MOLECULAR ECOLOGY

Molecular Ecology (2014) doi: 10.1111/mec.12803

Range-wide analysis of genetic structure in a widespread, highly mobile species (*Odocoileus hemionus*) reveals the importance of historical biogeography

EMILY K. LATCH,* DAWN M. REDING,†‡ JAMES R. HEFFELFINGER,§ CARLOS H. ALCALÁ-GALVÁN¶ and OLIN E. RHODES JR**

*Department of Biological Sciences, Behavioral and Molecular Ecology Research Group, University of Wisconsin-Milwaukee, 3209 N. Maryland Ave., Milwaukee, WI 53211, USA, †Department of Biology, Luther College, 700 College Dr., Decorah, IA 52101, USA, ‡Department of Ecology, Evolution, and Organismal Biology, Iowa State University, 251 Bessey Hall, Ames, IA 50011, USA, §Arizona Game and Fish Department, 555 N. Greasewood Rd., Tucson, AZ 85745, USA, ¶DICTUS-Universidad de Sonora, Blvd, Edif. 7H. Luis Donaldo Colosio s/n, Col. Centro, Hermosillo, Sonora 83100, México, **Savannah River Ecology Laboratory, PO Drawer E, Aiken, SC 29802, USA*

Abstract

Highly mobile species that thrive in a wide range of habitats are expected to show little genetic differentiation across their range. A limited but growing number of studies have revealed that patterns of broad-scale genetic differentiation can and do emerge in vagile, continuously distributed species. However, these patterns are complex and often shaped by both historical and ecological factors. Comprehensive surveys of genetic variation at a broad scale and at high resolution are useful for detecting cryptic spatial genetic structure and for investigating the relative roles of historical and ecological processes in structuring widespread, highly mobile species. In this study, we analysed 10 microsatellite loci from over 1900 samples collected across the full range of mule deer (*Odocoileus hemionus*), one of the most widely distributed and abundant of all large mammal species in North America. Through both individual- and population-based analyses, we found evidence for three main genetic lineages, one corresponding to the 'mule deer' morphological type and two to the 'black-tailed deer' type. Historical biogeographic events likely are the primary drivers of genetic divergence in this species; boundaries of the three lineages correspond well with predictions based on Pleistocene glacial cycles, and substructure within each lineage demonstrates island vicariance. However, across large geographic areas, including the entire mule deer lineage, we found that genetic variation fit an isolation-by-distance pattern rather than discrete clusters. A lack of genetic structure across wide geographic areas of the continental west indicates that ecological processes have not resulted in restrictions to gene flow sufficient for spatial genetic structure to emerge. Our results have important implications for our understanding of evolutionary mechanisms of divergence, as well as for taxonomy, conservation and management.

Keywords: gene flow, isolation by distance, landscape genetics, *Odocoileus hemionus*, phylogeography, spatial genetic structure

Received 6 August 2013; revision received 2 May 2014; accepted 11 May 2014

Introduction

Data on the spatial distribution of genetic diversity provide interesting and sometimes unexpected insights into mechanisms of evolutionary diversification in nature, as well as valuable resources for the development of efficient conservation strategies. For species that exist in isolated groups without direct genetic or ecological connections, delineating populations across a landscape can be straightforward. However, as connectivity among

Correspondence: Emily K. Latch, Fax: 414-229-3926; E-mail: latch@uwm.edu

groups increases, delineation of population boundaries and barriers to dispersal becomes complex and is perhaps most challenging in continuously distributed species with no apparent spatial subdivision (Waples & Gaggiotti 2006). High mobility and continuously distributed habitats are expected to produce considerable genetic connectivity over broad spatial scales, often resulting in an overall lack of genetic structure (panmixia) or an isolation-by-distance (IBD) pattern. This a priori expectation of limited deviation from panmixia or IBD is perhaps one reason that widespread species have received less attention from wildlife geneticists than species with patchy distributions (Frankham *et al.* 2002; Waples & Gaggiotti 2006).

However, a growing body of literature indicates that broad-scale patterns of discrete genetic differentiation can and do emerge in these widespread, highly mobile species. Historical processes (e.g. glaciation, island formation) and major topographic features (e.g. rivers, mountains) can leave lasting genetic signatures on contemporary patterns of spatial genetic structure that persist, even in the face of high rates of gene flow (Avise 2000; Lessa *et al.* 2003; Rueness *et al.* 2003; Weckworth *et al.* 2012). Alternatively, cryptic patterns of spatial genetic structure in widespread species sometimes cannot be explained by geographic isolation or historical events, leading to suppositions that gene flow restrictions may be imposed by major ecological transitions or habitat-associated variables such as climate, vegetation and prey specialization (Geffen *et al.* 2004; McRae *et al.* 2005; Pilot *et al.* 2006, 2012; Brown *et al.* 2007; Carmichael *et al.* 2007; Musiani *et al.* 2007; Reding *et al.* 2012).

Studies of widespread, highly mobile species emphasize that patterns of spatial genetic structure are dynamic, shaped by historical, ecological and behavioural processes acting on different spatial and temporal scales (Foll & Gaggiotti 2006). To detect the multiple, hierarchical layers of structure that exist, comprehensive surveys are warranted to understand the broad-scale partitioning of genetic variation and to resolve cryptic patterns of genetic structure. Studies that span large portions of a species' range, particularly when combined with high-resolution sampling, offer a powerful means to jointly investigate the roles of historical, ecological and behavioural processes in shaping genetic diversity and structure (Hoban *et al.* 2010; Reding *et al.* 2012).

Mule deer (*Odocoileus hemionus*) are one of the most widely distributed and abundant of all large mammal species in western North America. This species occurs in diverse habitats over 40 degrees latitude throughout every major biome except tundra. Mule deer are generalist browsers that utilize a wide variety of plant foods, with preferences that vary geographically and seasonally, and readily adapt to consumption of agricultural vegetation and landscape plantings (Hall 1981). The mating system in mule deer is polygynous, in which male reproductive success is correlated with body size (Miller 1974; Kucera 1978). Mule deer are capable of dispersing very long distances (>250 km), although dispersal distances vary considerably among individuals, seasons and study areas (Robinette 1966; Conner & Miller 2004). Historically, mule deer have remained common throughout North America, although population sizes have fluctuated considerably over time.

Like many North American mammals, the biogeographic history of mule deer is intertwined with Pleistocene glacial cycles and late Quaternary climate change. Mule deer are traditionally divided into two main types: mule deer (MD) and black-tailed deer (BTD). Morphological distinction of these two groups is supported by genetic data, which indicates divergence of 6–8% in mitochondrial DNA (Carr *et al.* 1986; Cronin *et al.* 1988; Cronin 1991a; Carr & Hughes 1993; Latch *et al.* 2009), among the highest intraspecific values reported for mammals. Nuclear markers exhibit less, although still notable, divergence between the two groups (Gavin & May 1988; Cronin 1991b; Cathey *et al.* 1998; Haynes & Latch 2012). Patterns of divergence likely reflect isolation in distinct glacial refugia during the Pleistocene, separated by the Cascade Mountain range in Oregon and Washington, USA (Latch *et al.* 2009). Following glacial retreat, these two lineages came back into contact, where they hybridize readily (Latch *et al.* 2011). Despite extensive admixture at the zone of contact, pure populations of MD and BTD remain outside the contact zone, and assortative mating may persist (Müller-Schwarze & Müller-Schwarze 1975).

Ecological drivers of gene flow may superimpose additional structure onto that generated by historical processes. Evidence that ecological factors may influence gene flow at broad spatial scales in mule deer comes from the boundaries of the 11 morphologically defined mule deer subspecies (*O. h. sitkensis* and *columbianus* within the BTD type, and *O. h. hemionus*, *fuliginatus*, *californicus*, *inyoensis*, *eremicus*, *crooki*, *peninsulae*, *sheldoni* and *cerrosensis* within the MD type) that roughly correspond to some of the major habitat transitions in North America (Hall 1981). In addition, Pease *et al.* (2009) found support for regional genetic structure in accordance with California bioregions, suggesting that ecological factors might be important in limiting gene flow among mule deer populations at regional scales.

Alternatively, behavioural mechanisms such as high gene flow may have eroded patterns of spatial genetic structure at a broad scale. Studies of genetic structure in mule deer have detected only low levels of genetic

differentiation at local scales and little evidence of substructure among geographically proximate populations, which is characteristic of widespread species with high gene flow (Smith *et al.* 1990; Scribner *et al.* 1991; Cullingham *et al.* 2011; Powell *et al.* 2013). Even when separated by formidable geographic barriers, divergence is low and only slightly less pronounced in the dispersing sex (males in mule deer; Cronin *et al.* 1991; Travis & Keim 1995; Powell *et al.* 2013). Mule deer exhibit extensive phenotypic and behavioural variation throughout their range, which has prohibited construction of a key to differentiate subspecies reliably (Cowan 1936). A lack of discrete variation across major ecological transitions provides additional support for the hypothesis that high gene flow may have eroded genetic structure imposed by historical and ecological processes.

In this study, we provide the first high-resolution, continent-scale assessment of nuclear genetic structure for an ungulate. Our primary goal was to quantify broad-scale patterns of spatial genetic structure in mule deer sampled across its full range in western North America and to determine whether historical biogeographic and contemporary ecological processes have resulted in cryptic spatial genetic structure. We also examine potential mechanisms involved in establishing and maintaining population divergence in this widespread, mobile ungulate. To achieve this goal, we sampled over 1900 georeferenced individuals collected throughout the mule deer range and assessed genetic variation and structure at 10 microsatellite loci using both individual- and population-based analyses.

Methods

Samples

Samples were collected from 1957 *Odocoileus hemionus* throughout the entire range of the species, including the USA, Canada and Mexico from 1995 to 2005 (Fig. 1 and Fig. S1, Supporting information). The sample set represented 10 of the 11 morphologically described subspecies; *O. h. cerrosensis*, endemic to Cedros Island, Baja California Sur, Mexico, was not sampled. Each sample was assigned to a subspecies based on both morphology of the animal and geographic location of sample collection. Samples included tissue (n = 1872), hair (n = 69), antler cores (n = 8) and blood (n = 8). The 71 sampling groups match those locations used in Latch *et al.* (2009), with the following exceptions: (i) sampling location BC-CR was omitted; (ii) sampling location WA-XX was further subdivided into three locations that better reflected geography (WA-XX, WA-NW and WA-SE; Fig. 1). Sampling groups never included the geographic range of multiple subspecies, and the geographic area of sampling groups was small relative to the total sample area (Fig. S1, Supporting information). Population codes are alphanumeric, with the prefix describing the state or province of collection and the suffix describing the specific population sample location. Spatial coordinates were obtained for each sample, either as GPS coordinates recorded at the time of sample collection or estimated based on written descriptions of the sampling locations. Explicit spatial coordinates were used for all individual-level analyses (Fig. S1, Supporting information). For population-level analyses, we used ARCGIS to calculate the mean centre of each of the 71 sampling groups and used those spatial coordinates in subsequent analyses (Fig. 1).

Laboratory analysis

We extracted DNA from all samples and all sample types using a modified ammonium acetate protocol (Latch *et al.* 2008a). We amplified 10 nuclear microsatellite loci for each sample in 10 μL multiplex reactions using the primers and conditions outlined in Latch *et al.* (2008b). All hair and antler core samples (n = 77) were independently extracted and genotyped twice for confirmation of results. We observed two instances of allelic dropout in hair samples and corrected these errors. Additionally, approximately 5% of the tissue samples (n = 96) were genotyped twice to estimate genotyping error rate. Of the 960 repeated genotypes, we observed 11 total mismatches (all instances of allelic dropout of the larger allele) across eight different loci, resulting in an estimated error rate of 1.1% for the entire data set. We culled any samples from the database for which we obtained fewer than six genotypes, that had inaccurate or missing spatial coordinates or that showed any genetic evidence of admixture with white-tailed deer (based on preliminary clustering analyses that separated these species; hybridization is known to occur in areas of sympatry; Derr 1991; Kay & Boe 1992; Hughes & Carr 1993; Cathey *et al.* 1998; Hornbeck & Mahoney 2000; Bradley *et al.* 2003). The final data set included 1831 georeferenced samples genotyped at 10 microsatellite loci. The percentage of genotypes missing from the final data set was 0.6% (110 missing genotypes of 18 310 attempted).

