectivity is considered much less often (Mech and Hallett, 2001; Horskins et al., 2006). When hotspots of high genetic connectivity are separated by large distances and intervening areas of high divergence, each may represent an independent, regional assemblage of gene pools for multiple species. Therefore, these areas are likely to represent unique reservoirs of genetic diversity that are irreplaceable and analogous.

Fig. 5 – (a) Overlay of protected lands on the diversity multi-species genetic landscape. (b) Protected status as a function of average genetic diversity, visualized as a LOWESS smoother with 10% span. (c) Histogram showing distribution of average genetic diversity values.
to distinct population segments within species (U. S. Fish and Wildlife Service and National Marine Fisheries Service, 1996). Such areas may warrant separate protection or coverage in the network of conservation lands, including regional Habitat Conservation Plans.

4.1.3. Diversity hotspots

Finally, higher genetic diversity may be expected in areas that contain older populations, as mutational pressures drive gene pools towards their equilibrium state for genetic diversity. Higher diversity may also be expected in areas where gene pools from formerly separated lineages now meet (see above).

One caveat to interpretation of the divergence multi-species genetic landscape is that divergence values are mapped to the midpoints between collection locations, and are thus dependent on the dispersion of collection locations across the landscape. When individual species datasets vary in sampling density (as with those compiled for this study), undoubtedly, additional error in the spatial location of hotspots is introduced. In order to avoid drawing false conclusions from this, we took care to examine each individual species’ genetic landscape with regard to collection locations and species’ biology before interpreting patterns of genetic divergence. We also restrict our interpretation of hotspots to broad regions of interest, rather than attempting to pinpoint exact locations.

4.2. Specific regions of interest

Many of the evolutionary hotspots that we identified in this study have been highlighted in previous work. For example, all but one of the divergence hotspots are located in the southern Transverse Range (Sierra Pelona, San Gabriel, and San Bernardino Mountains), an area that has previously been recognized as containing high lineage diversity (Ma tocq, 2002; Rissler et al., 2006; Chatzimanolis and Caterino, 2007), and vascular plant endemism (Munz, 1935; Raven and Axelrod, 1978). Additionally, Remington (1968) listed 25 animal and plant species known or suspected to hybridize across the extreme climatic gradient from the xeric desert slope to the mesic Pacific slope of the southern Transverse Range. Remington (1968) noted that contact sites throughout the Transverse Ranges tended to be localized at passes in the mountains and at the gaps between successive ranges. Our detection of divergence and diversity hotspots in these major passes (Sierra Pelona, San Gabriel, San Bernardino, San Bernardino, San Bernardino) is concordant with these findings. Some of the high genetic connectivity hotspots that we detected have also been previously identified in historical biogeographic studies. The northwestern Transverse Range connectivity hotspot was recognized by Peabody and Savage (1958). They hypothesized that concomitant retreat of a marine strait through the southern Coast Ranges, mountain uplift, and cooler climate during the Pleistocene created a “Coast Range corridor” which facilitated genetic exchange between the Coast and Transverse Ranges and allowed range expansion of Arcto-Tertiary species into southern California. Additionally, the southern Peninsular Range, connectivity hotspot lies within a band of high relictual plant species concentration that runs along the eastern slope of the San Jacinto and Santa Rosa Mountains into the Borrego Valley (Stebbins and Major, 1965). Raven and Axelrod (1978) hypothesized that comparatively high summer precipitation throughout these mountains has allowed the persistence of many species throughout this region that had wider ranges during the moister, cooler parts of the Quaternary. Although these studies examined distributional patterns rather than genetic divergence or diversity, the striking concordance between our findings and theirs suggests that significant evolutionary processes are occurring in the areas we have highlighted. Additional biogeographic patterns and potential explanatory factors contributing to each hotspot are presented in Table 3.

4.3. Gap analysis of protected lands

When spatial concordance in processes such as range expansion, lineage divergence and re-contact occurs among species, it suggests that protection of these regions may aid in preserving the evolutionary potential of the faunal assemblage. Because (1) our divergence multi-species genetic landscape maps the midpoints among sample locations, and more importantly, (2) evolutionary function may operate over broad spatial scales, we believe the optimal approach to protect these functions is to restore natural levels of connectivity across broad areas that include the regions of interest (Templeton et al., 2001; Hunter et al., 2003). Accordingly, we highlight five regions that may require further protection.

