Molecular Ecology (2004) **13**, 1265–1275 doi: 10.1111/j.1365-294X.2004.02110.x

Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history

BENJAMIN N. SACKS,* SARAH K. BROWN* and HOLLY B. ERNEST*†
*Wildlife and Ecology Unit, Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, Davis, CA 95616-8744, USA, †Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, USA

Abstract

Little is known about the relationship between animal movements and the emergent structure of populations, especially for species occupying large continuous distributions. Some such mammals disperse disproportionately into habitat similar to their natal habitat, a behavioural bias that might be expected to lead to habitat-conforming genetic structure. We hypothesized that coyotes (*Canis latrans*) would exhibit such natal-biased dispersal, and used 13 microsatellite loci to test, correspondingly, whether genetic structure conformed to major habitat breaks. First, we used a model-based approach to assign coyote genotypes to distinct genetic clusters irrespective of geographical location. Visualization on a geographical information system revealed a strong concordance between the locations of cluster assignments and habitat bioregions, not explainable in terms of physical dispersal barriers or intervening low-quality habitat. Next, we used a multiple Mantel test, which controlled for effects of geographical distance (itself, marginally significant; $P = 0.06$), to statistically determine that genetic distance was indeed higher between than within bioregions ($P < 0.001$). Whereas previously published examples of landscape effects on gene flow have typically been explainable in terms of species-wide habitat affinities or dispersal barriers, our finding that genetic subdivisions were associated with unobstructed boundaries between contiguous habitats suggests a role for intraspecific variability in habitat affinities as a factor underlying genetic structure. In addition, our data combined with previously published data suggest a pattern of genetic isolation-by-distance throughout western North America, consistent with independent evidence that the western half of the coyote range predates European settlement.

Keywords: *Canis latrans*, gene flow, genetic structure, habitat, isolation-by-distance, microsatellites

Received 23 September 2003; revision received 2 December 2003; accepted 2 December 2003

Introduction

How animals disperse across the landscape is an important determinant of the genetic structure of their populations (Lidicker & Patton 1987). Beyond this generality, little is known about the relationship between movements of individual animals, as observable on a fine scale through field study, and the emergent genetic structure of populations over larger scales. This is particularly true of habitat generalists with continuous distributions. Although such populations might be expected to exhibit simple patterns of genetic isolation-by-distance (Slatkin 1993), perhaps altered to account for obvious dispersal barriers (e.g. large waterways), it is also possible that more complex patterns of genetic structure arise due to heterogeneous, individual-specific habitat selection behaviour. Effects of habitat heterogeneity on genetic structure have been considered in the context of species-specific habitat affinities (e.g. Keyghobadi *et al*. 1999; Manel *et al*. 2003), but the role of ontogenetic factors, which can lead to intraspecific variability in habitat selection (e.g. 'habitat imprinting'; Vogl *et al*. 2002), has received little attention.

Field studies of some vertebrate species suggest that individuals tend to disperse preferentially to habitat similar to that in their natal home range even when the species

Correspondence: Benjamin N. Sacks. Fax: + 1-530-752-3556; E-mail: bnsacks@ucdavis.edu

© 2004 Blackwell Publishing Ltd

overall has broad habitat affinities (Wecker 1963; Olson & Van Horne 1998; Vogl *et al*. 2002). It is easy to imagine how such behaviour, in principle, could reduce gene flow between habitat types relative to that within habitat types, although the visible emergence of genetic structure from this behaviour would require the tendency to be strong within the species. Also, such structuring should be most evident where habitat types were discrete and differed in multiple ways (e.g. vegetation, topography, prey communities) increasing the likelihood that most individuals would perceive habitat boundaries similarly. Although there are only a few examples in the literature of such natal habitat-biased dispersal, this may reflect the paucity of longitudinal studies rather than rarity of the trait. To the extent that a species' survival depends on learning and experience, such behaviour would seem to be highly beneficial and therefore could be widespread among vertebrates.

Our primary objective was to test a prediction of the natal habitat-biased dispersal hypothesis in coyotes (*Canis latrans*) from central California, specifically, that genetic structure conforms to major habitat divisions. We took advantage of California's diverse landscape, which divides conveniently into several discrete bioregions on an appropriate scale for such a study of coyotes. We sampled coyotes primarily from four contiguous bioregions: the Sierra Nevada (SN), Northwestern (NW), Central Western (CW), and Great Valley (GV) (described by Hickman 1993). The first three bioregions are mountainous, whereas the GV is almost completely flat. The SN extends to 4418 m elevation, characterized by vegetation mosaics of oak woodland, grassland and chaparral at lower to medium elevations, and coniferous forests and alpine vegetation at higher elevations. This bioregion is separated from the other two mountainous bioregions by the GV, composed largely of agricultural cropland and grassland, punctuated with marsh and riparian vegetation. The other two mountainous bioregions (NW, CW) are composed of low-elevation mountains and foothills with similar habitats to the low to medium elevations of the SN. The border between these two similar coastal bioregions (as defined in this study) corresponded to the San Francisco Bay Estuary, undoubtedly an impenetrable barrier to gene flow. Coyotes were distributed continuously throughout these bioregions (including all vegetation types).

Previous studies of coyote population genetics based on mitochondrial DNA (mtDNA) and nuclear microsatellites indicated high gene flow throughout the range of the coyote and found no evidence of genetic isolation-by-distance (Lehman & Wayne 1991; Roy *et al*. 1994). This finding is a little surprising given that coyote dispersal is quite limited relative to the extent of the current geographical range (Bekoff 1982), most of which (e.g. the western extent) predates European exploration and, based on fossil evidence, may date back to the Pleistocene (Dobie 1949; Jackson 1951; Young 1951; Nowak 1978, 1979; Schmidt 1991). One possible explanation for this finding in the previous studies could be that the range of distances between sampling locations was too narrow and distances too great to detect isolation-by-distance. Isolation-by-distance should be easier to detect (i.e. due to greater statistical power) over larger ranges of distances, or with ranges composed of shorter separation distances (e.g. Forbes & Boyd 1997). Furthermore, the pattern may not occur at greater distances if genetic drift and gene flow have not had sufficient time to equilibrate, which could require several hundred thousand generations (Slatkin 1993). A second objective of our study, then, was to test the hypothesis of genetic isolation-by-distance.