Population-level analyses

For the total sample and each of the 71 sample groups with at least eight individuals (n = 65 sample groups), we used GENEPOP 4.0 (Rousset 2008) to perform tests of linkage disequilibrium for each pair of microsatellite loci, assess deviation from Hardy–Weinberg equilibrium (HWE) and estimate F_{IS} (Weir & Cockerham 1984).

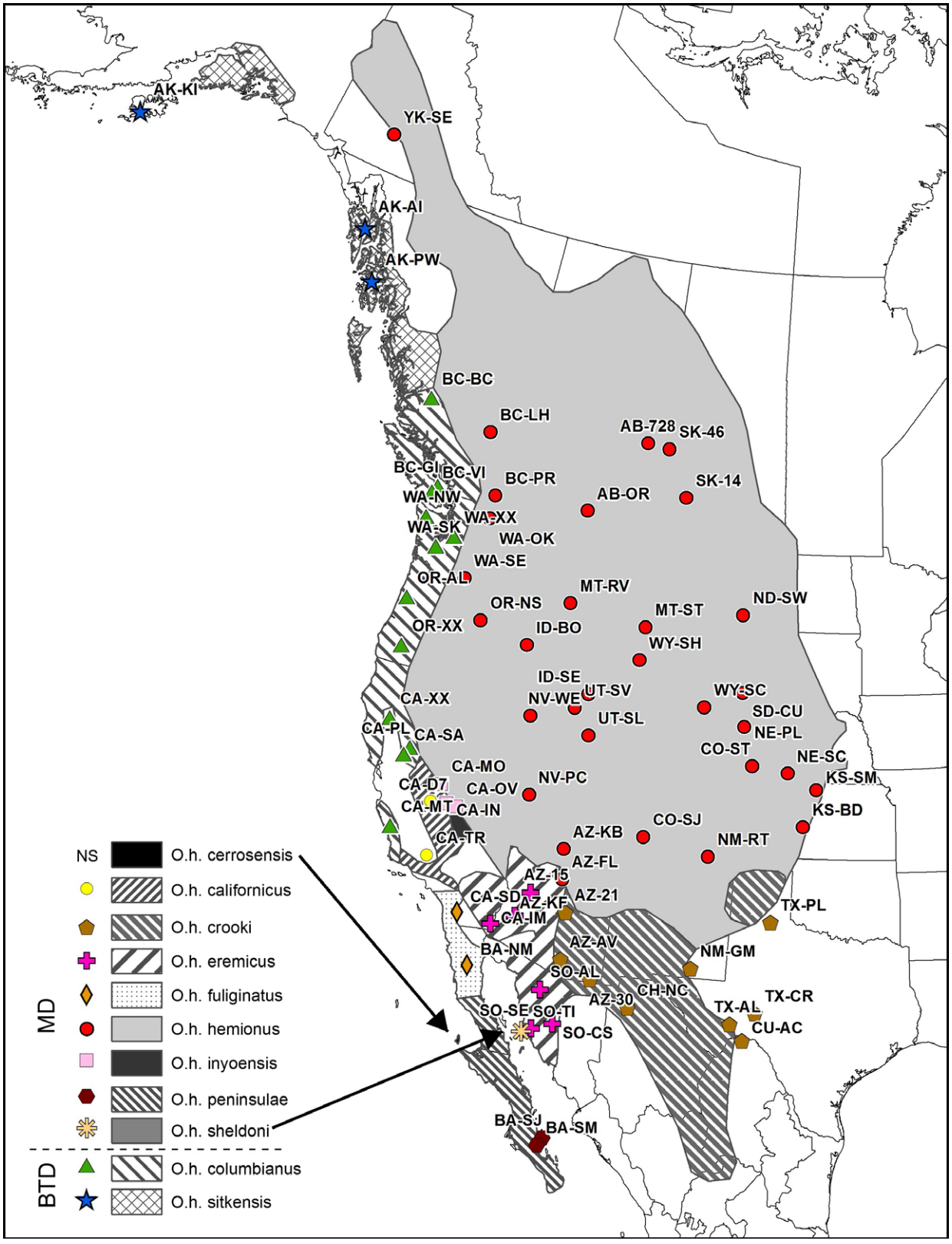

Fig. 1 Geographic locations of $n = 71$ mule deer (*Odocoileus hemionus*) sample groups. Symbols indicate a priori assignment to one of 10 morphological subspecies. Patterned polygons indicate the range of each subspecies. Population codes are alphanumeric, with the prefix describing the state or province of collection and the suffix describing the specific sample location.

We used ARLEQUIN 3.5 (Excoffier *et al.* 2005) to calculate expected (H_E) and observed (H_O) heterozygosities, and FSTAT 2.9.3 (Goudet 2001) to calculate allelic richness corrected for differences in sample size. In addition, we used ARLEQUIN to calculate pairwise F_{ST} as a measure of differentiation among sample groups, as well as to test for a pattern of isolation by distance (IBD) by conducting a Mantel test (Mantel 1967) between a genetic distance matrix ($F_{ST}/(1-F_{ST})$) and a geographic distance matrix (*ln*-transformed km) with 1000 permutations. To more effectively view the patterns of genetic differentiation among sample groups, we used GENALEX 6.5 (Peakall & Smouse 2012) to construct a principal coordinate analysis (PCoA) graph via the covariance matrix with data standardization derived from a matrix of mean genotypic genetic distance. In all assessments of significance, we used the sequential Bonferroni method to adjust for multiple tests (Rice 1989).

Individual-level Bayesian clustering

Complimentary analytical approaches are often useful in characterizing patterns of population genetic structure (Balkenhol *et al.* 2009). Thus, we employed both spatially informed and nonspatial approaches to characterize genetic structure across the data set. In our spatially informed approach, we employed Bayesian clustering using BAPS 5 (Corander *et al.* 2008). As a number of samples in the data set have identical coordinates, we performed a mixture analysis using the 'spatial clustering of groups' model. Mixture analysis was performed for five iterations of each value of K_{max} (the maximum number of populations) for the range K_{max} = 1–65. In this step, the maximum likelihood and highest posterior probability were used to determine the optimum number of genetic clusters in the sample, and then each individual was assigned to a cluster. In the second step, we performed an admixture analysis based on the results of the mixture clustering, using 500 iterations, 200 reference individuals per population and 100 iterations for each reference individual.

In our nonspatial approach, we used STRUCTURE (version 2.3.4; Pritchard *et al.* 2000; Falush *et al.* 2003; Hubisz *et al.* 2009) in an iterative analysis. We performed clustering for all 1831 individuals, for 10 iterations of each value of K for the range K = 1–65. Each run consisted of 500 000 replicates of the MCMC after a burn-in of 50 000 replicates. We used the admixture model and allowed allele frequencies to be correlated. We observed the common phenomenon that once the true K is reached, likelihoods for larger Ks plateau and the variance among runs increase (Pritchard & Wen 2003). Thus, the optimal value for K was determined both by inspecting the likelihoods for each set of runs and by calculating the ad hoc ΔK statistic (Evanno *et al.* 2005; Fig. S2, Supporting information). To assign individuals to inferred clusters, we performed a final set of 10 runs (1 000 000 burn-in and 1 000 000 stored replicates) at the optimal K. Values of q, the proportion of an individual's sampled genome characteristic of each cluster, were averaged across the 10 runs using CLUMPP software (Jakobsson & Rosenberg 2007). Each individual was assigned to the cluster in which it had the highest probability of membership. The entire clustering process was repeated iteratively for each of the inferred clusters, to identify any additional genetic substructure (Evanno *et al.* 2005; Latch *et al.* 2006). These iterative runs were undertaken for each value of K for the range K = 1 to K = the number of sampling locations represented in the subsample.

To visualize spatial genetic structure detected using STRUCTURE, we used the ancestry coefficients (q) from the STRUCTURE analysis to create a genetic surface for each of the K = 3 inferred clusters (see Results). Following Murphy *et al.* (2008), we constructed these surfaces in ARCGIS using inverse distance-weighted interpolation (IDW) (Cressie 1993) with a cell size of 12 km, a power function of 2 and a variable search radius of 12 neighbouring points. For each raster, we also used ARCGIS to generate contour lines to identify regions of high assignment ($q \geq 0.75$) to a given cluster. As another way to visualize putative subpopulation 'boundaries', we selected all individuals for which $q \geq 0.75$ for a given cluster and used the *kde* function in Geospatial Modelling Environment (GME; Beyer 2012) to calculate kernel density estimates. We used the plug-in algorithm for bandwidth estimation because it has been shown to perform well and is not as sensitive to points with identical coordinates as some alternative algorithms (Beyer 2012). We used a cell size of 1,500, based on the manual's suggestion for calculating a reasonable cell size. We subsequently used the *isopleth* function in GME to create 90% polygons (i.e. contains 90% of the volume of the surface) for each of the kernel density estimate rasters. We considered these polygons to reflect the core areas of each subpopulation.

Individual-level principal components analysis (PCA and sPCA)

To complement the clustering analyses, we also performed individual-based principal components analysis (PCA) and spatial principal components analysis (sPCA) using the ADEGENET 1.3-4 package (Jombart 2008) of the software R (R Development Core Team 2012). An advantage of PCA-type approaches is that they do not rely on Hardy–Weinberg or linkage equilibrium and thus are useful for detecting clines and other complex

genetic structures (Jombart *et al.* 2008). The PCA implementation in ADEGENET calls upon the *dudi.pca* function from the ADE4 package (Dray & Dufour 2007). When using individuals as samples, this method performs a PCA on a matrix of allele frequencies, with each row representing an individual and each column representing an allele (coded 0 if absent, 0.5 if heterozygous or 1 if homozygous). As recommended in the ADEGENET manual, missing data were replaced with the mean of the concerned allele frequency. We also used the software to produce a scatterplot of individual genotypes along the first two principal components and construct 95% inertia ellipses for the 10 subspecies to visualize how much separation (and thus differentiation) exists among the groups.

The sPCA technique provides principal component scores that summarize both the nonspatial genetic variability and the spatial autocorrelation structure among individual genotypes (as implemented here) or populations (Jombart *et al.* 2008). In this method, highly positive eigenvalues reflect axes with both a large variance and global (i.e. positive spatial autocorrelation) structure. Consistent with the genetic surface approach, we used a nearest neighbour ($k = 12$) connection network. We also examined a distance-based connection network in which individuals separated by ≤250 km (approximately, the maximum recorded natal dispersal distance for this species; Robinette 1966) were considered neighbours. Smaller and larger distances were also evaluated, but yielded concordant results. We present only the findings from the nearest neighbour network as the two approaches provided highly similar results. We used the *screeplot* function to graph the eigenvalues, each decomposed into its variance and spatial autocorrelation (Moran's *I*) components, and used the graph to choose how many principal components to interpret. Jombart *et al.* (2008) recommend interpreting only the principal components containing sufficient variability and spatial structuring and that can be well distinguished from all other principal components based on a sharp decay between successive eigenvalues. We conducted a permutation procedure with 999 randomizations to test for a significant global pattern in the data. In visualizations, we used the lagged spatial principle component scores (i.e. averaged with its neighbours) rather than the original scores to smooth the underlying pattern.