4.3.1. Sierra Pelona

This divergence hotspot almost completely is unprotected (Fig. 4a). It is bounded to the northeast by the largely urbanized San Fernando Valley and to the southwest by the rapidly growing city of Palmdale in the Antelope Valley, and is bisected by Highway 14, which connects them. However, some areas in the pass between the Sierra Pelona and San Gabriel Mountains are still relatively undeveloped. For example, upland areas around the headwaters of the Santa Clara River have recently been interpreted as a narrow mesic corridor that could provide a connection between populations of invertebrates on either side (Chatzimanolis and Caterino, 2007).

4.3.2. San Bernardino Mountains

Although areas within the San Bernardino and San Jacinto Mountains are well protected, the divergence and diversity hotspots at the San Gorgonio Pass between them are unprotected (San Bernardino, Fig. 4a). Interstate Highway 10 constitutes a major transportation corridor through this pass, so that efforts to protect this hotspot may require restoring connectivity across this barrier (Ernest et al., 2003; Hunter et al., 2003).

4.3.3. Los Angeles Basin

The Los Angeles basin contains both a hotspot of divergence and a region of high connectivity for several taxa. The southern polygon in the San Gabriel Mountains hotspot (San Gabriel) lies in the midst of urban Los Angeles. Additionally,
<table>
<thead>
<tr>
<th>Hotspot</th>
<th>Description</th>
<th>Regional features and historical biogeography</th>
<th>Concordance from other studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWT</td>
<td>High connectivity along an east–west region bounded by the Tehachapi Mts (east) and Santa Ynez Mts (west), and a north–south corridor bounded by the Southern Coast Ranges (north) and Western Transverse Ranges (south)</td>
<td>High gene flow and range expansion of Arcto-Tertiary species into southern CA during the Pleistocene, due to marine waters retreating, mountain uplift, and cooler climate. Relatively low abiotic complexity compared to elsewhere in the So. Cal. Coast. Ecorgn.: primarily low hills (&lt;1500 m), no nearby desert, and climatic buffering by the Pacific Ocean.</td>
<td>Sepedophillus castaneus²</td>
</tr>
<tr>
<td>CT</td>
<td>High diversity within the Transverse Range bounded by the San Emigdio Mts (north), Santa Clara River valley (south), and the Tehachapi Mts (northeast)</td>
<td>High habitat diversity: sits at the junction between five of the major ecoregions in CA², is geologically complex, and includes three major faults that were active during the Pleistocene. Possible point of origin for southern CA lineages of some species. Speciation hotspot for plants².</td>
<td>Vascular plants⁵</td>
</tr>
<tr>
<td>SP</td>
<td>High divergence at the juncture between the Sierra Pelona and San Gabriel Mts and the Desert and California Floristic Provinces</td>
<td>Barriers include Santa Clara River currently, and a Pliocene marine embayment (ca. 5–2.5 MYA⁶). Mountain uplift across this divide beginning ca. 3 MYA⁷.</td>
<td>Diadophis punctatus⁸, Sepedophillus castaneus²</td>
</tr>
<tr>
<td>SG₁</td>
<td>High divergence along the northern edge of the mountain range between lowland/desert taxa in the Mojave Desert and populations west of the mountains. For other species, high divergence between the San Gabriel and San Bernardino Mts</td>
<td>San Andreas Fault on the western edge of the San Bernardino Mts, through the Cajon Pass to the northern edge of the San Gabriel Mts. Land west of the rift zone has been steadily moving north over the last 12 MY⁹. Land movement along fault zones may impact salamanders¹⁰,¹¹. Additional factors may be more important in more mobile species: mountain uplift across this divide beginning ca. 3 MYA⁷, or more recent Pleistocene climate shifts (e.g., Chaemaea fasciata¹²).</td>
<td>Multiple species: contact zones in the pacific slope suture zone concentrated in low lying mountain passes¹³</td>
</tr>
<tr>
<td>SG₂</td>
<td>High divergence in along the southern edge of the San Gabriel Mts between populations within the San Gabriel and/or San Bernardino Mts, and the Santa Ana Mts to the southwest</td>
<td>Population fragmentation in lower elevations due to intermittent marine incursions (most recently during the Pleistocene and early Holocene), and contemporary widespread urbanization¹⁴.</td>
<td>Same as SG2 above</td>
</tr>
<tr>
<td>SB₁</td>
<td>High divergence in the San Bernardino Mts west of Big Bear Lake, reflecting divergence between the San Bernardino and Santa Ana Mts</td>
<td>Low elevation San Gorgonio Pass bordered by the highest peaks in southern CA (San Gorgonio Mt.: 3505 m; San Jacinto Mt.: 3293 m) represents one of the most extreme elevation and habitat clines in southern CA. Lies at the juncture of four floristic subregions from two floristic provinces³. Regional uplift beginning ca. 3 MYA⁷, and numerous faults (including the San Andreas Fault). Intermittent flooding from the mid-Pliocene through the Quaternary by the Sea of Cortez and the Los Angeles Basin⁶,¹⁵. More recent fragmentation by Interstate Route 10. Numerous subspecies/highly divergent clades meet here¹⁶,¹⁷,¹⁸.</td>
<td>Multiple species: contact zones in the pacific slope suture zone concentrated in low lying mountain passes¹³</td>
</tr>
<tr>
<td>SB₂,₃</td>
<td>High divergence between lowland distributed species on desert and coastal sides of the San Bernardino Mts. For montane species, high divergence between the San Bernardino and San Jacinto Mts across the low elevation, xeric San Gorgonio Pass. Also, high intrapopulation genetic diversity in the San Gorgonio Pass</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
the most significant deficit in connectivity hotspot protection (i.e., the largest dip in Fig. 4b) corresponds largely to the Los Angeles Basin, in which undeveloped lands occur mainly as small fragments. Due to high levels of urbanization, there is likely little chance to further protect this region. However, opportunities to protect the unique types here or restore connectivity among existing reserves may still exist, particularly along riparian corridors (e.g., Santa Ana River).