Materials and methods

Field sampling

Coyote specimens (n = 457) were obtained from trappers employed by the U.S. Department of Agriculture/Wildlife Services and the Santa Clara County Vector Control District. Coyotes were obtained by wildlife specialists as part of livestock depredation control and public health programmes and were not killed for research purposes. Coyote carcasses (n = 337) were kept frozen until necropsy when a muscle specimen was removed. Blood specimens from additional coyotes (n = 120) were collected in the field on filter paper strips (Nobuto Blood Filter Strips®, Advantec Manufacturing Inc.), allowed to dry, and stored at room temperature in manila envelopes. Locations of coyotes were described by trappers in the field in terms of distances and directions from nearby landmarks (usually < 10 km) and translated to spatial coordinates by us. Coyotes in diverse habitats of California have similarly sized territories, ~5 km^2, and transient coyotes use substantially larger areas (Shivik 1995; Sacks *et al*. 1999; Riley *et al*. 2003). Therefore, the accuracy of the recorded specimen locations should have been adequate relative to the scale of coyote space use.

Microsatellite genotyping

We extracted DNA from muscle specimens using the DNeasy® tissue kit (Qiagen Inc.) and for dried blood specimens collected on filter paper strips, used a DNA extraction service (Lucy Whittier Molecular and Diagnostic Core Facility, University of California, Davis, CA), which used the following protocol. Two dried blood spots (49 mm^2 each) from each strip were digested in 20 μg proteinase K (Invitrogen) and 150 μL NucPrep digestion buffer (Applied Biosystems) and incubated for 60 min at 56 °C, after which 500 μL of NucPrep DNA purification solution

was added and mixed. A 6700 automated nucleic acid workstation (Applied Biosystems) was used to extract DNA from the tissue lysates.

Individuals were initially genotyped at 14 microsatellite loci (*FH2001*, *FH2004*, *FH2010*, *FH2054*, *FH2079*, *FH2088*, *FH2096*, *FH2100*, *FH2161*, *FH2289*, *FH2328*, *FH2380*, *FH2441* and *FH2457*; Breen *et al*. 2001). Six and eight loci were genotyped simultaneously in each of two multiplex polymerase chain reactions (PCR). However, one locus (*FH2441*) exhibited a 36% heterozygote deficiency, suggesting the presence of a null allele, and was therefore excluded from all analyses (i.e. 13 loci were used in analyses). Two pairs of loci occurred on the same chromosome: *FH2004* and *FH2096* (CFA 11) and *FH2010* and *FH2079* (CFA 24); the other nine loci were on chromosomes not shared with any other locus used in this study. The total PCR volume was 17 μL, including 0.34 mM dNTPs, 4.3 mM $MgCl_2$, 0.7 U *Taq* DNA polymerase, 1.7 μg bovine serum albumin, 1× PCR buffer, 2 μL template DNA, and primers in concentrations ranging 0.05–0.60 μM. Forward primers were fluorescently labelled (6-FAM, VIC, NED; Applied Biosystems). PCRs were carried out in PTC-100 Peltier thermal cyclers (MJ Research), with the following PCR profile: 95 °C for 10 min, 85 °C for 10 min, 33 cycles of 95 °C for 1 min, 64 °C for 30 s and 72 °C for 45 s, and a final 30 min extension at 72 °C. The PCR products were visualized using an ABI 377® automated sequencer with GENESCAN ANALYSIS 3.1® in conjunction with an internal size standard, GeneScan 500 LIZ (Applied Biosystems). Alleles were scored using STRAND v2.2.30 (Veterinary Genetics Laboratory, University of California, Davis).

Genotype-based inferences about population structure

We used program STRUCTURE v2.0 (Pritchard *et al*. 2000), a Bayesian model-based approach to detecting genetic structure, which uses discontinuities in multilocus allele profiles (e.g. detected in terms of Hardy–Weinberg equilibrium and linkage disequilibrium) to assign each multilocus genotype to a genetic cluster. Because geographical information is not used to make the assignments, assignments can be visualized on a map and used as an unbiased indicator of geographical patterns of population structure. Because we expected substantial gene flow, we used the population admixture model, which assigns to each genotype a probability of membership in each cluster. We defined a cluster assignment as that for which the membership assignment probability was greatest. We also presented cluster assignments for the subset of coyotes with cluster assignment probabilities ≥ 80% to illuminate geographical patterns of structure while reducing clutter due to admixture. The 80% cut-off was arbitrary and was selected for this data set to optimize the trade-off between sample size and cluster assignment confidence. Cluster assignments were displayed in reference to the four bioregions described above and major highways.

The number of clusters (K) must be specified by the user and, for situations with high gene flow, should reflect the minimum value that does not sacrifice substantial explanatory power (Pritchard & Wen 2002). We used two metrics of explanatory power, a strictly statistical one recommended by the authors, log $Pr(X|K)$ statistic (Pritchard *et al*. 2000), and a 'geographical index' developed by us, calculated as the average geographical distance between locations within clusters divided by the average distance between locations irrespective of cluster. While the log $Pr(X|K)$ statistic is useful in terms of maximizing the genetic distance between clusters, it indicates nothing about the reasonableness of clusters in a geographical sense. In contrast, the geographical index gives an indication of how well genetic clusters correspond to spatial clusters. In the case of no correspondence (e.g. under panmixia), the expected value for the geographical index is 1, because the expected average geographical distance between locations within clusters equals the average distance between all locations (irrespective of cluster). Thus, smaller values of the geographical index correspond to greater geographical explanatory power of clusters. The geographical index should be most useful when isolation-by-distance is a component of the structure, but any type of spatial structure should be associated with a reduction in the average geographical distance between within-cluster locations relative to that among all pairs of locations.

Allele-frequency-based hypothesis tests

Although the model-based approach had the distinct advantages of using all data and enabling the genotypic data to suggest patterns of genetic structure free of investigator biases, an important disadvantage of this approach was that it assumed a priori that structure existed. If, to the contrary, allele frequencies varied continuously over space consistent with an isolation-by-distance model (Slatkin 1993), cluster assignments produced by the model could have been misleading (Pritchard *et al*. 2000). Therefore, we needed to evaluate structural hypotheses with respect to the potentially confounding effect of geographical distance.