Evaluation of classification schemes

To evaluate significant geographic divisions of hypothesized a priori groupings based on taxonomy and ecological regions, we used hierarchical analyses of molecular variance (AMOVA; Excoffier *et al.* 1992) in ARLEQUIN (Excoffier & Lischer 2010). This analysis divides total genetic variance into variance components via differences among groups (Φ_{CT}), among populations within groups (Φ_{SC}) and within populations (Φ_{ST}). Hypothesized AMOVA designs based on taxonomic and ecological data were contrasted to a posteriori groups implied from analyses of nuclear genetic data in this study and from previous analyses of mtDNA data. The optimal genetic subdivisions are expected to maximize the between-group variance (Φ_{CT}) compared to within-group variance (Φ_{SC}). We used 16,000 permutations of the data, 1 000 000 steps in the Markov chain and 100 000 dememorization steps in the analysis (Guo & Thompson 1992). A total of six models were evaluated: Model A – a priori, taxonomically designated 10 subspecies (*O. h. sitkensis, columbianus, hemionus, californicus, fuliginatus, inyoensis, eremicus, crooki, peninsulae and sheldoni*); Model B – a priori, taxonomically designated mule deer (MD), Sitka black-tailed deer (*O. h. sitkensis*) and Columbian black-tailed deer (*O. h. columbianus*); Model C – a priori, taxonomically designated mule deer (MD) and black-tailed deer (BTD); Model D – a posteriori grouping according to mtDNA haplogroup (10 groupings based on Latch *et al.* 2009); Model E – a posteriori grouping according to highest order STRUCTURE results (three lineages; see Results); Model F – a priori grouping according to the Commission for Environmental Cooperation (CEC) Level 1 Ecoregions (CEC 1997). These broad ecological regions highlight areas with similarities in geology, physiography, vegetation, climate, soils, land use, wildlife distributions and hydrology. Level 1 classification divides North America into 15 broad ecoregions; seven of which were represented among the 65 sample groups with at least 8 samples. We also performed AMOVA analysis on only the MD group, testing models that split samples based on subspecies, mtDNA haplogroup and ecoregion (analogous to Models A, D and F above), but did not find strong support for any of the models.

Results

Population-level analyses

In the global sample, all loci showed a significant heterozygote deficit and all pairs of loci showed significant departure from linkage equilibrium, naturally due to population subdivision. Basic genetic diversity statistics for the 65 sample groups with greater than eight individuals sampled indicate that allelic richness was lowest in island populations and highest in California and the Pacific Northwest (Table 1). Genetic variation in island populations overall (AK-KI, AK-AI, AK-PW, BC-GI, BC-VI, SO-TI) was reduced relative to mainland populations

Table 1 Summary of genetic variation, averaged across 10 microsatellite loci, for 65 populations with $n > 8$ samples. Sample size (n), allelic richness (A_R), mean expected (H_E) and observed (H_O) heterozygosities, and F_{IS} for 10 microsatellite loci

Population	n	A_R	H_E (SE)	H_O (SE)	F_{IS} (SE)
AK-KI	175	1.87	0.297 (0.072)	0.283 (0.069)	**0.044 (0.025)**
AK-AI	44	1.70	0.195 (0.065)	0.178 (0.059)	**0.074 (0.040)**
AK-PW	35	2.10	0.304 (0.087)	0.268 (0.076)	**0.084 (0.041)**
BC-BC	18	3.46	0.606 (0.039)	0.533 (0.046)	**0.126 (0.031)**
BC-GI	27	2.35	0.375 (0.086)	0.323 (0.072)	**0.094 (0.053)**
BC-VI	30	3.08	0.477 (0.090)	0.401 (0.080)	**0.201 (0.096)**
WA-NW	13	3.50	0.547 (0.064)	0.563 (0.065)	−0.050 (0.055)
WA-XX	14	3.93	0.644 (0.043)	0.717 (0.037)	**−0.135 (0.055)**
WA-SK	25	4.04	0.653 (0.040)	0.656 (0.049)	0.000 (0.034)
OR-AL	30	4.08	0.645 (0.047)	0.668 (0.052)	−0.038 (0.051)
OR-XX	43	4.38	0.670 (0.055)	0.594 (0.054)	**0.109 (0.041)**
CA-XX	10	4.54	0.664 (0.048)	0.649 (0.057)	0.029 (0.038)
CA-MT	25	4.52	0.703 (0.043)	0.658 (0.035)	0.037 (0.066)
CA-D7	23	4.48	0.706 (0.038)	0.698 (0.052)	0.008 (0.053)
CA-TR	30	4.33	0.687 (0.054)	0.653 (0.061)	**0.051 (0.046)**
AZ-21	23	3.85	0.625 (0.054)	0.595 (0.073)	0.029 (0.086)
TX-PL	11	3.53	0.529 (0.067)	0.500 (0.085)	0.052 (0.081)
AZ-AV	28	3.29	0.588 (0.049)	0.561 (0.063)	0.048 (0.067)
NM-GM	27	3.72	0.603 (0.054)	0.611 (0.057)	−0.028 (0.063)
AZ-30	30	3.78	0.606 (0.060)	0.564 (0.067)	0.057 (0.060)
CH-NC	21	3.81	0.599 (0.060)	0.542 (0.083)	**0.179 (0.103)**
TX-CR	20	3.77	0.611 (0.067)	0.542 (0.080)	**0.109 (0.082)**
TX-AL	30	3.68	0.595 (0.066)	0.561 (0.075)	0.043 (0.074)
AZ-15	14	3.92	0.632 (0.056)	0.629 (0.071)	−0.007 (0.071)
AZ-KF	30	3.63	0.608 (0.067)	0.587 (0.079)	0.028 (0.069)
CA-IM	30	3.09	0.561 (0.046)	0.560 (0.070)	0.006 (0.078)
SO-AL	22	4.01	0.637 (0.051)	0.618 (0.061)	0.008 (0.074)
SO-CS	23	3.61	0.577 (0.062)	0.539 (0.077)	0.050 (0.093)
CA-SD	28	3.93	0.675 (0.047)	0.648 (0.044)	0.027 (0.041)
BA-NM	10	3.53	0.556 (0.077)	0.581 (0.094)	−0.053 (0.080)
YK-SE	30	3.37	0.580 (0.058)	0.533 (0.057)	**0.103 (0.074)**
BC-LH	28	4.11	0.651 (0.044)	0.645 (0.050)	0.007 (0.041)
AB-728	29	4.13	0.683 (0.038)	0.634 (0.058)	**0.071 (0.059)**
SK-46	29	3.87	0.617 (0.059)	0.600 (0.054)	0.005 (0.049)
BC-PR	20	3.87	0.650 (0.047)	0.619 (0.050)	0.041 (0.050)
SK-14	30	3.80	0.631 (0.052)	0.577 (0.066)	**0.098 (0.057)**
AB-OR	29	4.13	0.661 (0.044)	0.621 (0.066)	0.056 (0.074)
WA-OK	29	3.93	0.658 (0.041)	0.644 (0.048)	**0.031 (0.027)**
WA-SE	17	4.31	0.694 (0.042)	0.682 (0.060)	0.027 (0.048)
MT-RV	28	4.07	0.648 (0.046)	0.625 (0.053)	0.034 (0.045)
ND-SW	30	4.01	0.652 (0.047)	0.633 (0.065)	0.029 (0.068)
OR-NS	28	4.02	0.667 (0.035)	0.657 (0.044)	0.017 (0.031)
MT-ST	30	4.27	0.665 (0.058)	0.650 (0.070)	0.030 (0.043)
ID-BO	25	3.90	0.623 (0.050)	0.637 (0.057)	−0.027 (0.043)
WY-SH	30	4.17	0.667 (0.046)	0.607 (0.063)	**0.093 (0.062)**
SD-CU	30	4.01	0.656 (0.041)	0.673 (0.050)	−0.034 (0.053)
ID-SE	28	4.06	0.646 (0.046)	0.610 (0.051)	**0.061 (0.041)**
WY-SC	19	4.05	0.644 (0.035)	0.574 (0.039)	**0.103 (0.053)**
UT-SV	23	3.91	0.637 (0.042)	0.646 (0.043)	−0.021 (0.038)
NV-WE	30	4.10	0.656 (0.044)	0.660 (0.045)	−0.012 (0.036)
NE-PL	30	3.72	0.625 (0.049)	0.603 (0.052)	0.026 (0.052)
UT-SL	23	3.99	0.665 (0.045)	0.657 (0.048)	−0.001 (0.052)
CO-ST	26	3.95	0.622 (0.052)	0.613 (0.075)	0.032 (0.073)
NE-SC	29	3.92	0.608 (0.057)	0.614 (0.065)	−0.019 (0.064)

Table 1 *Continued*

Population	*n*	A_R	H_E (SE)	H_O (SE)	F_{IS} (SE)
KS-SM	27	3.85	0.601 (0.052)	0.596 (0.052)	0.000 (0.035)
NV-PC	27	4.07	0.677 (0.039)	0.659 (0.038)	0.015 (0.047)
KS-BD	31	3.93	0.611 (0.058)	0.606 (0.054)	−0.002 (0.032)
CO-SJ	29	4.27	0.670 (0.041)	0.604 (0.060)	**0.102 (0.063)**
AZ-KB	30	3.60	0.653 (0.029)	0.613 (0.047)	**0.066 (0.047)**
NM-RT	29	3.89	0.613 (0.052)	0.583 (0.067)	**0.069 (0.054)**
AZ-FL	26	3.91	0.636 (0.045)	0.572 (0.060)	**0.089 (0.080)**
CA-MO	25	4.19	0.680 (0.037)	0.720 (0.048)	**−0.063 (0.053)**
CA-OV	20	4.03	0.671 (0.035)	0.658 (0.054)	0.010 (0.068)
BA-SM	15	3.41	0.549 (0.090)	0.475 (0.084)	**0.105 (0.067)**
SO-TI	8	1.79	0.277 (0.055)	0.279 (0.087)	0.036 (0.184)

Standard errors for estimates of H_O, H_E and F_{IS} are provided in parentheses. F_{IS} values significantly different from zero are in bold font.

($A_{Rislands}$ = 2.15, SD = 0.513; $A_{Rmainland}$ = 3.92, SD = 0.298; paired *t*-test P = 0.0003). Deviations from Hardy–Weinberg equilibrium were evident in many sample groups (Table 1). The highest F_{IS} values were often found in island populations (suggesting inbreeding), in sample groups that covered large geographic areas or in areas of inferred admixture between divergent lineages (suggesting Wahlund effect; Table 1).

Estimates of pairwise F_{ST} ranged from −0.00156 (between NE-SC and KS-SM) to 0.64377 (between SO-TI and AK-AI; Fig. S3, Supporting information), with the largest values generally occurring in comparisons between Sitka black-tailed deer (*O. h. sitkensis*) sample groups vs. sample groups comprised of alternative subspecies (Fig. S3, Supporting information). A significant pattern of isolation by distance was observed overall (r = 0.452, P = 0.001). The principal coordinate analysis revealed a high degree of genetic differentiation between the three *sitkensis* sample groups and all others (Fig. 2). The PCoA also provided support for differentiation between Columbian black-tailed deer (*O. h. columbianus*) sample groups and others, although there was some overlap with mule deer sample groups. Within the eight mule deer subspecies, there was little evidence for differentiation other than the separation of the lone *sheldoni* sample group. However, there was a north–south trend observable among the mule deer subspecies, which was particularly noticeable along the third axis (Fig. S4, Supporting information). Scores from the first three principal coordinate axes accounted for 57.3% of the total variation in the data (first axis 34.6%; second axis 13.9%; third axis 8.8%).

