Social Structure and Genetic Connectivity in the Southern Mule Deer: Implications for Management

Andrew J. Bohonak, San Diego State University
Anna Mitelberg, San Diego State University

Final Report: April 16, 2014

This is a final report for the project "Social Structure and Genetic Connectivity in the Southern Mule Deer: Implications for Management" to Andrew J. Bohonak from California Department of Fish and Wildlife (CDFW); April 1, 2012 to March 1, 2014; SDSURF Fund 57103A; CDFW grant agreement P1182117).

Background

Due to a lack of resources, land managers often rely on either common sense assumptions or on non-local data to extrapolate the effect of development on species. For example, highly vagile generalists are thought to respond fairly well to urbanization, particularly in presence of ameliatory measures such as wildlife corridors (Alcaide et al., 2009; Mech et al., 2001; Puttker et al., 2008). The southern mule deer, *Odocoileus hemionus fuliginatus,* is a mobile, long-lived generalist assumed to be doing relatively well based on regional wildlife corridor studies and inferences from other deer species in fragmented urban landscapes (CBI, 2002; CBI, 2003; Doerner et al., 2005; Hayden, 2002; Leberg and Ellsworth, 1999; Markovchick-Nicholls et al., 2008; Peles et al., 1999). Prior to this project, Bohonak's graduate student Anna Mitelberg conducted a microsatellite DNA fingerprinting study of 184 scat piles collected from 2005 to 2007 (Mitelberg, 2010). She found evidence for limited dispersal, population structure that corresponds to major freeways, and population bottlenecks within the past 60 years. No long-distance dispersal events were observed, and the southern mule deer herd had lower levels of genetic diversity than other deer lineages within California (Pease et al., 2009). She hypothesized that urbanization (especially freeways) has reduced the frequency and magnitude of long distance dispersal events.

The Mitelberg study only included six microsatellite markers, which limits resolution for many of the analyses that can be conducted in a high-gene flow landscape using population genetic data. Because the statistical power of the study was low (only six loci), it could not be published in the peer-reviewed literature. Currently, studies with 12-30 independent microsatellites are common. Although some conclusions in the original study were well supported (particular those based on individual movement as inferred from genetic fingerprints), statistically supported conclusions about the scale of population differentiation were limited. The goals of this study were to:

1. Improve the laboratory methods to include more markers.
2. Genetically analyze both old and new mule deer samples with the full set of genetic markers.
3. Make management recommendations based on population genetic analyses, including how future changes in connectivity might be detected.

Field methods

Southern mule deer scat was collected at predefined sampling sites within a region of approximately 80 km x 60 km in San Diego County (Figure 1, Table 1). The primary goal of the sampling scheme was to assess movement patterns, particularly in the suburbs north of Miramar, and with particular attention to open spaces and wildlife corridors. These major wildlife corridors run approximately east-west, connecting coastal islands of habitat with inland core areas. Scat samples were obtained by Mitelberg and volunteers from the San Diego Tracking Team. Field protocols are provided in Appendix A (adapted from Valero, 2004; Mitelberg, 2010). We targeted locations where deer signs (such as tracks, scat or browse) had been previously recorded, and animals were tracked until fresh signs, and eventually, fresh scat deposits were found. Between 20 and 30 pellets were collected from each scat pile. To minimize duplicate samples from the same individual on the same date, a maximum of 20 samples were collected per site each day.

Between May 2012 and August 2013, 205 samples were obtained (202 scat piles and 3 tissue samples from road-killed individuals). These new samples supplemented the 168 archived samples collected between April 2006 and April 2007 by Mitelberg (2010). Although sampling was proposed for 22 sites in this project, the actual number of sites was increased to 27 (Figure 1, Table 1). The average distance between two samples collected at the same site was 2.4 km, compared to 20.9 km for pairs of samples from different sites.

Not all of the newly collected samples were of sufficient quality, and only some of Mitelberg's (2010) samples contained enough material for new analysis. As described below, a total of 364 DNA extractions were processed, and genetic data from 240 were found to be reliable enough for analysis. These 240 samples represented 173 individuals (as determined by DNA fingerprinting). The other 67 samples were either same-day duplicates, or recaptures from different days.