We did this in terms of partial correlations among pairwise distance matrices calculated from 12 sampling locations (Fig. 1), using a multiple factor Mantel permutation test (Smouse *et al*. 1986) in ARLEQUIN v2.000 (Schneider *et al*. 2000). We addressed the question, is genetic distance greater between sampling locations in different bioregions than between sampling locations in the same bioregion, over and above that expected due to geographical distance? For this analysis, the dependent variable matrix was of genetic distance (expressed as $\hat{F}_{ST}/(1-\hat{F}_{ST})$, Slatkin 1995) and the independent variable matrices were of

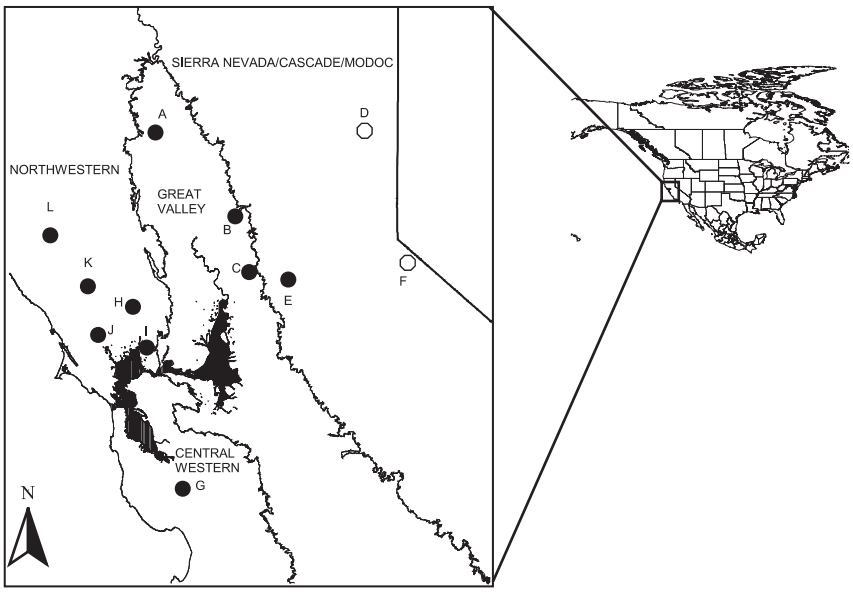

Fig. 1 Locations of 12 coyote sampling locations in central California in four bioregions. Coyotes belong to the putative subspecies *Canis latrans ochropus* (solid circles) and *C. l. lestes* (open circles).

geographical distance (km) and Bioregion. The bioregion matrix was composed of zeros and ones, where 0 indicated two sampling locations in the same bioregion and 1 indicated two sampling locations in different bioregions. Because the partial correlation coefficient in this analysis between genetic distance and bioregion effectively controlled for effects of geographical distance, its significance would support the hypothesis that bioregions accounted for genetic distance over and above that explained by geographical distance alone.

The genetic distance matrix was calculated in ARLEQUIN v2.000. Calculation of the geographical distance matrix involved calculation of centroids (average coordinates among data points) for each sampling location, and then calculation of distances between centroids. Most paths (all but that from sampling location F) to sampling location G were measured via sampling location E (Fig. 1) to account for the San Francisco Bay Estuary (including San Pablo Bay and delta), which were assumed to completely obstruct dispersal. Bioregions were based on those described by Hickman (1993) except that a small portion of the 'Central Western' bioregion that was north of the San Pablo Bay was included with the 'Northwestern' bioregion, and small portions of the 'Cascades' and 'Modoc' bioregions were combined with the 'Sierra Nevada' bioregion. We based bioregions on those described by Hickman (1993) for convenience because they were designated a priori and incorporated coarse vegetative and topographic aspects of habitat, but this did not imply that the particular designations were necessarily the best ones possible.

Although for coyotes, subspecies designations have had few adherents in recent years (Nowak 1978), to be thorough, we tested for a population-genetic basis of putative subspecies in our study region. We did this with the Mantel test as described above except that we substituted a subspecies matrix (0 = same putative subspecies, 1 = different putative subspecies) for the bioregion matrix. Determination of subspecies membership was done according to Grinnell *et al*. (1937).

Hardy–Weinberg and linkage equilibria

ARLEQUIN v2.000 was used to test Hardy–Weinberg equilibrium (Guo & Thompson 1992) and GENEPOP v3.3 was used to test for linkage disequilibrium (Raymond & Rousset 1995) within sampling locations. We considered deviations from Hardy–Weinberg equilibrium and linkage disequilibrium to be significant when $P < 0.05/c$, where c was the number of loci, locus pairs, or populations tested (i.e. the Bonferroni-corrected P-value corresponding to alpha = 0.05; Zar 1999).

Violations of Hardy–Weinberg and linkage equilibrium assumptions can indicate: (i) genotyping errors or selection, likely to be locus specific; (ii) physical linkage, likely to be locus pair specific; or (iii) genetic heterogeneity within sampling locations, which may or may not be sampling-location specific. To assess the first issue, sampling locations were used as replicates to evaluate loci and locus pairs with respect to Hardy–Weinberg equilibrium and linkage equilibrium, respectively. If all loci or locus pairs were equally prone to occasionally deviate from Hardy–Weinberg or linkage equilibrium, respectively (e.g. due to nonrandom mating), we would expect the frequency of such deviations to vary according to a Poisson distribution, which we tested using a Chi-square goodness-of-fit test (Zar 1999). Significant lack of fit of to a Poisson distribution

would suggest that certain loci or locus pairs, respectively, were incorrectly genotyped, under selection, or physically linked. (Locus pairs were not independent and therefore were pseudoreplicates. However, this only increased the likelihood of a type I error, which, in this case, meant incorrectly rejecting the hypothesis that linkage disequlibrium was random with respect to loci. A negative test result would still be valid.)

To assess genetic heterogeneity within sampling locations, Hardy–Weinberg and linkage equilibria were examined across loci and locus pairs. High frequency of Hardy–Weinberg and linkage disequilibria in a sampling location would indicate a nonrandomly breeding group, due either to geographical divisions or generally high inbreeding. We selected three sampling locations with relatively high numbers of deviations from Hardy–Weinberg and linkage equilibria and high sample size to examine for spatial structure. This was done by using STRUCTURE v2.0 as described above except that each sampling location was analysed independently and the number of clusters set at $K = 2$. Also, an inbreeding coefficient (F_{15}) was calculated for the entire population (Hartl & Clark 1989).

Results

Genotype-based inferences about population structure

All loci were polymorphic, with allele counts ranging from 7 to 27 (average 14 alleles). For analysis in STRUCTURE, we chose $K = 4$ clusters, which seemed the highest value justifiable given the data. Beyond $K = 4$, gains in explanatory power were minimal according to both indexes (Fig. 2).