Individual-level Bayesian clustering

In the BAPS analysis, K = 14 was the optimal solution (Fig. S5, Supporting information). The mode of the posterior distribution of K was at K = 15; however, one of the clusters contained only a few individuals geographically scattered in the northwestern United States and was not considered further (Corander *et al.* 2006; Latch *et al.* 2006). As described using population-based analyses, samples collected from *O. h. sitkensis* (two clusters) and *O. h. columbianus* (two clusters) were identified as distinct from the remaining mule deer samples (10 clusters). The presence of spatially explicit clusters across the sample area suggests that the pattern of genetic variation is distributed spatially. However, isolation-by-distance patterns of gene flow can produce clustering patterns that mimic discrete spatial structure in some cases (Frantz *et al.* 2009). We used the spatial model for BAPS, which imposes a prior distribution that puts more weight on geographically homogeneous partitions of the sample. This model does not allow estimation of the magnitude and scale of spatial correlations in the presence of IBD (Francois & Durand 2010), which could result in the designation of distinct clusters in areas characterized by a gap in sampling (Frantz *et al.* 2009).

Hierarchical analyses in STRUCTURE clearly indicated three main genetic groups, corresponding to black-tailed deer of subspecies *O. h. sitkensis* ($BTD_{sitkensis}$), black-tailed deer of subspecies *O. h. columbianus* ($BTD_{columbianus}$) and mule deer of all remaining subspecies (MD; Fig. 3). The genetic surfaces and associated contour lines support the distinctiveness of $BTD_{sitkensis}$ and show a zone of admixture between mule deer and $BTD_{columbianus}$ that extends north–south (i.e. between the red and green contour lines in Fig. S6, Supporting information) predominantly along the Cascade–Sierra chain. This admixture zone is also evident via the overlapping 90% polygons from the kernel density estimates (Fig. S7, Supporting information). Hierarchical analyses revealed additional substructure within each of the three main groups (Fig. 4). $BTD_{sitkensis}$ substructure separated the introduced population of deer on Kodiak Island from the rest of the Alexander Archipelago. A limited

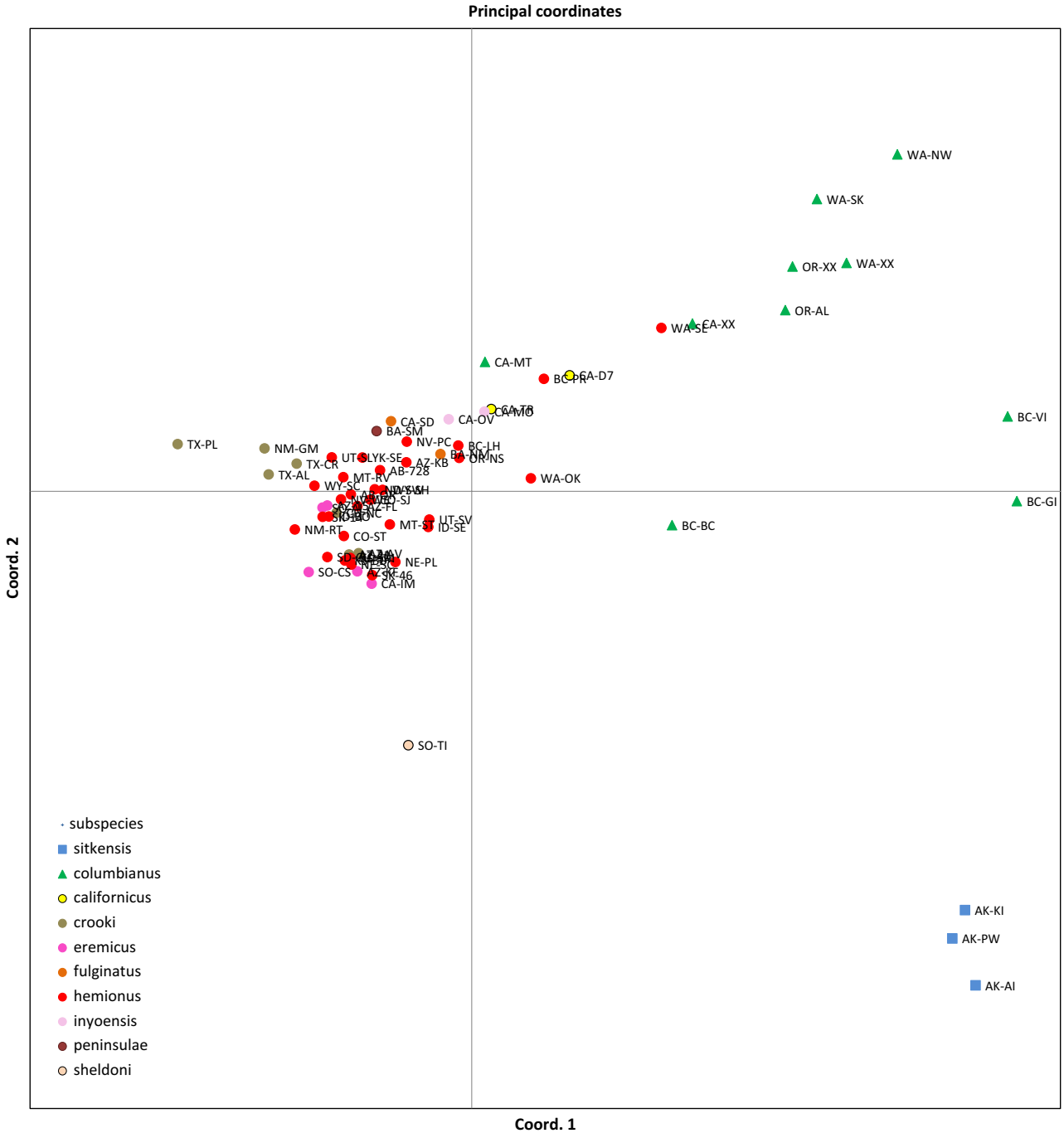

Fig. 2 Principal coordinate analysis of the 65 sample groups with $n \geq 8$, colour-coded according to a priori classification to one of ten morphological subspecies. The first two axes are plotted.

number of cross-assignments were observed and are likely due to the use of Alexander Archipelago deer as a source for the original introduction to Kodiak Island. $BTD_{columbianus}$ was further divided into two subclusters that separated Vancouver and Gabriola Islands from mainland samples. MD was subdivided into perhaps seven subclusters, although the evidence for seven subclusters over other numbers of subclusters was not strongly supported (Fig. S2, Supporting information). The seven subclusters corresponded only weakly to geography and exhibited diffuse boundaries between subclusters (Fig. 4). Alternative numbers of subclusters did not yield populations with distributions more closely associated with geography.

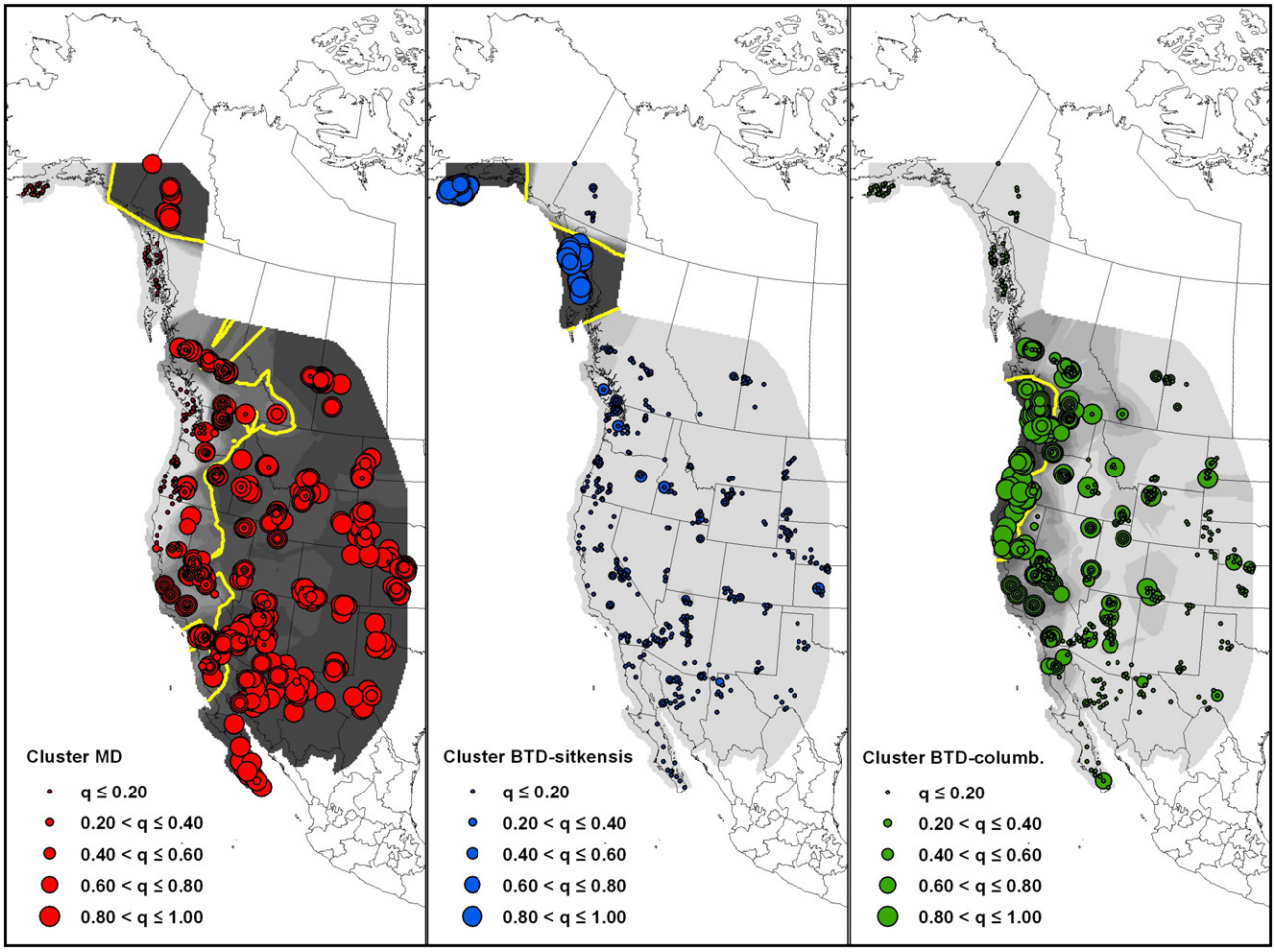

Fig. 3 Results of the first round of STRUCTURE analysis, where the size of each circle reflects the ancestry coefficient (*q*-value) for a given sample. Panels (left to right) provide results for the mule deer (MD), $BTD_{sitkensis}$ and $BTD_{columbianus}$ lineages. The genetic surfaces (darker reflects higher *q*-value) underlying the samples are shown. The lines mark the 75% contour line (i.e. $q = 0.75$) for the genetic surface.

Despite differences in the underlying analytical methodologies, spatial and nonspatial Bayesian clustering approaches yielded results that were broadly concordant. We identified 14 clusters using BAPS and 11 clusters using the iterative STRUCTURE analysis. The BTD lineage was consistently divided into four groups using both approaches, and individual assignments were nearly identical. Within the MD lineage, there was general agreement in the overall geographic distribution of clusters, although there were differences both in the number of clusters (10 using BAPS and 7 using STRUCTURE) and the locations of boundaries between individual clusters. Aside from the fundamental difference in the dependency on spatial data in the two approaches, our implementation of BAPS and STRUCTURE also differs in the unit of analysis. Because multiple samples in our data set have identical geographic coordinates, the spatial clustering analysis in BAPS was carried out with the population as the unit of analysis ('spatial clustering of groups'). This approach identifies differences between the underlying allele frequency distributions of pregrouped populations. This is in contrast to the individual-based STRUCTURE analysis, in which individuals are partitioned based on Hardy–Weinberg equilibrium. This may help to explain some of the seemingly sharp transitions between clusters in the BAPS analysis that were not reflected in the STRUCTURE analysis. Nonetheless, there was broad-scale concordance between the approaches.