Laboratory methods

A. Protocol improvement

Deer pellets were air dried for two days at room temperature, refrigerated, and processed in the laboratory within two weeks of collection. An outer surface wash was performed using phosphate buffered saline solution (1X PBS) sequentially on 12-20 pellets per pile, in the same aliquot of PBS. DNA was then extracted from this concentrated wash using a DNA IQ kit (Promega, #DC6700). To help detect contamination problems should they arise, each day's DNA extractions included a negative control with only PBS. Step by step extraction protocols are provided in Appendix B.

Mitelberg's (2010) study used six microsatellites, and a seventh genetic marker to identify gender (whether the pellet came from a male or female deer; protocol adapted from Gilson et al., 1998). She used a 25 µl polymerase chain reaction (PCR) reaction for each gene, requiring 175 µl per extraction. For this study, we were able to increase the number of microsatellite markers to 15 (optimized from our previous work, as well as Pease et al., 2009). We optimized the PCR protocols and were able to screen the entire set of 15 microsatellites, plus the gender marker, using only one 5 µl reaction. This resulted in a significant decrease in supply costs and technician time.

The final PCR protocol developed for this project was as follows: a single multiplex PCR was carried out using the Qiagen Multiplex PCR Plus kit (Qiagen, # 206152). Each 5 µl reaction contained 1.5 µl DNA, 2.5 µl Master Mix, 0.375 µl of primer mix containing all 16 primer sets in optimized proportions (Appendix C), and 0.625 µl of water. PCR was performed on an Eppendorf Mastercycler: 95°C for 5 min; followed by 37 cycles of 95°C for 30 s, 56°C for 3 min, 72°C for 30 s; and a final 68°C extension for 30 min. Each set of PCR reactions included two positive controls (one male extraction and one female extraction) and two negative controls. Fragment analysis was performed on an ABI 3730xl by Eton Biosciences following submission of a 1.5 µl aliquot of the PCR product, combined with 10 µl formamide and 0.5 µl of GeneScan 500 LIZ size standard (Life Technologies, #4322682). The raw data (chromatographs) returned from Eton Biosciences were interpreted as genotypes using GeneMapper v. 4.0 (Applied Biosystems).

B. Effect of environmental exposure (not contracted)

Although beyond the scope of the contract, we ran an additional experiment to estimate the effect of field conditions on the reliability of scat piles for DNA fingerprints. This experiment was conducted between October and November 2012 using two large mule deer scat piles that were collected fresh, split into multiple treatments, and then aged under field conditions for specific periods of time.

Two freshly deposited scat piles were collected on the morning of October 15, 2012. Each scat pile was "sampled" on October 15 (Day 0 of the experiment). For this experiment, “sampling” consisted of removing five pellets per pile, and performing a pellet wash and DNA extraction as described above. (The only change was that five pellets were washed per sample in the experiment, rather than 12-16 pellets used for field-collected scat.)

After removing fresh Day 0 pellets, those that remained in each pile were split into three equal groups corresponding to three treatments:

1. Standard Protocol control. Pellets were stored indoors at room temperature for two days, and then refrigerated. On days 3, 7, 14 and 6 weeks, pellets were sampled from the refrigerator and processed for genetic analysis. To account for technician error, the Standard Protocol control group was replicated twice for each scat sample, by a different technician.
2. Room Temperature. Pellets were stored at room temperature for the entire course of the experiment. On days 1, 2, 3, 7 and 14, one sample was taken from each scat sample and processed for genetic analysis.
3. Field Exposure. Pellets were stored outside in the SDSU greenhouse for the course of the experiment. The pellets were placed on a bench in an area of the greenhouse with no roof, where they were exposed to sun in the morning and shade in the afternoon. Pellets were "plated out" in perforated plastic containers, raised from the container bottom to allow for drainage in case of precipitation. It rained on day 4 of the experiment. On days 1, 2, 3, 7 and 14, one sample was taken from each treatment and processed for genetic analysis.