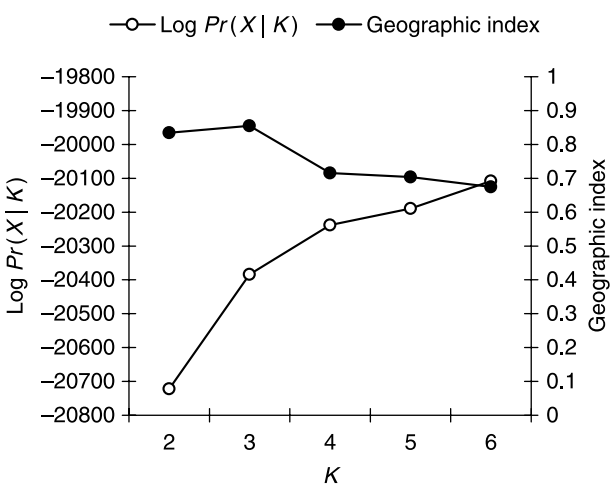

Fig. 2 Two indexes of explanatory power of program STRUCTURE cluster assignments in relation to the number of clusters (K), illustrating that substantial explanatory power (i.e. maximizing log $Pr(X \mid K)$ and minimizing the geographical index) is gained by increasing K to $K = 4$, above which little explanatory power is gained.

Based on the subset of 241 coyotes for which assignment to a genetic cluster was ≥ 80%, there was a clear geographical correspondence (Fig. 3A). Major highways were not consistently associated with genetic breaks but bioregions appeared to be. The pattern was similar when all coyotes ($n = 457$) were examined (Fig. 3B) although, as expected, more admixture was apparent. This was especially evident in the southeastern section of the Northwestern bioregion.

Allele frequency-based hypothesis tests

There were significant univariate correlations between genetic and geographical distance ($r = 0.62$; $P = 0.002$) and between genetic distance and bioregion ($r = 0.70$; $P < 0.001$). Genetic distance between sampling locations was significantly greater for pairs in different bioregions than for pairs in the same bioregion even when geographical distance was accounted for (Fig. 4). Genetic distance between sampling locations did not differ significantly according to whether pairs consisted of same or different putative subspecies ($r_{Y1-2} = -0.25$; $P = 0.92$).

Hardy–Weinberg and linkage equilibria

Among the thirteen loci, four, two and one loci deviated significantly from Hardy–Weinberg equilibrium in one, two, and four sampling locations, respectively, which did not differ significantly from Poisson expectations (? = 0.57, $P = 0.45$). Of the 78 locus pairs, 23, 13, 4 and 2 pairs exhibited significant linkage disequilibrium in 1, 2, 3 and 4 sampling locations, respectively, which also did not differ significantly from Poisson expectations (? = 2.56, $P = 0.46$). Only one instance (in one sampling location) of linkage disequilibrium occurred between loci on the same chromosome (*FH2004*, *FH2096*), further suggesting no physical linkage. Thus, instances of Hardy–Weinberg and linkage disequilibria were approximately random and likely reflected nonrandom mating (e.g. due to inbreeding) rather than locus-specific genotyping errors, selection or physical linkage.

Considerable variability among sampling locations in the frequency of significant Hardy–Weinberg and linkage disequilibria occurred but was at least partly due to differences in sample size (Table 1). However, the heterozygote deficiency, which is unbiased by sample size, was similar across sampling locations. Analysis using STRUCTURE of three sampling locations with relatively high frequency of significant Hardy–Weinberg and linkage disequilibria (B, E, G) indicated no clear geographical correspondence to genetic heterogeneity, suggesting that genetic heterogeneity within sampling locations reflected inbreeding or over-sampling of related individuals. The inbreeding coefficient (?) averaged over the total population was 0.09 (SE = 0.01).

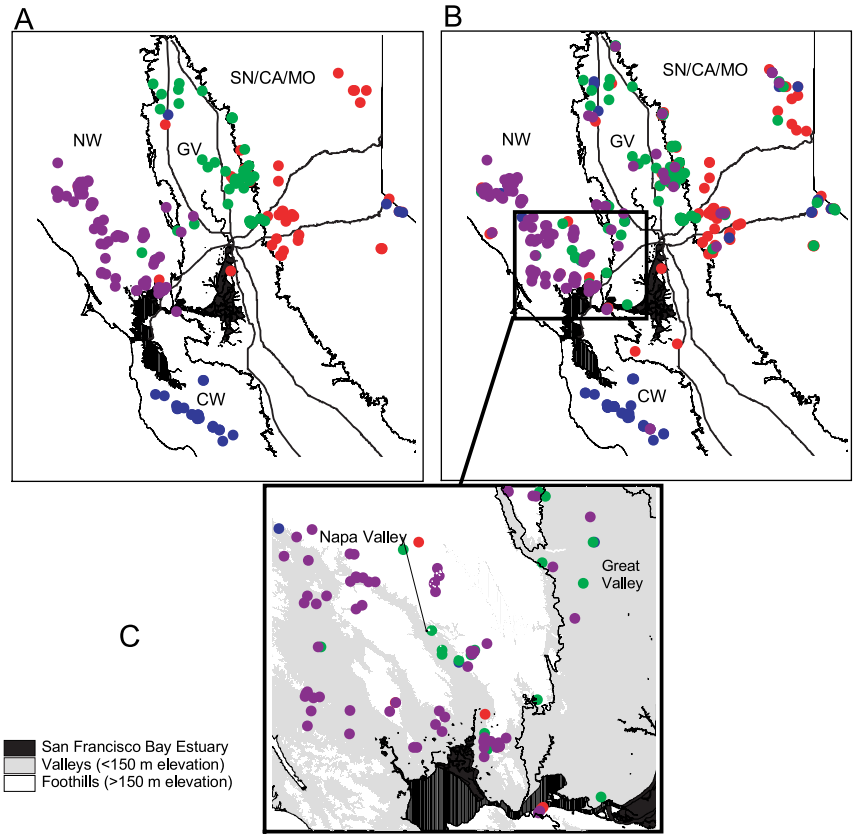

Fig. 3 Assignments using STRUCTURE v2.0 of central California coyotes to four genetic clusters (colour-coded), illustrating habitat-specific breaks corresponding to bioregions (light lines) but not to major highways (dark lines): (A) subset of coyotes with assignment probability > 80% (n = 241); (B) all coyotes (n = 457); (C) blow-up of southern end of Northwestern bioregion showing secondary valleys. Bioregions are abbreviated as follows: Northwestern (NW), Great Valley (GV), Sierra Nevada/Cascade/Modoc (SN/CA/MO) and Central Western (CW).