Individual-level principal components analysis (PCA and sPCA)

Similar to the PCoA and Bayesian clustering findings, the standard PCA distinguished three main genetic groups. The scatterplot of the first two principal components showed *O. h. sitkensis* to be the most genetically distinct group, with more overlap between *O. h. columbianus* and mule deer. Much of the overlap between

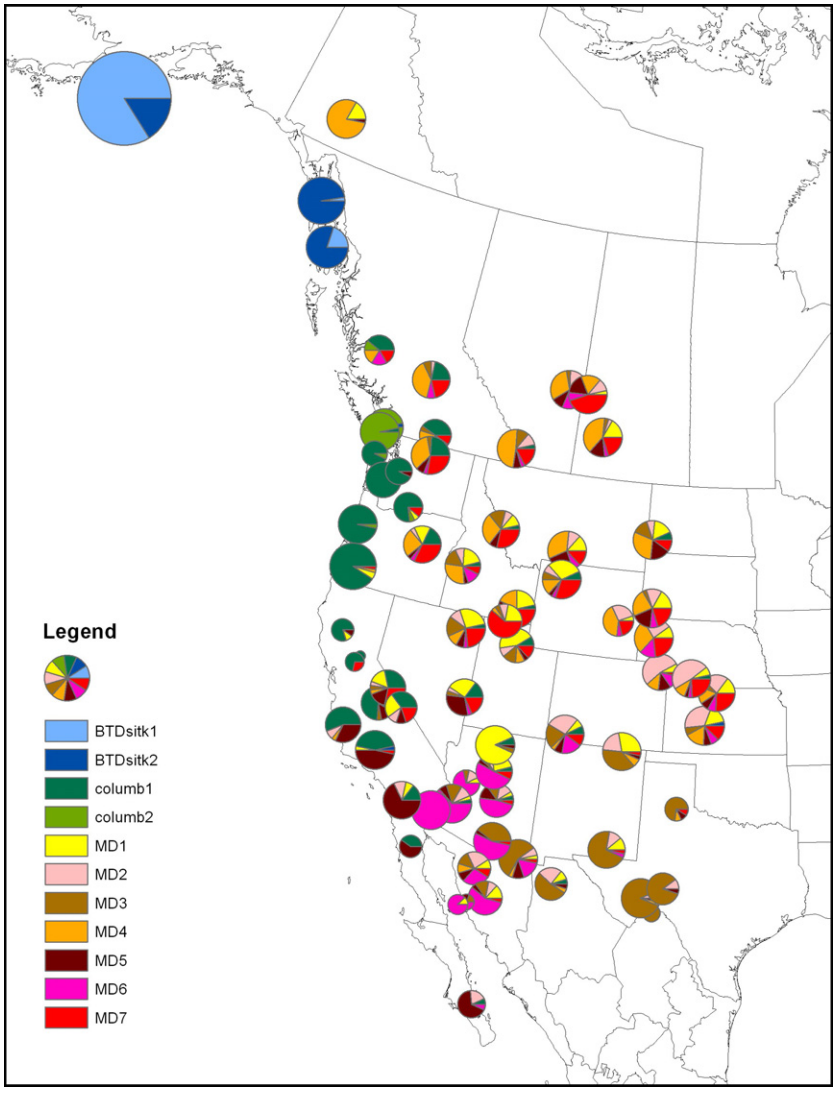

Fig. 4 Map showing results from the full iterative STRUCTURE analysis, with each sample group depicted as a pie chart. The size of the pie chart reflects the number of samples.

O. h. columbianus and mule deer occurred in samples collected from the known zone of admixture between these two lineages (Latch *et al.* 2011). In contrast, the PCA provided little resolution among the mule deer subspecies (Fig. 5).

In the sPCA, the screeplot of the decomposed eigenvalues showed that the first eigenvalue was considerably larger than all others. It exhibited a strong signal of positive spatial autocorrelation (I = 0.815) and represented a considerable proportion of the entire genetic variability (var = 0.616; Fig. S8, Supporting information). However, the first four eigenvalues could be easily distinguished from all others, and thus, the corresponding principal components were interpreted. The scores from the first spatial principal component (sPC) distinguish mule deer from both subspecies of black-tailed deer (Fig. 6). The scores from the second sPC distinguished *columbianus* and *peninsulae* from the other subspecies. The scores from the third sPC appeared to predominantly distinguish northern MD from southern MD, as well as to differentiate the two *sitkensis* groups from one another. The scores from the fourth sPC were difficult to interpret, but appeared to be separating out Baja and southern California samples, and again distinguishing the two *sitkensis* groups from one another. The permutation test confirmed the existence of at least one global pattern (P = 0.028).

Evaluation of classification schemes

Based on the 3-level AMOVAs (region, population, individuals), the greatest proportion of variation was explained by Model B, in which samples were grouped into three a priori taxonomic groups (16.93% of variance explained among groups). Model E, in which samples were grouped a posteriori based on their membership

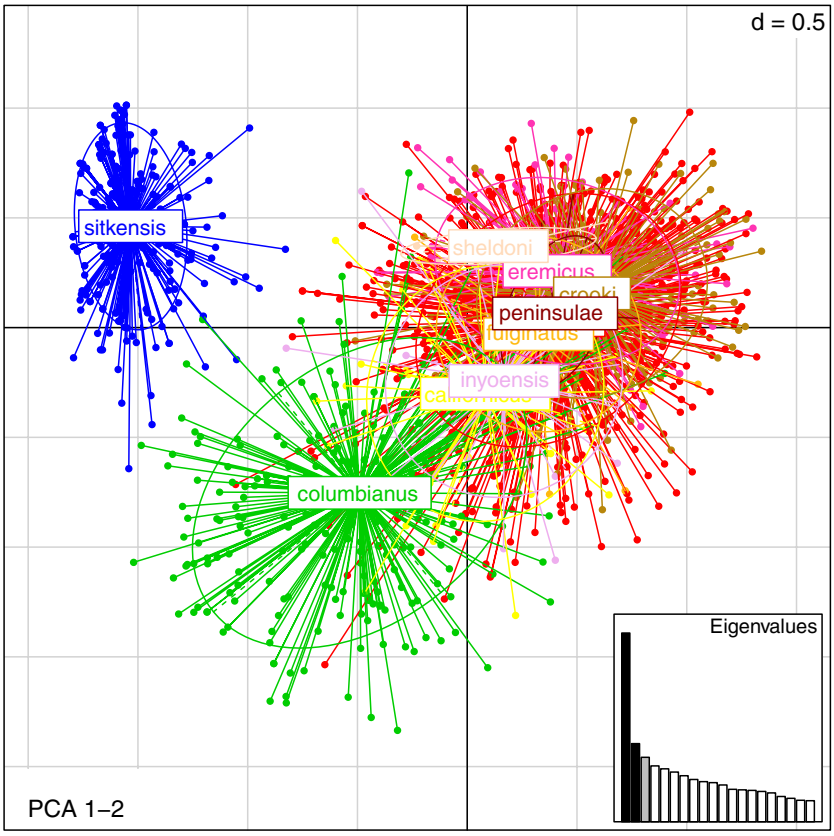

Fig. 5 Scatterplot showing results from individual-based principle components analysis. Points represent genotypes of individual mule deer, colour-coded according to a priori classification to one of the ten morphological subspecies. Subspecies are labelled inside their 95% inertia ellipses. Inset shows a bar chart of the eigenvalues with corresponding components in black. Grid distance corresponds to a value of 0.5.

to one of three genetically defined clusters, also explained a comparable portion of the variation among groups (15.80%; Table 2). Models based on larger (Model A) or smaller (Model C) numbers of subspecies groups, and ecologically based models (Model F) explained less of the variation among groups than other models (Table 2).

Discussion

Mule deer possess many characteristics that are predicted to limit genetic differentiation at a broad scale. They are currently and historically abundant, continuously distributed, highly mobile and exist at high population densities. Despite the potential for limited genetic differentiation in mule deer, our data revealed the presence of broad-scale spatial genetic structure. Both individual-based and population-based analyses supported the existence of three main genetic lineages and resolved additional substructure within each lineage. The geographic distributions of these three primary lineages correspond well to predictions based on the biogeographic history of mule deer and suggest that the broad-scale patterns of spatial genetic structure we observed are largely the result of historical restrictions to gene flow. Specifically, Pleistocene glacial cycles left distinct and lasting signatures on this species and are most evident in the strong genetic differentiation observed between mule deer and black-tailed deer. The strongest influence on secondary genetic substructure was island vicariance, which resulted in detectable patterns of genetic divergence and reduced genetic variation of island populations relative to mainland populations. Although mule deer can swim distances of several miles (Taylor 1956; Robinette 1966; Reimchen *et al.* 2003; Lopez 2006), our data suggest that gene flow between mainland and island populations is infrequent. The patterns of population genetic structure that emerged as a result of these historical biogeographic events have been maintained despite the evidence we observed for high levels of gene flow and low levels of differentiation across large geographic areas of mainland habitat.

The distinctiveness of mule deer (MD) and black-tailed deer (BTD) has long been recognized and is supported by morphological (Taylor 1956; Wallmo 1981), behavioural (Müller-Schwarze & Müller-Schwarze 1975) and genetic (Latch *et al.* 2009, 2011; Haynes & Latch 2012) data. In this study, patterns of genetic differentiation between MD and BTD were supported by both population-based and individual-based analyses. These two lineages likely diverged in allopatric refugia on