Each extraction (44 total, including two negative extraction controls) was then amplified four times for the entire set of microsatellites, and genotyped according to the protocol described above. Q-scores (quality scores) were calculated for each locus and sample as described in Miquel et al. (2006). Individual chromatographs were assigned a score of 0 if they were not successfully scored, or if they resulted in an erroneous genotype when compared with the

consensus genotype. A score of 1 was assigned to all successfully scored chromatographs inferred to be correct based on the consensus.

C. Minimizing genotyping error rates

For field-collected scat in the primary data set, we used repeated analyses on the same sample, two types of negative controls in the PCR reactions, and a variety of computational methods to identify and eliminate errors that are inherent to microsatellite analyses, as well as errors due to sample degradation or contamination. These errors are present in all genetic analyses of natural populations, and when high enough, will lead to erroneous conclusions and interpretations of population structure (see Section E below). Our protocols to minimize error rates meet or exceed those used in nearly all published studies.

Each of the 364 extractions was initially genotyped twice (*Genotyping Round 1*), for a total of 30 chromatographs (15 microsatellite loci x 2 replicates) per extraction. To eliminate poor quality samples, we discarded all extractions where less than 50% of the 30 locus-extraction combinations were successful. For the remaining extractions, PCRs were replicated if there were less than two successful genotypes per locus (*Genotyping Round 2*). Data from these first two rounds of genotyping were combined and analyzed with RELIOTYPE (Miller, 2002), which uses a maximum likelihood algorithm to assess how reliable an inferred multilocus genotype is, and help decide whether further replication is needed. For each sample, each allele was only accepted if it had been observed twice. We accepted all samples with reliability greater than 95%, and limited the sample-wide incidence of false inclusions to less than 5% with a 95% probability. After correcting for multiple tests, a 99.49% reliability was required of each scat pile. Samples failing to pass the reliability criteria were again replicated according to the replication strategy suggested by RELIOTYPE (*Genotyping Round 3*). Samples were discarded from the final data set if they did not yield a reliable genotype after this third round of replicates. Individual PCR cocktails in Genotyping Rounds 2 and 3 included only the necessary loci, rather than all 16.

Of the 364 scat piles analyzed, two contained 3-4 alleles for multiple loci, even after several rounds of genotyping. A diploid individual can have only one or two alleles per locus. We therefore assumed that these samples contained scat pellets from two different individuals, and discarded them from the data set.

We found 240 samples that met the 99.49% reliability criteria. GIMLET 1.3.3 (Valière, 2002) was then used to 1) reconstruct consensus genotypes for each scat pile, and 2) identify single PCR errors: allelic dropout (ADO) and false alleles (FA). Error rates were calculated following Broquet and Petit (2004), and are reported in Table 2.

Throughout the duration of the project, approximately 50 PCR "negative controls" were run (containing all PCR reagents, substituting water for DNA extraction). Five of these showed a chromatogram peak for locus V, but its signal strength was much weaker than any actual samples. An additional 40 "extraction control" PCRs were also run. In these controls, extractions containing no DNA or scat were substituted for normal extractions in the PCR reaction. One extraction control showed evidence of contamination for all loci, but the signal strength was very weak. Following standard laboratory procedure, all reagents were replaced with new aliquots. No additional contamination problems were detected.

D. Identifying recaptures

The 240 consensus genotypes were subsequently regrouped in GIMLET to detect which scat piles had identical genotypes (with missing alleles treated as any other allele). For every pair of scat piles with identical genotypes, the second was labeled a "duplicate" if the piles were collected at the same site and on the same day. The second identical genotype was labeled a "recapture" if it was collected on a different day. We removed 50 duplicates and 15 recaptures detected by GIMLET, leaving 173 individuals with unique genotypes. Thirteen had been recaptured on one occasion, and two on two different occasions. All recaptures occurred within the same population and the same sampling site; no identical genotypes were collected at different sites on the *same* day.

CERVUS v.3.0.3 (Kalinowski et al., 2007) was used to calculate the probability of identity (P_{ID}), which is the probability that two different deer within the study area are identical using the 15 microsatellites from this study.