Within-bioregion heterozygosity

Using sampling locations as sample units, there were no significant differences in expected heterozygosity among the Sierra Nevada/Cascade/Modoc (H_E = 0.76), Great Valley (H_E = 0.74) and Northwestern (H_E = 0.74) bioregions (Kruskal–Wallis H = 1.04, P = 0.59). The Central Western bioregion could not be included in the statistical comparison because it was composed of a single sampling location, although the point estimate was similar to the other bioregions (H_E = 0.73).

Discussion

Habitat-specific genetic breaks

Our findings suggest that coyote populations in central California were subdivided according to macrohabitat breaks. Although actual subdivisions may not correspond perfectly to bioregions as defined in this study (see below), the genetic discontinuities in the coyote population seemed related to habitat affinities rather than to physical dispersal barriers such as major highways, waterways (except for the San Francisco Bay Estuary), or intervening low-quality habitat. Although we cannot definitively rule out the possibility that physical barriers caused the observed subdivisions, it seems highly unlikely. For example, fences are not only highly permeable to coyote movements (Thompson 1978; Sacks 1996), but these are distributed in a fine-grained manner with respect to the

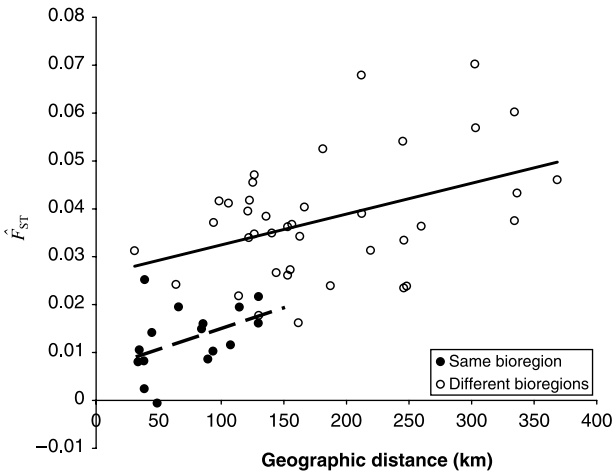

Fig. 4 Relationship between genetic distance and geographical distance of within- vs. between-bioregion coyote sampling location pairs. Regression lines are shown for within-bioregion (dashed) and between-bioregion (solid) distances. The partial correlation between genetic distance and bioregion (r_{Y2-1} = 0.54; P < 0.001) was greater than the partial correlation between genetic distance and geographical distance (r_{Y1-2} = 0.36; P = 0.06).

Table 1 Hardy–Weinberg equilibrium (HWE), linkage disequilibrium (LD), and heterozygote deficiency in 12 sampling locations of variable sample size. Sampling locations are shown in Fig. 1

Sampling location	Proportion of loci out of HWE	Proportion of locus pairs with LD	Average (across loci) $H_E - H_O$	No. coyotes
A	0	0	0.08	14
B	0.31	0.29	0.09	71
C	0	0.01	0.11	27
D	0	0	0.11	19
E	0.15	0.31	0.03	46
F	0	0.14	0.04	19
G	0.23	0.12	0.12	47
H	0	0	0.05	25
I	0	0	0.03	25
J	0.08	0	0.04	21
K	0	0	0.08	19
L	0.08	0.01	0.04	77

study area. The Sacramento River and major highways, respectively, compose the largest waterways and roadways in the study region, none of which coincided with bioregional boundaries or genetic subdivisions (Fig. 3). Finally, although there were some large gaps in our spatial sampling, this was due to our inability to acquire samples from those locations and did not indicate coyote absence or intervening low-quality habitat. In fact, coyotes were distributed approximately continuously throughout the study region. Therefore, any effect of geographical distance on genetic distance should have been adequately controlled for in the Mantel test, which indicated an even stronger effect of macrohabitat than of geographical distance (at least on the scale studied). Our results should nevertheless be confirmed through expanded sampling and through more direct tests of the natal-habitat-biased dispersal hypothesis.

Few studies have examined the effect of habitat on the population structure of broadly distributed terrestrial animals, and these typically have revealed findings explainable in terms of species-wide habitat affinities as opposed to intraspecific differences in habitat affinities, e.g. due to natal-habitat-biased dispersal. For example, one study found that gene flow in alpine butterflies (*Parnassius smintheus*) was reduced through forest vegetation relative to meadow vegetation, due to slower movement rates (Keyghobadi *et al*. 1999). A study of mountain lions (*Puma concolor*) found reduced gene flow between mountain ranges separated by the Great Valley of California relative to gene flow within mountain ranges, presumably due to the rarity of dispersal through the Great Valley (Ernest *et al*. 2003). In both of these cases, gene flow was restricted through an intervening habitat (effectively a partial barrier), which could be most parsimoniously explained in terms of species-wide habitat-specific behaviours or affinities.

However, future investigation could reveal natal-habitat-biased dispersal to be a common factor underlying genetic structure. Some examples in other carnivores support this possibility. Ernest *et al*. (2003) also observed a genetic subdivision associated with the crest of the Sierra Nevada Mountains that was known to be crossable by seasonally migratory individuals from the east (Pierce *et al*. 1999). Similarly, a break in gene flow among wolves (*Canis lupus*) corresponded to a river, which should have been physically passable 6–8 months of the year when it was frozen (Carmichael *et al*. 2001). In both examples, carnivores exhibited migratory behaviour on at least one side of the subdivision associated with a migratory prey population. These examples suggest the possibility of natal-habitat-biased dispersal, in particular, where prey represented the essential aspect of the habitat.

Understanding the causes of genetic subdivision in broadly distributed taxa may provide key insights into the ecological context of past speciation events. To the extent that natal-habitat-biased dispersal (e.g. habitat imprinting) translates to genetic structure, this relationship illustrates how genetic spatial heterogeneity can arise within a species due to an internal (i.e. self-organizing) process. Although genetic subdivisions in our study were clearly not sufficiently deep to reflect reproductive isolation, they illuminate the possibility that natal-habitat-biased dispersal, in concert with other factors and in other species, could contribute to its development. Our findings therefore not only emphasize the potential importance of experience and ontogeny as intraspecific factors affecting genetic structure, but, more generally, suggest another means whereby intraspecific, self-organizing phenomena may underlie speciation (see also Bolnick *et al*. 2003).