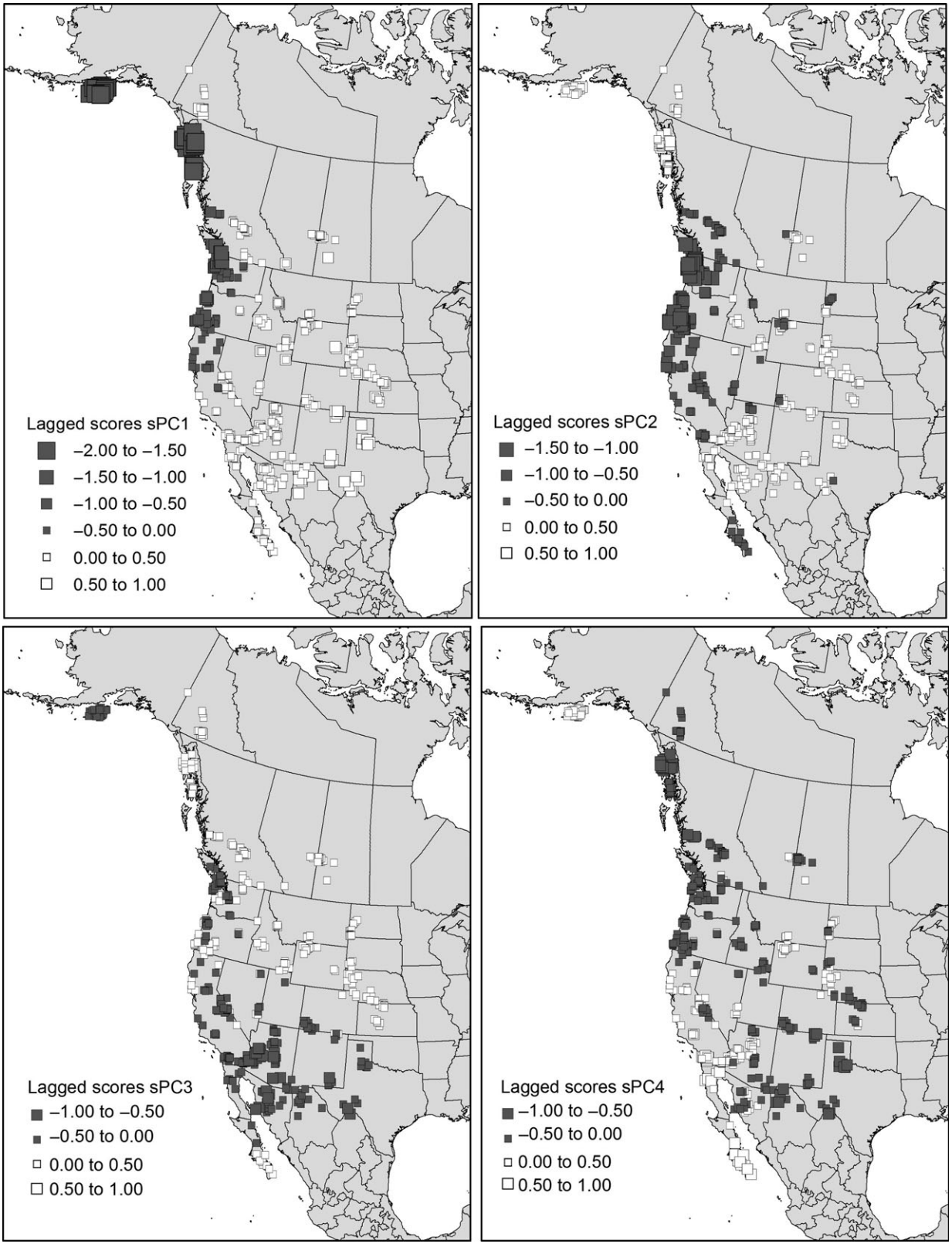

Fig. 6 Plots showing lagged spatial principal components analysis (sPCA) scores for the first four spatial principle components, using a connection network based on 12 nearest neighbours. Each square represents the score of an individual genotype (grey indicates negative and white positive values, and larger squares reflect greater absolute values) and is positioned by its spatial coordinate.

Table 2 Analysis of molecular variance (AMOVA) based on 10 microsatellite loci. Hypothesized groupings represent morphological (Model A–C), haplotypic (Model D; Latch *et al.* 2009), nuclear microsatellite (Model E) and ecological (Model F) variation

Model	Hypothesized groupings	Φ_{SC}	Φ_{ST}	Φ_{CT}	% among groups
A: 10 subspp	[*sitkensis*] [*columbianus*] [*californicus*] [*crooki*] [*eremicus*] [*fuliginatus*] [*hemionus*] [*inyoensis*] [*peninsulae*] [*sheldoni*]	**0.061**	**0.182**	**0.129**	12.93
B: 3 subspp	[*sitkensis*] [*columbianus*] [*californicus*, *crooki*, *eremicus*, *fuliginatus*, *hemionus*, *inyoensis*, *peninsulae*, *sheldoni*]	**0.078**	**0.234**	**0.169**	16.93
C: 2 subspp	[*sitkensis*, *columbianus*] [*californicus*, *crooki*, *eremicus*, *fuliginatus*, *hemionus*, *inyoensis*, *peninsulae*, *sheldoni*]	**0.102**	**0.220**	**0.131**	13.06
D: Hap	[BTD-A] [BTD-B] [MD-A] [MD-B] [MD-D] [MD-E] [MD-F] [MD-H] [MD-I] [MD-J]	**0.095**	**0.164**	**0.076**	7.60
E: 3 clusters	[$BTD_{sitkensis}$] [$BTD_{columbianus}$] [MD]	**0.078**	**0.224**	**0.158**	15.80
F: Ecoregion	[6 – Northwestern Forested Mountains] [7 – Marine West Coast Forest] [9 – Great Plains] [10 – North American Deserts] [11 – Mediterranean California] [12 – Southern Semi-arid Highlands] [13 – Temperate Sierras]	**0.094**	**0.166**	**0.080**	7.97

BTD, black-tailed deer; MD, mule deer. Bold values indicate significant *P*-value ($P < 0.05$). Models B (three morphologically described subspecies) and E (three genetic clusters) yielded the highest variance partitioned among groups.

either side of the Cascade Mountain range during the last glacial maximum (Latch *et al.* 2009). Postglacial recolonization resulted in secondary contact between lineages. Genetic admixture at secondary contact zones is a common consequence of postglacial recolonization from allopatric refugia, and mule deer and black-tailed deer hybridize readily at the zone of contact (Latch *et al.* 2011). This zone of secondary contact along the crest of the Cascade Mountains corresponds with a mammalian contact zone hot spot (Swenson & Howard 2005), suggesting that the crest of the Cascade Mountains served as a strong barrier to gene flow during glacial maxima. In our study, the clearest evidence for admixture between MD and BTD is provided by our STRUCTURE analysis and most easily visualized by the contour lines and kernel boundaries for the genetic surfaces of the ancestry coefficients. Both visualizations indicate a zone of overlapping lineage distributions that correspond to the zone of secondary contact. Despite extensive hybridization between MD and BTD, parental lineages remain genetically distinct and today occupy very different habitats; black-tailed deer in coastal coniferous rainforest west of the mountains and mule deer in grassland/shrubland and forested mountains to the east.

The presence of a third main genetic lineage, separating black-tailed deer into two distinct groups ($BTD_{sitkensis}$ and $BTD_{columbianus}$), is supported taxonomically and by genetic data illustrating considerable mitochondrial sequence divergence between these subspecies (Latch *et al.* 2009). This pattern of divergence is confirmed by our nuclear microsatellite data and by different analytical approaches. The PCA, PCoA and F_{ST} analyses consistently indicated that $BTD_{sitkensis}$ exhibited greater divergence from MD than did $BTD_{columbianus}$. Bayesian clustering analyses also indicated strong divergence of $BTD_{sitkensis}$, and neither approach assigned individuals sampled within the geographic range of *O. h. sitkensis* to other genetic lineages. There exists a gap in sampling between these BTD subspecies, and clustering algorithms may infer barriers that coincide with sampling gaps (Serre & Pääbo 2004; Rosenberg *et al.* 2005; Frantz *et al.* 2009). Similarly, sampling only island populations of $BTD_{sitkensis}$ may have exaggerated the effects of genetic drift resulting from island vicariance. Both sampling artefacts may have yielded overestimates of the amount of genetic differentiation between Sitka and Columbian black-tailed deer. However, our data do not suggest that the observed pattern of genetic differentiation is solely attributable to these sampling artefacts. Genetic differentiation instead may be driven by an overall low level of genetic diversity within $BTD_{sitkensis}$, fitting with a historical bottleneck dating to the origination of the subspecies (Latch *et al.* 2008b, 2009; Colson *et al.* 2013). Morphological and behavioural differences between Sitka and Columbian black-tailed deer subspecies provide further support for evolutionary differentiation over sampling artefacts (Wallmo 1981). Characterizing genetic variation in mainland Sitka black-tailed deer populations and in the unsampled region of northwestern British Columbia and southern Alaska would help to distinguish among these alternative hypotheses.

Within the three main genetic lineages, our data indicated additional spatial genetic substructure. In black-tailed deer ($BTD_{sitkensis}$ and $BTD_{columbianus}$), inferred subclusters corresponded to islands. Island populations encounter a physical impediment to gene flow and are

thus expected to diverge from mainland populations (Mayr 1963). Furthermore, island populations tend to be smaller than continental populations, resulting in a more rapid fixation of neutral alleles on islands and reduced overall levels of genetic variation (Frankham 1997). $BTD_{sitkensis}$ was divided into two genetically distinct clusters associated with the Alexander Archipelago (AK-AI and AK-PW) and Kodiak Island (AK-KI) in Alaska. Black-tailed deer on Kodiak Island were introduced to the island beginning in 1924 and used only 25 animals from the Alexander Archipelago as founders (Smith 1979; Van Daele 2005; Latch *et al.* 2008b). Low levels of genetic variation in our large sample from Kodiak Island and genetic differentiation of Kodiak Island deer from Alexander Archipelago deer are therefore expected as a result of the founding event and subsequent genetic drift. Rapid population growth following the initial introduction to Kodiak Island limited further loss of genetic variation relative to founding populations in the Alexander Archipelago (Latch *et al.* 2008b). However, founding populations in the Alexander Archipelago also exhibit low levels of genetic variation compared with populations from other lineages. Consequently, all $BTD_{sitkensis}$ populations were characterized by very low levels of genetic variation, irrespective of their management history.

The $BTD_{columbianus}$ lineage was separated into two clusters that divided mainland samples and island samples (Vancouver Island, BC-VI, and Gabriola Island, BC-GI). This pattern was reflected in both spatial and nonspatial Bayesian clustering algorithms. The island populations were characterized by low genetic variation compared with mainland populations, similar to the pattern we observed for island populations of $BTD_{sitkensis}$. In addition to low variation at nuclear microsatellites, BC-GI has a single mtDNA haplotype (Latch *et al.* 2009), suggesting that colonization of Gabriola Island likely involved very few individuals. BC-VI exhibited low diversity at nuclear microsatellites but retains haplotype diversity (Latch *et al.* 2009), a signature of population bottlenecks followed by rapid population expansion and accumulation of mutations (Avise *et al.* 1984; Grant & Bowen 1998). Such a pattern would be predicted on an island where deer have been successful following initial colonization and expanded rapidly following population declines in the 1970s (J. MacDermott, personal communication). Overall, mainland $BTD_{columbianus}$ samples exhibited high levels of genetic variation, suggesting that this lineage has maintained robust population sizes without severe declines. Geographically, individuals genetically characterized as mainland $BTD_{columbianus}$ were concentrated in the Pacific Northwest and northwestern California. However, individuals assigned to this lineage were spread throughout the species' range and were found as far south as southern Baja California Sur and east to central Kansas. Such a wide geographic spread of the mainland $BTD_{columbianus}$ lineage might indicate that these alleles represent shared ancestral variation that has not yet experienced complete lineage sorting. This explanation is supported by mtDNA data that indicate very strong differentiation between $BTD_{columbianus}$ and MD lineages (Latch *et al.* 2009) and a complete lack of shared variation. More rapid differentiation of mtDNA haplotypes would be expected given the smaller effective population size relative to nuclear alleles.

Subclustering within the mule deer lineage (MD) was less discrete than for BTD. Similar to the patterns we observed in black-tailed deer, the single island population sampled within the mule deer lineage (SO-TI; Tiburon Island, Sonora, Mexico) exhibited considerable differentiation from the rest of the lineage and severely reduced genetic variation. This population is our sole representative of the morphologically described subspecies *O. h. sheldoni* (Goldman 1939), endemic to Tiburon Island. Although our data indicate considerable genetic differentiation between SO-TI and mainland mule deer populations, the mechanisms driving this divergence (i.e. island vicariance and genetic drift or divergent selection) are unclear (Alminas 2013).

Among mainland MD populations, weak spatial genetic structure was apparent. Iterative STRUCTURE analyses revealed genetic substructure within the MD lineage that was weakly correlated with geography. PCoA also indicated slight aggregation of populations from proximate geographic locations. However, clear boundaries between clusters were not present, and a large portion of the genetic variation within the MD cluster could be explained using a simple IBD pattern of gene flow. In addition, the third axis of the PCoA analysis revealed separation among populations within the MD lineage, accentuating a north–south gradient. A lack of clear boundaries and clinal variation suggests that IBD is a stronger driver of genetic subdivision within the MD lineage than alternative explanations involving behavioural or ecological isolating mechanisms. Bayesian clustering methods do not consider IBD explicitly and thus can overestimate genetic structure when analysing data sets characterized by IBD (Serre & Pääbo 2004; Frantz *et al.* 2009). Although this bias is most often observed when sampling is population based, or when a small number of markers are used, both spatial and nonspatial algorithms superimpose extraneous clusters at high levels of IBD (Serre & Pääbo 2004; Frantz *et al.* 2009; Schwartz & McKelvey 2009). Strong IBD patterns are consistent with the natural history of a widespread, high gene flow species such as mule deer, and suggest a lack of discrete barriers

across large geographic distances encompassed in the mainland MD lineage.