E. Delineating population boundaries, estimating relatedness, minimizing errors

The vast majority of laboratory artifacts and degraded samples were detected using the methods described in parts C and D above, and not analyzed further. This reduced the final data set to a group of 173 reliable, non-identical samples (Table 1). Following standard procedures, these were analyzed for departure from Hardy-Weinberg genotyping proportions, which are expected in a randomly mating population. Departures from Hardy-Weinberg are typically interpreted as one or more of the following:

1. Many samples are degraded or highly divergent, resulting in incorrect genotypes across many genes due to methodological artifacts (addressed above).
2. One or more genes has incorrect genotypes across many samples, due to methodological artifacts (such as null alleles).
3. The samples analyzed as a single population actually derive from two or more different biological populations (gene pools) that do not freely interbreed. This is known as a Wahlund effect.
4. Samples were taken in a manner that is biased with regards to relatedness (e.g., samples include more close relatives than would be expected in a random sample).
5. Although sampling may be random, mating within the population is not (i.e., there is inbreeding or outbreeding).

To minimize item #3 (incorrectly pooling samples from multiple populations), a series of preliminary analyses focused on objectively delineating population boundaries. This was challenging due to the spatially clustered nature of our sampling scheme (Figure 1) and uneven sample sizes at the 28 sampling sites (Table 1). Preliminary individual clustering analyses were carried out with STRUCTURE v. 2.2.3 (Pritchard et al., 2000; Falush et al., 2003; non-spatial) and BAPS v.5.4 (Corander et al., 2008; spatial information included). The results of these analyses were interpreted in the context of spatial proximity between samples, landscape features (canyons, corridors, open space, freeways) and known aspects of deer behavior and mobility.

We interpreted the STRUCTURE and BAPS analyses in terms of nine putative populations subsumed within two regional gene pools. These correspond to five sites in the western, predominantly urban portion of the study area (western gene pool) and four sites with high potential for connectivity in an eastern gene pool (Table 1, Figure 2).

Locus B consistently failed to amplify in several individuals (even when the PCR reaction did not contain other primer sets). In addition, preliminary results in MicroChecker, v.2.2.3 (van Oosterhout et al., 2004) suggested the presence of null alleles for this locus in two of the nine populations. We eliminated Locus B due to the high likelihood of methodological artifacts (item #2 above), and conducted all further analyses with the remaining fourteen loci. Following removal of Locus B, tests conducted in TFPGA v.1.3 showed no statistically significant departures from Hardy-Weinberg genotype proportions in the nine populations (Miller, 1997; conventional Monte Carlo Method, 100 batches/analysis, 2000 permutations/batch).

The over-representation of relatives in a sample violates the random sampling assumption of most population genetic analyses (item #4 above). Individual-based clustering software such as STRUCTURE is particularly sensitive to this violation (Rodriguez-Ramilo and Wang, 2012). Female mule deer rarely travel far from their natal sites, and tend to cluster near their immediate female relatives when they do disperse (Porter et al., 1991). Because of this social structure, our sampling sites are very likely to contain relatives. To circumvent this issue, the pedigree reconstruction software COLONY v.2.0.5.0 (Wang, 2004; Wang and Santure, 2009; Jones and Wang, 2010) was used to identify and remove full siblings and parents. For each locus, COLONY requires ADO and FA error rates (Table 2). Because our consensus genotypes are estimated to be 99.49% reliable from RELIOTYPE (see section C above), we multiplied the ADO and FA error rates in Table 2 by 0.0051 for the COLONY analyses. We considered both genders to potentially be polygamous, and all 173 individuals to be both potential parents and potential offspring.

When analyzing the 173 unique individual data set, COLONY detected 36 full sibling "dyad" pairs. For each of these dyads, one sibling was randomly removed. (There were two additional sibling pairs in which both individuals were retained, because they occurred in different sampling sites, and the removal of either would have made the site-level sample size too low for further analysis.) We removed an additional six individuals which COLONY identified as parents of offspring in the data set, and which were not removed with the full sibling dyad screen. This resulted in an “unbiased data set” of 133 relatively unrelated individuals that was used in further analyses.