Understanding the causes of population structure in continuously distributed species is also important for practical reasons including conservation and epidemiology. For example, there has been much effort to model the spread of zoonoses such as rabies through wildlife populations (e.g. Anderson *et al*. 1981; Childs *et al*. 2000). These models typically assume that reservoir hosts disperse

randomly with respect to direction and habitat. While such assumptions may be of little consequence on continental scales, habitat-specific dispersal tendencies could be very important on local and regional scales, such as when planning efforts to combat budding epidemics. Natal-habitat-biased dispersal also could be important to endangered species reintroduction or translocation efforts, both in terms of source population selection and in predictions about routes of re-colonization.

The importance of habitat as an influence on gene flow likely depends on spatial scale. For example, although highways in this study did not detectably affect genetic structure in central California, generally, it is likely that some stretches of highway in urban areas reduce coyote dispersal more so than habitat boundaries, locally. On a larger scale, such as throughout North America, it is similarly likely that habitat of the grain examined in this study would be less important than distance in determining gene flow, simply because the range of distances would be orders of magnitude larger. However, higher order habitat divisions could be comparably important. For example, pairs of coyote sampling locations on either side of the Rocky Mountains were associated with greater Nei's unbiased genetic distances than were those on the same side of the Rockies (Roy *et al*. 1994), suggesting that the crest of the Rockies might slow gene flow among coyote subpopulations. Rueness *et al*. (2003) made similar observations of lynx (*Lynx canadensis*) with respect to the Rocky Mountains. However, in both cases, any genetic subdivision associated with the crest of the Rockies could have been due to physical constraints on dispersal rather than habitat affinity.

Local patterns of gene flow and recent history

In this study, there was an interesting pattern of admixture in the southern end of the Northwestern bioregion (Fig. 3C). In particular, the secondary valleys, particularly the Napa Valley, contained a disproportionate number of cases of Northwestern bioregion coyotes assigned to the Great Valley bioregion cluster, suggesting that secondary valleys may represent intermediate habitats where mountain coyotes (e.g. Northwestern) and valley coyotes (Great Valley) more readily blend.

Outside the Napa Valley, however, the Northwestern bioregion coyotes appeared genetically distinct from the Great Valley coyotes, which is especially interesting in light of recent history. The abundance of coyotes in the Northwestern bioregion was extremely low relative to neighbouring bioregions before the 1970s, during which coyote numbers increased dramatically in this major sheep-producing region (Ferrell *et al*. 1953; Coolahan 1990; Hackett 1990). This increase was likely related to legal bans in the early 1970s on the widespread use of toxicants to reduce coyote populations, which may have created a coyote vacuum. The genetic distinctiveness of the Northwestern coyotes relative to the Great Valley coyotes suggests that this vacuum was filled by the decedents of coyotes from the Northwestern bioregion rather than from an influx of coyotes from the (then) higher density neighbouring bioregion. If so, this would imply a strong reluctance on the part of valley coyotes to disperse into the mountains. Future genetic analyses using nonrecombining markers would be especially useful in testing this hypothesis and further illuminating the recent history of coyotes in the Northwestern bioregion.

Coyotes in western North America

Early characterizations of the coyote as a 'prairie wolf' and anecdotal accounts of coyotes entering into small densely forested areas after Europeans introduced roads and clear cuts (Grinnell *et al*. 1937; Dobie 1949; Young 1951) have apparently fostered a popular misconception that the pre-European range of the coyote was restricted to the central part of the continent (e.g. Moore & Parker 1992; Parker 1995). While the coyote range clearly has expanded eastward recently, numerous accounts by European explorers indicate that the southwestern most extent of the current coyote range (coastal British Columbia to coastal Mexico) predates European settlement (e.g. Dobie 1949; Jackson 1951; Young 1951; Schmidt 1991) and fossils suggest that it could date back to the early Pleistocene (Nowak 1978, 1979). The ages of the southern-most and northern-most extents of the coyote range are less certain (Dobie 1949; Young 1951), although evidence suggests these too may be pre-European (Jackson 1951; Nowak 1978). Overall, it seems the pre-European coyote range was at least two-thirds its current area (Dobie 1949; Young 1951), suggesting a considerably less pronounced range expansion than is commonly presumed.

Also, our findings are consistent with the pre-European existence of coyotes in western North America. Combining our data with those from a previous, continental-scale microsatellite study of coyotes (Roy *et al*. 1994) provides strong support for the isolation-by-distance pattern overall (Fig. 5). Although different loci were used in the two studies, in principle, different sets of loci should yield similar estimates of *Nm* (number of migrants exchanged per generation). The estimates of *Nm* by Roy *et al*. (1994) were approximately of the magnitude predicted by extrapolation from this study, suggesting that they are consistent with isolation-by-distance despite the high levels of gene flow indicated. If so, this supports the observation that the time since establishment of the western portion of the coyote range was well beyond that since European colonization. For example, after 260 years post expansion, the maximum distance for which genetic isolation-by-distance would be expected to be evident would be 570 km (the log

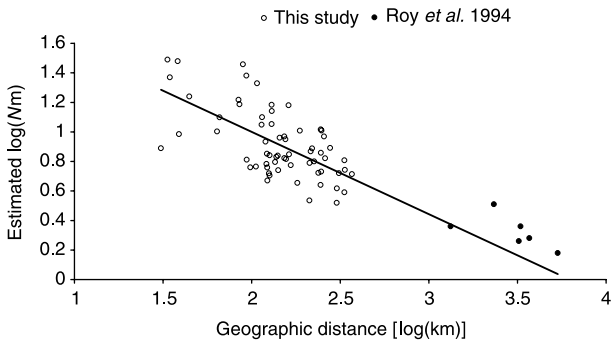

Fig. 5 Relationship between estimated number of coyote migrants per generation (*Nm*) and geographical distance using data from studies on different geographical scales. *Nm* was calculated as per Slatkin (1993) as 0.25*(⁻¹ – 1). Regression line is based on data from the current study. The *y*-intercept of a regression line calculated from combined data sets (not shown) was *Nm* = 87 (R^2 = 0.61). Data from Roy *et al*. (1994) are shown only for pairs of coyote populations in western North America that do not hybridize with wolves (*C. lupus*, *C. rufus*).

of which is 2.75, for comparison to Fig. 5), based on the equation i_{max} = (Slatkin 1993), where *Nm* = the effective number of females in a neighbourhood [literally, the number of migrants per generation at log(distance) = 0], estimated as the y-intercept of the regression shown in Fig. 5 (this study); τ is in units of *Nm* generations (assumed to be 2 years); i_{max} is in units of numbers of neighbourhood lengths, calculated as the square root of the area containing *Nm* females (assuming density = 0.10 breeding females/km²; Bekoff 1982; Sacks 1996). Given the known existence of coyotes along the Pacific coast from fossils as recent as 8000 years ago (Nowak 1979), our findings indicate no reason to doubt the continuous presence of coyotes in the western part of their range since the Pleistocene.