### Table 3 – continued

<table>
<thead>
<tr>
<th>Hotspot</th>
<th>Description</th>
<th>Regional features and historical biogeography</th>
<th>Concordance from other studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAB₁,₂,₃</td>
<td>High connectivity within parts of Los Angeles and northern Orange Co.</td>
<td>Contained within a single floral assemblage (South Coast floristic subregion) and almost uniformly low elevation. Historic marine inundations, with the most recent sizeable incursion occurring ca. 100,000 years ago. Pleistocene habitat fragmentation, post-isolation expansion and recent anthropogenic fragmentation are differentially expressed in each taxon. For example, genetic homogeneity in most species (recent range expansions) contrasts with high divergence in <em>Sorex ornatus</em> (Supplementary material, Fig. S1)</td>
<td></td>
</tr>
<tr>
<td>WS</td>
<td>Very high divergence south of the San Jacinto and Santa Rosa Mts, and northwest of Clark and Borrego Valleys. Differentiation in some species between the Santa Rosa Mts to the east and populations to the west, complemented in other species by a north/south break near the headwaters of the San Luis Rey and Santa Margarita rivers</td>
<td>The hotspot generally sits between two mountain ranges, but it is topographically heterogeneous with altitudes exceeding 2500 m on many peaks. Strong climatic differences along a coast-to-mountain gradient; uplift of the Peninsular Range created a permanent rain shadow across much of southern California by ca. 2 MYA. Annual precipitation drops from 50 cm to &lt;15 cm in the deserts immediately east of this hotspot. The most divergent areas within the hotspot are strikingly coincident with local increases in precipitation to &gt;60 cm annually. Also a tectonically complex area of the San Jacinto fault zone</td>
<td></td>
</tr>
<tr>
<td>PR₁</td>
<td>High connectivity across the Fish Creek Mts (northeast) spanning Anza Borrego Desert to the Tierra Blanca Mts (southwest)</td>
<td>For desert-distributed taxa, similarity extends through the Colorado Desert (although unmapped in Fig. 2). For other species, relics from cooler, wetter Pleistocene cycles persist due to comparatively high warm season precipitation.</td>
<td>High relictual plant species diversity</td>
</tr>
<tr>
<td>PR₂,₃</td>
<td>High variance in divergence in the southern Peninsular Range in San Diego Co., bounded by the South Coast floral subregion on the west and the Laguna Mts to the east. High genetic diversity is found southwest of connectivity hotspot PR₁, overlapping slightly with high variance area PR₂ to the west</td>
<td>Abiotic environment may be more homogeneous than other hotspots of high divergence, and variation among species near the border with Mexico reflects widely divergent evolutionary histories. Some taxa have few barriers to genetic exchange in the region. For others, secondary contact zones between previously isolated lineages occurs north of the hotspot (<em>Phrynosoma coronatum</em>), or within it (<em>Lichinura trivirgata</em>, <em>Masticophis flagellum</em>; Online Supplementary material, Fig. S1)</td>
<td>High relictual plant species diversity</td>
</tr>
</tbody>
</table>