A lack of ecologically based genetic structure across very large geographic areas that we observed for mule deer indicates that historical factors have left a lasting mark on current population genetic structure. This pattern is in contrast to the patterns of spatial genetic structure observed for many other continuously distributed, highly mobile mammals. Contrasting patterns of genetic structure across the landscape may stem from underlying differences in population ecology. Most of the vagile mammals studied to date have been large carnivores, which typically inhabit large home ranges and exist at low population densities. At low population densities, the effect of genetic drift may be strongly relative to gene flow, and even weak effects of landscape permeability on gene flow could play a critical role in generating spatial genetic structure (Wright 1946; Heywood 1991; Vekemans & Hardy 2004). Ungulates, in contrast, are often found at high population densities, where drift is weak relative to gene flow, limiting the development of population genetic structure unless landscape barriers are strong. High historical and contemporary population sizes combined with a lack of strong landscape barriers have been suggested as a potential mechanism for the maintenance of genetic connectivity across large geographic areas in Arctic foxes, a small, highly mobile carnivore with minimal social structure (Carmichael *et al.* 2007), elk (Hicks *et al.* 2007), and in caribou when population sizes are large (Serrouya *et al.* 2012; Weckworth *et al.* 2012).

An absence of genetic structure across large areas of the continental west may also be a reflection of locally influential behavioural responses to the landscape. A strong behavioural response to habitats, such as natal habitat-biased dispersal (Davis & Stamps 2004), might explain why some recent studies have found spatial genetic structure for large and highly mobile mammals in heterogeneous habitats, but not in homogeneous habitats (McRae *et al.* 2005; Pilot *et al.* 2012). For example, landscape-scale structuring in coyotes mirrored habitat subdivision in California, USA, but was not present in more homogeneous habitats in Colorado (Sacks *et al.* 2008). Our data suggest there may be a parallel in mule deer, although the mechanisms driving the observed pattern may be quite different. Habitat-specific population structure is supported at a regional scale in California mule deer, where the pattern of spatial genetic structure is consistent with natal habitat-biased dispersal (Pease *et al.* 2009). Increased spatial heterogeneity of the landscape is also tightly correlated with decreased home range size in mule deer, suggesting an alternative mechanism by which habitat heterogeneity might determine the potential for gene flow in mule deer (Kie *et al.* 2002). Alternative behavioural mechanisms in response to changes in resource availability, such as winter herding behaviour or migratory behaviour, also may contribute opportunities for gene flow in mule deer and help to explain the blurred genetic structure across the mainland range that we observed.

Regardless of the combination of factors that ultimately drive population genetic structure in highly mobile taxa, our species-wide investigation suggests that regional patterns of genetic structure may not translate to other portions of the range characterized by different habitat complexity, such as in interior portions of the western United States. By undertaking a level of sampling typically carried out at a regional scale, and applying it across the entire range, we provide a more comprehensive view of the patterns and potential processes influencing genetic structure in a species. As additional population genetic data sets collected across broad spatial scales and at high resolution become available for large mammals with varied social structures and habitat requirements, we will learn more about the potential drivers of spatial genetic structure in widespread, highly mobile species.

Conclusions

Our results, based on a high-resolution sample of mule deer collected at a broad scale, clearly demonstrate that historical biogeographic events are paramount in the formation and maintenance of spatial genetic structure in mule deer. Both individual- and population-based analyses clearly indicated that historical biogeographic events, specifically Pleistocene glaciation and island vicariance, are the primary mechanisms promoting genetic divergence in this species. Within lineages, very little genetic structure was detected among mainland populations. Panmixia is predicted for continuously distributed, highly mobile species, where unimpeded gene flow is expected to limit the emergence of spatial genetic structure. However, ecological processes often superimpose additional structure onto that generated by historical processes, or erode historical patterns of structure. Although ecologically based patterns of genetic structure have been described for several species, our data indicate a lack of ecologically based structure for mule deer at broad spatial scales. Given the low degree of population structuring observed across much of the interior west, it seems likely that large population sizes and high gene flow have resulted in well-connected populations and an overall lack of genetic isolation. This connectivity may be reinforced by species-specific behaviours such as winter herding and migration, and potentially by weak behavioural responses to a fairly homogeneous landscape. Overall,

these findings highlight the benefit of high-resolution, broad-scale studies of spatial genetic structure.

The hierarchical patterns of range-wide genetic structure in mule deer we observed could have important implications for taxonomy. Based on genetic data alone, it is feasible that mule deer could be better described as distinct populations nested within two (MD and BTD) or three (MD, $BTD_{sitkensis}$ and $BTD_{columbianus}$) main lineages, rather than as 11 distinct subspecies. We urge caution, however, in interpreting our results as evidence that taxonomic revision is warranted, because such considerations must consider multiple criteria in addition to discrete differences in neutral genetic variation, ideally including morphological, ecological, physiological, behavioural and adaptive genetic variation (Haig *et al.* 2006).

Our findings also provide valuable insight into the design of optimal strategies for mule deer conservation and management, which in this case warrant close cooperation among managers given the lack of discrete dispersal boundaries (Heffelfinger *et al.* 2003). This cooperation is particularly important given recent declines in mule deer populations throughout the western United States. If the observed genetic structuring reflects large-scale use of the landscape, then any management action that inhibits movement may have unforeseen negative consequences on mule deer residing far from the location of management activity. Similarly, our results imply that the maintenance of habitat quality and migration corridors over large areas may be critical to the conservation of mule deer populations.

Acknowledgements

We appreciate the support provided by the Boone & Crockett Club, Pope and Young Club, Campfire Conservation Fund of the Camp Fire Club, National Fish and Wildlife Foundation, University of Wisconsin-Milwaukee, Purdue University, University of Arizona, Arizona Game and Fish Department, California Deer Association, Dallas Safari Club, Safari Club International (National and Seattle Chapter), Eldon 'Buck' Buckner, James deVos, Scott Fitkin, Richard Green, Jeff Gronauer, Winifred Kessler, Paul Krausman, Mike Schlegel and Don Whittaker. Samples from the entire range were collected by more than 150 volunteers, which, regrettably, are too numerous to mention individually. Francisco Abarca was instrumental in helping with international coordination. We appreciate helpful comments provided by Lisette Waits and anonymous reviewers on an earlier draft of the manuscript.

References

Alminas OSV (2013) *Phylogeographic inference of insular mule deer* (Odocoileus hemionus) *divergence in North America's desert southwest*. M. S. Thesis, University of Wisconsin-Milwaukee.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge.

Avise JC, Neigel JE, Arnold J (1984) Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *Journal of Molecular Evolution*, **20**, 99–105.

Balkenhol N, Waits LP, Dezzani RJ (2009) Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography*, **32**, 818–830.

Beyer HL (2012). *Geospatial Modelling Environment (Version 0.7.1.0). (software)*. URL: http://www.spatialecology.com/gme. (accessed 1 August 2012).

Bradley RD, Bryant FC, Bradley LC *et al.* (2003) Implications of hybridization between white-tailed deer and mule deer. *The Southwestern Naturalist*, **48**, 654–660.

Brown DM, Brenneman RA, Koepfli K-P *et al.* (2007) Extensive population genetic structure in the giraffe. *BMC Biology*, **5**, 57.

Carmichael LE, Krizan J, Nagy JA *et al.* (2007) Historical and ecological determinants of genetic structure in arctic canids. *Molecular Ecology*, **16**, 3466–3483.

Carr SM, Hughes GA (1993) Direction of introgressive hybridization between species of North American deer (*Odocoileus*) as inferred from mitochondrial cytochrome-b sequences. *Journal of Mammalogy*, **74**, 331–342.

Carr SM, Ballinger SW, Derr JN, Blankenship LH, Bickham JW (1986) Mitochondrial DNA analysis of hybridization between sympatric white-tailed deer and mule deer in west Texas. *Proceedings of the National Academy of Sciences of the USA*, **48**, 9576–9580.

Cathey JC, Bickham JW, Patton JC (1998) Introgressive hybridization and nonconcordant evolutionary history of maternal and paternal lineages in North American deer. *Evolution*, **52**, 1224–1229.

Colson KE, Brinkman TJ, Person DK, Hundertmark KJ (2013) Fine-scale social and spatial genetic structure in Sitka black-tailed deer. *Conservation Genetics*, **14**, 439–449.

Commission for Environmental Cooperation (1997) *Ecological Regions of North America: Toward a Common Perspective*. Commission for Environmental Cooperation, Montreal, Quebec.

Conner MM, Miller MW (2004) Movement patterns and spatial epidemiology of a prion disease in mule deer population units. *Ecological Applications*, **14**, 1870–1881.

Corander J, Marttinen P, Mäntyniemi S (2006) Bayesian identification of stock mixtures from molecular marker data. *Fishery Bulletin*, **104**, 550–558.

Corander J, Siren J, Arjas E (2008) Bayesian spatial modeling of genetic population structure. *Computational Statistics*, **23**, 111–129.

Cowan IM (1936) Distribution and variation in deer (Genus *Odocoileus*) of the Pacific Coast region of North America. *California Fish and Game*, **22**, 155–246.

Cressie NAC (1993) *Statistics for Spatial Data* (revised). John Wiley & Sons, New York.

Cronin MA (1991a) Mitochondrial and nuclear genetic relationships of deer (*Odocoileus* spp.) in western North America. *Canadian Journal of Zoology*, **69**, 1270–1279.

Cronin MA (1991b) Mitochondrial DNA phylogeny of deer (Cervidae). *Journal of Mammalogy*, **72**, 533–566.

Cronin MA, Vyse ER, Cameron DG (1988) Genetic relationships between mule deer and white-tailed deer in Montana. *Journal of Wildlife Management*, **52**, 320–328.

Cronin MA, Nelson ME, Pac DF (1991) Spatial heterogeneity of mitochondrial DNA and allozymes among populations of

white-tailed deer and mule deer. *Journal of Heredity*, **82**, 118–127.

Cullingham CI, Nakada SM, Merrill EH *et al.* (2011) Multiscale population genetic analysis of mule deer (*Odocoileus hemionus hemionus*) in western Canada sheds new light on the spread of chronic wasting disease. *Canadian Journal of Zoology*, **89**, 134–147.

Davis JM, Stamps JA (2004) The effect of natal experience on habitat preferences. *Trends in Ecology and Evolution*, **19**, 411–416.

Derr JN (1991) Genetic interactions between white-tailed and mule deer in the southwestern United States. *Journal of Wildlife Management*, **55**, 228–237.

Dray S, Dufour A-B (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, **22**, 1–20.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

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

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

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

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Foll M, Gaggiotti O (2006) Identifying the environmental factors that determine the genetic structure of populations. *Genetics*, **174**, 875–891.

Francois O, Durand E (2010) Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources*, **10**, 773–784.

Frankham R (1997) Do island populations have less genetic variation than mainland populations? *Heredity*, **78**, 311–327.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.

Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology*, **46**, 493–505.

Gavin TA, May B (1988) Taxonomic status and genetic purity of Columbian white-tailed deer. *Journal of Wildlife Management*, **52**, 1–10.

Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology*, **13**, 2481–2490.

Goldman EA (1939) A new mule deer from Sonora. *Journal of Mammalogy*, **20**, 496–497.

Goudet J (2001) *FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3)*. Available from http://www2.unil.ch/popgen/softwares/fstat.htm (accessed 01 July 2013).

Grant WAS, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, **89**, 415–426.

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

Haig SM, Beever EA, Chambers SM *et al.* (2006) Taxonomic considerations in listing subspecies under the U.S. Endangered Species Act. *Conservation Biology*, **20**, 1584–1594.

Hall ER (1981) *The mammals of North America*, 2nd edn. John Wiley and Sons, New York.

Haynes GD, Latch EK (2012) Identification of novel single nucleotide polymorphisms (SNPs) in deer (*Odocoileus* spp.) using the BovineSNP50 BeadChip. *PLoS One*, **7**, e36536.

Heffelfinger JR, Carpenter LH, Bender LC *et al.* (2003) Ecoregional differences in population dynamics. In: *Mule Deer Conservation: Issues and Management Strategies* (eds DeVos JC, Conover MR & Headrick NE), pp. 63–91. Berryman Institute Press, Utah State University, Logan, Utah.

Heywood J (1991) Spatial analysis of genetic variation in plant populations. *Annual Review of Ecology and Systematics*, **22**, 335–355.

Hicks JF, Rachlow JL, Rhodes OE, Williams CL, Waits LP (2007) Reintroduction and genetic variation: Rocky Mountain elk in Yellowstone and the western states. *Journal of Mammalogy*, **88**, 129–138.

Hoban SM, Borkowski DS, Brosi SL *et al.* (2010) Range-wide distribution of genetic diversity in the North American tree *Juglans cinerea*: a product of range shirts, not ecological marginality or recent population decline. *Molecular Ecology*, **19**, 4876–4891.

Hornbeck GE, Mahoney JM (2000) Introgressive hybridization of mule deer and white-tailed deer in southwestern Alberta. *Wildlife Society Bulletin*, **28**, 1012–1015.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Hughes GA, Carr SM (1993) Reciprocal hybridization between white-tailed deer (*Odocoileus virginianus*) and mule deer (*O. hemionus*) in western Canada: evidence from serum-albumin and mtDNA sequences. *Canadian Journal of Zoology*, **71**, 524–530.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart T, Devillard S, Dufour A-B, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, **101**, 92–103.

Kay CE, Boe E (1992) Hybrids of white-tailed and mule deer in western Wyoming. *Great Basin Naturalist*, **52**, 290–292.

Kie JG, Bowyer RT, Nicholson MC, Borowski BB, Loft ER (2002) Landscape heterogeneity at differing scales: effects of spatial distribution of mule deer. *Ecology*, **83**, 530–544.

Kucera TE (1978) Social behavior and breeding system of the desert mule deer. *Journal of Mammalogy*, **59**, 463–476.

Latch EK, Dharmarajan D, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment

at low levels of population differentiation. *Conservation Genetics*, **7**, 295–302.

Latch EK, Scognamillo DG, Fike JA, Chamberlain MB, Rhodes OE (2008a) Deciphering ecological barriers to North American river otter (*Lontra canadensis*) gene flow in the Louisiana landscape. *Journal of Heredity*, **99**, 265–274.

Latch EK, Amann RA, Jacobson JP, Rhodes OE (2008b) Competing hypotheses for the etiology of cryptorchidism in Sitka blacktailed deer: an evaluation of evolutionary alternatives. *Animal Conservation*, **11**, 234–246.

Latch EK, Heffelfinger JR, Fike JA, Rhodes OE (2009) Species-wide phylogeography of North American mule deer (*Odocoileus hemionus*): cryptic glacial refugia and postglacial recolonization. *Molecular Ecology*, **18**, 1730–1745.

Latch EK, Kierepka EM, Heffelfinger JR, Rhodes OE (2011) Hybrid swarm between divergent lineages of mule deer (*Odocoileus hemionus*). *Molecular Ecology*, **20**, 5265–5279.

Lessa EP, Cook JA, Patton JL (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the late Quaternary. *Proceedings of the National Academy of Sciences of the USA*, **1000**, 10331–10334.

Lopez R (2006) *Genetic structuring of Coues white-tailed deer in the southwestern United States*. M. S. Thesis, Northern Arizona University.

Mantel N (1967) Detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Mayr E (1963) *Animal Species and Evolution*. Belknap Press of Harvard University Press, Cambridge.

McRae BH, Beier P, Dewald LE, Huynh LY, Keim P (2005) Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American puma. *Molecular Ecology*, **14**, 1965–1977.

Miller FL (1974) Four types of territoriality observed in a herd of black-tailed deer. In: *The Behavior of Ungulates and its Relation to Management* (eds Geist V, Walther FR), pp. 644–660. IUCN Publications, New Series No. 24, Morges.

Müller-Schwarze D, Müller-Schwarze C (1975) Subspecies specificity of a response to a mammalian social odor. *Journal of Chemical Ecology*, **1**, 125–131.

Murphy MA, Evans JS, Cushman SA, Storfer A (2008) Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. *Ecography*, **31**, 685–697.

Musiani M, Leonard JA, Cluff HD *et al.* (2007) Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat color and association with migratory caribou. *Molecular Ecology*, **16**, 4149–4170.

Peakall R, Smouse PE (2012) GENALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*, **28**, 2537–2539.

Pease KM, Freedman AH, Pollinger JP *et al.* (2009) Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation. *Molecular Ecology*, **18**, 1848–1862.

Pilot M, Jedrzejewski W, Branicki W *et al.* (2006) Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology*, **15**, 4533–4553.

Pilot M, Jedrzejewski W, Sidorovich VE *et al.* (2012) Dietary differentiation and the evolution of population genetic structure in a highly mobile carnivore. *PLoS One*, **7**, e39341.

Powell JH, Kalinowski ST, Higgs MD, Ebinger MR, Vu NV, Cross PC (2013) Microsatellites indicate minimal barriers to mule deer *Odocoileus hemionus* dispersal across Montana, USA. *Wildlife Biology*, **19**, 102–110.

Pritchard JK, Wen W (2003) *Documentation for STRUCTURE software: version 2*. Available from http://pritch.bsd.uchicago.edu (accessed 15 November 2013).

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

R Development Core Team (2012) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/ (accessed 15 November 2013).

Reding DM, Bronikowski AM, Johnson WE, Clark WE (2012) Pleistocene and ecological effects on continental-scale genetic differentiation in the bobcat (*Lynx rufus*). *Molecular Ecology*, **21**, 3078–3093.

Reimchen TE, Nelson RJ, Smith CT (2003) Estimating deer colonization rate to offshore islands of Haida Gwaii using microsatellite markers. In: *Lessons from the Islands: Introduced Species and What They Tell Us About How Ecosystems Work* (eds Gaston AJ, Golumbia TE, Martin JL, Sharpe ST), pp. 117–120. Proceedings from the Research Group on Introduced Species 2002 Conference, Queen Charlotte City, BC, Canadian Wildlife Service Occasional Papers Series, Ottawa, Ontario.

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

Robinette L (1966) Mule deer home range and dispersal in Utah. *Journal of Wildlife Management*, **30**, 335–349.

Rosenberg NA, Mahajan S, Ramachandran S *et al.* (2005) Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genetics*, **1**, e70.

Rousset F (2008) Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Rueness EK, Jorde PE, Hellborg L *et al.* (2003) Cryptic population structure in a large, mobile mammalian predator: the Scandinavian lynx. *Molecular Ecology*, **12**, 2623–2633.

Sacks BN, Bannasch DL, Chomel BB, Ernest HB (2008) Coyotes demonstrate how habitat specialization by individuals of a generalist species can diversify populations in a heterogeneous ecoregion. *Molecular Biology and Evolution*, **25**, 1384–1394.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

Scribner KT, Smith MH, Garrott RA, Carpenter LH (1991) Temporal, spatial, and age-specific changes in genotypic composition of mule deer. *Journal of Mammalogy*, **72**, 126–137.

Serre D, Pääbo S (2004) Evidence for gradients of human genetic diversity within and among continents. *Genome Research*, **14**, 1679–1685.

Serrouya R, Paetkau D, McLellan BN *et al.* (2012) Population size and major valleys explain microsatellite variation better than taxonomic units for caribou in western Canada. *Molecular Ecology*, **21**, 2588–2601.

Smith RB (1979) History and current status of Sitka black-tailed deer in the Kodiak Archipelago. In: *Sitka Black-Tailed Deer, Proceedings of a Conference* (eds Wallmo OC & Schoen SW), pp. 184–195. Alaska Department of Fish and Game, Juneau, Alaska.

Smith MH, Scribner KT, Carpenter LH, Garrott RA (1990) Genetic characteristics of Colorado mule deer (*Odocoileus hemionus*) and comparisons with other cervids. *Southwestern Naturalist*, **35**, 1–8.

Swenson NG, Howard DG (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *American Naturalist*, **166**, 581–591.

Taylor WP (1956) *The Deer of North America*. Stackpole Books, Harrisburg, Pennsylvania.

Travis SE, Keim P (1995) Differentiating individuals and populations of mule deer using DNA. *Journal of Wildlife Management*, **59**, 824–831.

Van Daele LJ (2001) Unit 8 deer management report. In: *Deer Management Report of Survey and Inventory Activities 1 July 2002–30 June 2004* (ed Brown C), pp. 109–125. Alaska Department of Fish and Game, Juneau, Alaska.

Vekemans X, Hardy O (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.

Wallmo OC (1981) *Mule and Black-Tailed Deer of North America*. University of Nebraska Press, Lincoln, Nebraska.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419–1439.

Weckworth BV, Musiani M, McDevitt AD, Hebbelwhite M, Mariani S (2012) Reconstruction of caribou evolutionary history in Western North America and its implications for conservation. *Molecular Ecology*, **21**, 3610–3624.

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

Wright S (1946) Isolation by distance under diverse systems of mating. *Genetics*, **31**, 39–59.

E.K.L. contributed to research design, performed the research, contributed to the analysis approach, analysis and interpretation of the data and wrote the manuscript. D.M.R. performed much of the data analysis and interpretation and contributed to the writing of the manuscript. J.R.H. contributed to research design, sample collection, analysis approach, data interpretation and contributed to writing the manuscript. C.H.A-G contributed to sample collection and data interpretation. O.E.R. contributed to research design, analysis approach and interpretation of the data.

Data accessibility

Data are available at http://datadryad.org: Dryad doi:10.5061/dryad.ns2jn.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Geographic locations of $n = 1831$ samples subdivided into $n = 71$ sample groups.

Fig. S2 Structure plots for the full data set ($n = 1831$) and each of the three main clusters (MD, $n = 1{,}241$; $BTD_{columbianus}$, $n = 334$; $BTD_{sitkensis}$, $n = 256$).

Fig. S3 Matrix of pairwise F_{ST} values between the 65 sample groups with $n \geq 8$, depicted as a heat map.

Fig. S4 Principal coordinate analysis of the 65 sample groups with $n \geq 8$, colour-coded according to a priori classification to one of ten morphological subspecies.

Fig. S5 Results of BAPS admixture analysis, with each individual depicted as a pie chart reflecting ancestry coefficients (q) for each of the 15 inferred populations.

Fig. S6 Results of the first round of STRUCTURE analysis, showing the 75% contour lines for all three main lineages ($BTD_{sitkensis}$ in blue, $BTD_{columbianus}$ in green and MD in red).

Fig. S7 Map showing all samples assigned with high probability ($q \geq 0.75$) to one of the three main lineages and the corresponding 90% kernels for each lineage.

Fig. S8 Results of spatial PCA analysis, including connection network and screeplot.