The quantitative results of this study, as well as those of previous studies describing gene flow in coyotes (Lehman & Wayne 1991; Roy *et al*. 1994), may not be as reliable as the qualitative ones. For example, estimates of gene flow between distant populations of coyotes have been reported to be higher than estimates for wolves (Lehman & Wayne 1991; Roy *et al*. 1994), brown bears (*Ursus arctos*; Paetkau *et al*. 1998) and mountain lions (Ernest *et al*. 2003), all larger carnivores with higher vagility. Although this is certainly plausible, for example, reflecting the more continuous range of the coyote relative to the larger carnivores (Wayne 1996), it is also possible that these interspecies comparisons are confounded by differential biases (Forbes & Hogg 1999). Coyotes presently occur at densities 1–2 orders of magnitude higher than those of the larger carnivores (e.g. Mech 1970; Bekoff 1982) and mitochondrial analyses suggest that coyotes historically existed at a considerably higher density than wolves (Vila *et al*. 1999). Genetic divergence due to genetic drift slows with increasing population density, such that estimates of gene flow in nonequilibrium (between migration and drift) populations will be disproportionately overestimated in denser populations (Hartl & Clark 1989). Also, even if populations are at migration–drift equilibrium, microsatellite analyses are susceptible to upwardly biased estimates of gene flow due to homoplasy or 'homogenizing mutation', biases which are expected to be greatest in higher density populations (Nauta & Weissing 1996; Forbes & Hogg 1999). Interestingly, however, mitochondrial analyses give estimates of gene flow in coyotes similar to those estimated using microsatellites (Lehman & Wayne 1991; Roy *et al*. 1994). Future studies using microsatellites from nonrecombining portions of the Y chromosome would be useful in further clarifying the level of gene flow among coyote populations.

Conclusions

Our most significant finding was that coyote genetic structure corresponded to habitat-specific breaks, a pattern expected if coyotes tended to disperse preferentially to habitat similar to their natal habitat. Future population genetic sampling in the study region and more direct tests of natal-habitat-biased dispersal in coyotes are necessary to confirm this putative relationship between behavioural development of individuals and genetic structure of populations in this species. Nevertheless, examples supporting similar relationships in other species, along with the potentially important implications to population ecology and conservation, seem to warrant greater investigation of natal-habitat-biased dispersal, in general, and as a factor underlying population genetic structure, in particular.

The advent of model-based approaches to analysis of genetic structure (such as that used here), which provide a powerful means of uncovering cryptic population structure (Manel *et al*. 2003), should greatly facilitate future investigations of natal-habitat-biased dispersal as a cause of genetic structure. However, we might have achieved the same qualitative results (albeit with less strength) based on the multiple Mantel test alone given that the study region was easily divided a priori into discrete bioregions of the appropriate scale. In part, we were driven to conduct this study because of the existence of the discrete bioregions. However, perceptions of habitat by humans and other species may often disagree, and there may be other cases in which population subdivision arises in a population from natal-habitat-biased dispersal in a way not easy for a human observer to anticipate. This problem might be especially important for smaller animals (e.g. small mammals, herps, invertebrates), which likely perceive their environment on very different scales than humans. In such cases, the best way to determine such structure would be to confront the genetic data free of a priori assumptions about meaningful habitat units.

Acknowledgements

Samples were contributed by researchers, Bruno Chomel, Vicki Kramer and Brian Mitchell, employees of the Santa Clara County Vector Control District, Dairen Simpson and Noor Tietje, and wildlife specialists employed by the USDA/APHIS/Wildlife Services: Gary Abreo, John Chandler, Jim Coleman, Dan Davis, Mark Frederick, Eddie Goymerac, Gary Johnson, Mike Muller, Jody Nicholas, Jim O'Brien, Fred Radkey and Mike Smith. We also thank Monty Slatkin, Bob Wayne, Louis Bernatchez and three anonymous reviewers for helpful comments. Alison Ruhe, Katy Robertson, other members of the Veterinary Genetics Laboratory and Jennifer Leonard provided valuable laboratory methodological advice. We are indebted to Niels Pedersen, who made the research possible. Funding was provided through the Veterinary Genetics Laboratory in the UC Davis School of Veterinary Medicine and grants from the UC Davis, Genetic Resources Conservation Program.

References

Anderson RM, Jackson HC, May RM, Smith AM (1981) Population dynamics of fox rabies in Europe. *Nature*, **289**, 765–771.

Bekoff M (1982) Coyote. In: *Wild Mammals of North America* (eds Chapman JA, Feldhamer GA), pp. 447–459. Johns Hopkins University Press, Baltimore.

Bolnick DI, Svanback R, Fordyce JA *et al.* (2003) The ecology of individuals: incidence and implications of individual specialization. *American Naturalist*, **161**, 1–28.

Breen M, Jouquand S, Renier C *et al.* (2001) Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map if the domestic dog genome to all chromosomes. *Genome Research*, **11**, 1784–1795.

Carmichael LE, Nagy JA, Larter NC, Strobeck C (2001) Prey specialization may influence patterns of gene flow in wolves of the Canadian Northwest. *Molecular Ecology*, **10**, 2787–2798.

Childs JE, Curns AT, Dey ME *et al.* (2000) Predicting the local dynamics of epizootic rabies among raccoons in the United States. *Proceedings of the National Academy of Sciences of the USA*, **97**, 13666–13671.

Coolahan C (1990) The North Coast animal damage control program. In: *Predator Management in North Coastal California* (eds Giusti GA, Timm RM, Schmidt RH), pp. 16–22. University of California, Hopland Field Station Publication 101.

Dobie FJ (1949) *The Voice of the Coyote*. University of Nebraska Press, Lincoln.

Ernest HE, Boyce WM, Bleich VC *et al.* (2003) Genetic structure of mountain lion (*Puma concolor*) populations in California. *Conservation Genetics*, **4**, 353–366.

Ferrell CM, Leach HR, Tillotson DF (1953) Food habits of the coyote in California. *California Fish and Game*, **39**, 301–341.

Forbes SH, Boyd DK (1997) Genetic structure and migration in native and reintroduced rocky mountain wolf populations. *Conservation Biology*, **11**, 1226–1234.