4.3.4. Warner Springs

This hotspot is the largest within our divergence multi-species genetic landscape, and sits at the crossroads between several mountain ranges and the desert. Based on its biogeographic role (Table 3) and the sampling coverage in this hotspot, landscape connectivity could be extended through this hotspot from the foothills of the San Bernardino Mountains to Palomar Mountain and the Santa Ana Range. Approximately 50% of the land in this region is not protected, with the largest contiguous unprotected section occurring in the northern portion of the hotspot (Fig. 4a). This corresponds to the foothills and valleys just south of the San Bernardino National Forest. Much of this land falls within the East County Multi-species Conservation Plan in San Diego County (www.dfg.ca.gov/habcon/nccp/status.html) and may be protected through that planning effort.

4.3.5. Southern Peninsular Range

Although approximately half of the area in the southern Peninsular Range’s high variance region (Peninsular Range) is protected in the Cleveland National Forest, the most genetically diverse portion of diversity hotspot Peninsular Range is almost entirely unprotected (Fig. 5a). This area at the south end of the Laguna Mountains contains largely rural, unincorporated communities and tribal lands, and is bounded to the west by encroaching suburban development. Finally, although our analysis is limited to areas north of the U.S. Mexico border, hotspots of diversity and divergence likely extend to both sides. The unique evolutionary potential represented by the fauna in this region is a natural resource that requires binational attention despite the accompanying political challenges (White et al., 2006).

4.4. Conclusions and future directions

Our synthesis of genetic studies encompassing disparate taxa into two multi-species genetic landscapes explicitly is spatial in nature. Thus, it allows for a visual means of determining geographical areas of high evolutionary potential without relying on spatial surrogates, or overlaying predetermined location labels on phylogenetic trees. Spatial concordance among intraspecific genetic patterns allows us to identify regions where similar evolutionary processes have occurred repeatedly in the past, and presumably this is indicative of high potential for future evolutionary change as well. Most of the hotspots fall in higher elevation areas along the edges of the southern California coastal ecoregion that are already well protected. With the exception of the Los Angeles Basin, areas that are not currently well protected are (to the best of our knowledge) relatively undeveloped, but may be threatened by future development pressures (e.g., exurban development: Radeloff et al., 2005). This encouraging finding suggests that (at least for the animals studied here) there is still potential to maintain processes such as adaptation and lineage diversification with proper reserve planning. Admittedly, our gap analysis of currently protected lands is only an initial step to identify regions where further protection may be warranted. The consideration of evolutionary hotspots for reserve design should complement other criteria (e.g., species richness and diversity, rare species or lineages), as advocated by Williams et al. (1996). Spatial methods of systematic conservation assessment incorporating measures of conservation value (e.g., species diversity, complimentarity, etc.) and costs (e.g., land availability, restoration costs) have been developed (Pressey, 1999; Ball and Possingham, 2000; Possingham et al., 2000; Cowling et al., 2003) and some detailed modeling using these tools is underway in southern California (Beier et al., 2006). With proper interpretation, the multi-species genetic landscape approach may provide an avenue to readily incorporate some measure of evolutionary process into these GIS-based assessment and planning activities.

Acknowledgements

We thank Kevin Burns, Marshal Hedin, Elizabeth Jockusch, Rob Lovich, Jesus Maldonado, Marjorie Matocq, Daniel Rubinoff, Jonathan Richmond, Javier Rodriguez-Robles, and Jonathan Spinks for providing us with their published data. We especially thank Jim Archie, Adam Leaché, Milan Mitrovich, Jim Patton and Dustin Wood for sharing unpublished data. Rick Rayburn, Craig Moritz and others organized and led initial workshops that helped to motivate this work. Tod Reeder and Janet Franklin provided comments that much improved an earlier draft of this manuscript. This work was supported by the U.S. Geological Survey Western Ecological Research Center, California Department of Parks and Recreation, and NSF grant DBI-0204447. Use of trade names does not imply U.S. Geological Survey endorsement.

Appendix A. Supplementary material


REFERENCES


Hall, C.A., 2002. Nearshore marine paleoclimatic regions, increasing zoogeographic provinciality, molluscan extinctions, and paleoshores, California: late Oligocene (27 MA) to late Pliocene (2.5 MA), Geological Society of America, Special Paper.