Forbes SH, Hogg JT (1999) Assessing population structure at high levels of differentiation: microsatellite comparisons of bighorn sheep and large carnivores. *Animal Conservation*, **2**, 223–233.

Grinnell J, Dixon JS, Linsdale JM (1937) *Fur Bearing Mammals of California: Their Natural History, Systematics, and Relations to Man*, Vol. II. University of California Press, Berkeley.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hackett D (1990) Predator problems on California's north coast. In: *Predator Management in North Coastal California* (eds Giusti GA, Timm RM, Schmidt RH), pp. 23–27. University of California, Hopland Field Station Publication 101.

Hartl DL, Clark AG (1989) *Principles of Population Genetics*, 2nd edn. Sinauer Associates, Sunderland, MA.

Hickman JC (1993) *The Jepson Manual: Higher Plants of California*. University of California Press, Berkeley.

Jackson HHT (1951) Classification of the races of the coyote. In: *The Clever Coyote* (eds Young SP, Jackson HHT), pp. 229–407. University of Nebraska Press, Lincoln.

Keyghobadi N, Roland J, Strobeck C (1999) Influence of landscape on the population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae). *Molecular Ecology*, **8**, 1481–1495.

Lehman N, Wayne RK (1991) Analysis of coyote mitochondrial DNA genotype frequencies: estimation of effective number of alleles. *Genetics*, **128**, 405–416.

Lidicker WZ, Patton JL (1987) Patterns of dispersal and genetic structure in populations of small rodents. In: *Mammalian Dispersal Patterns: The Effects of Social Structure on Population Genetics* (eds Chepko-Sade BD, Halpin ZT), pp. 144–161. University of Chicago Press, Chicago.

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

Mech LD (1970) *The Wolf: The Ecology and Behavior of an Endangered Species*. University of Minnesota Press, Minneapolis.

Moore GC, Parker GR (1992) Colonization by the eastern coyote (*Canis latrans*). In: *Ecology and Management of the Eastern Coyote* (ed. Boer AH), pp. 23–37. Wildlife Research Unit, University of New Brunswick, Fredericton.

Nauta MJ, Weissing FJ (1996) Constraints on allele size at microsatellite loci: implications for genetic differentiation. *Genetics*, **143**, 1021–1032.

Nowak RM (1978) Evolution and taxonomy of coyotes and related *Canis*. In: *Coyotes: Biology, Behavior, and Management* (ed. Bekoff M), pp. 3–16. Academic Press, New York.

Nowak RM (1979) *North American Quaternary Canis*. Museum of Natural History, University of Kansas, Kansas.

Olson GS, Van Horne B (1998) Dispersal patterns of juvenile Townsend's ground squirrels in southwestern Idaho. *Canadian Journal of Zoology*, **76**, 2084–2089.

Paetkau D, Waits LP, Clarkson PL *et al.* (1998) Variation in genetic diversity across the range of North American brown bears. *Conservation Biology*, **12**, 418–429.

Parker GR (1995) *Eastern Coyote: The Story of its Success*. Nimbus Publishing, Halifax, NS.

Pierce BM, Bleich BC, Wehausen JD, Bowyer RT (1999) Migratory patterns of mountain lions: implications for social regulation and conservation. *Journal of Mammology*, **80**, 986–992.

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

Pritchard JK, Wen W (2002) *Documentation for Structure Software Version 2* http://pritch.bsd.uchicago.edu

Raymond M, Rousset F (1995) GENEPOP Version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Riley SPD, Sauvajot RM, Fuller TK *et al.* (2003) Effects of urbanization and habitat fragmentation on bobcats and coyotes in southern California. *Conservation Biology*, **17**, 566–576.

Roy MS, Geffen E, Smith D, Ostrander EA, Wayne RK (1994) Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution*, **11**, 553–570.

Rueness EK, Stenseth NC, O'Donoghue M, Boutin S, Ellegren H, Jakobsen KS (2003) Ecological and genetic spatial structuring in the Canadian lynx. *Nature*, **425**, 69–72.

Sacks BN (1996) *Ecology and Behavior of Coyotes in Relation to Depredation and Control on a California Sheep Ranch*. MSc thesis, University of California, Berkeley.

Sacks BN, Jaeger MM, Neale JCC, McCullough DR (1999) Territoriality and breeding status of coyotes relative to sheep predation. *Journal of Wildlife Management*, **63**, 593–605.

Schmidt RH (1991) Gray wolves in California: their presence and absence. *California Fish and Game*, **77**, 79–85.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, *Version 2.000: A Software for Population Genetics Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Shivik JA (1995) *Factors Influencing Space Use and Activity of Sagehen Basin Coyotes*. MSc thesis, University of California, Berkeley.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, **35**, 627–632.

Thompson BC (1978) Fence-crossing behavior exhibited by coyotes. *Wildlife Society Bulletin*, **6**, 14–17.

Vila C, Amorim IR, Leonard JA *et al.* (1999) Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. *Molecular Ecology*, **8**, 2089–2103.

Vogl W, Taborsky M, Taborsky B, Teuschl Y, Honza M (2002) Cuckoo females preferentially use specific habitats when searching for host nests. *Animal Behaviour*, **64**, 843–850.

Wayne RK (1996) Conservation genetics in the Canidae. In: *Conservation Genetics: Case Histories from Nature* (eds Avise JC, Hamrick JL), pp. 75–118. Chapman & Hall, New York.

Wecker SC (1963) The role of early experience in habitat selection by the prairie deer mouse, *Peromyscus maniculatus bairdi*. *Ecological Monographs*, **33**, 307–325.

Young SP (1951) Its history, life habits, economic status, and control. In: *The Clever Coyote* (eds Young SP, Jackson HHT), pp. 229–407. University of Nebraska Press, Lincoln.

Zar JH (1999) *Biostatistical Analysis*, 4th edn. Prentice Hall, Upper Saddle River, NJ.

This study is part of Ben Sacks's postdoctoral research on population structure of California canids. This research is being conducted in the Wildlife and Ecology Unit of the Veterinary Genetics Laboratory headed by Holly Ernest and is aimed at understanding how large-scale movement patterns map onto the landscape and to gain insights into behavioural tendencies of individual canids. The research builds on Dr Sacks's MSc and PhD work on canid behavioural ecology, population biology, diseases, and parasites and complements ongoing research in the lab on landscape ecology and genetics and genetics of disease susceptibility of California's wildlife, including black bears, mountain lions, Swainson's hawks and red-tailed hawks. Sarah Brown's MSc research focused on applying population genetics to management of the endangered salt marsh harvest mouse.