F. Intrapopulation diversity statistics

A final test for deviations from Hardy Weinberg genotype proportions was conducted on the "unbiased data set” in TFPGA (133 individuals, 14 loci, conventional Monte Carlo Method, 100 batches/analysis, 2000 permutations/batch). FSTAT v.2.9.3.3 (Goudet, 2001) was used to calculate allele frequencies, allelic richness, heterozygosity, F_{IS}, and to test for significant deviations from non-random mating (126,000 randomizations).

G. Divergence among populations

Two types of approaches were used to verify that the nine putative populations were genetically unique. Because assessment of population structure is a complex task, the consensus of multiple analyses based on different algorithms suggests that any population structure is likely to be "real" (Pearse and Crandall, 2004). First, contingency tests were used to assess whether allele frequencies were significantly different among all possible population pairs (Raymond and Rousset, 1995; also referred to as "exact tests of population differentiation"). This analysis was conducted in TFPGA (2000 dememorization steps, 20 batches, 2000 permutations per batch) on the unbiased data set. To quantify which populations were the most divergent, Wright's (1931) F_{ST} was estimated among all possible population pairs with FSTAT v.2.9.3.3 (Goudet, 2001).

Second, we implemented an individual-based clustering analysis using STRUCTURE, with and without sampling site as a prior (K = 1-10 clusters; 3 replicates per K; 100,000 burnin, 500,000 MCMC replicates after burnin; admixture model; Hubisz et al., 2009). STRUCTURE HARVESTER (Earl and vonHoldt, 2012) was used to implement the Evanno et al. (2005) method for inferring the optimal number of clusters from multiple runs.

H. Fine-scale population structure, including relatives

As mentioned above, female mule deer rarely travel far from their natal sites, and tend to cluster near their immediate female relatives when they do disperse (Porter et al.'s 1991 "rose petal hypothesis"). To test for the presence of rose-petal fine-scale population structure in our mule deer population, we conducted tests of genetic spatial autocorrelation patterns among the 173 individual data set, using GENALEX 6.501 (Peakall and Smouse 2006, 2012). This analysis calculates spatial autocorrelation values (i.e., multivariate genetic distances) within predefined distance classes. Genetic distance between each pair of individuals in the study is plotted against geographic distance classes (also known as lags or intervals). For example, the distance class of 0.6 - 3.0 km would include genetic distances between all pairs of individuals (or for this study, scat piles) separated by 600 m to 3000 m. Following Beck et al. (2008; but using individual data, rather than site level data), we used the nearest neighbor option in GENALEX and a frequency plot of pairwise distances to decide on meaningful distance classes; both 0.1 km and 0.6 km were tried as the first distance class for comparison. Because the results were similar, we present only the 0.1 km results here. Based on recommendations by Banks and Peakall (2012), spatial autocorrelation was analyzed at multiple levels: the entire data set, western sites only, and eastern sites only. We analyzed males and females separately to test for sex-biased dispersal. Statistical significance was assessed using three test statistics:

1. Rather than raw genetic distances, the autocorrelation coefficient *r* was used (*r* is bounded by [-1,1] and is closely related to the common autocorrelation statistic Moran's *I*). Confidence intervals (95% CIs) were calculated using 1000 bootstraps for each distance class. The null hypothesis is overlapping CIs between the sexes, particularly in the first distance class. The alternative hypothesis is nonoverlap of the CIs, with correlations significantly greater in the philopatric sex (females).
2. A squared paired-sample t-test statistic 'T2' was calculated for each distance class, and significance was assessed with 1000 bootstraps.

3. A correlogram wide 'Omega' (ω) was calculated across all distance classes. The null hypothesis is homogeneity between the spatial correlograms of the two sexes, and the alternative is heterogeneity. Significance was assessed with 1000 bootstraps.

Per Banks and Peakall (2012), the heterogeneity tests (T2 and ω) were considered statistically significant at a p-value of less than 0.01.

I. Effective population size (N_e)

Based on the frequencies of full and half sib dyads identified in COLONY, this program estimates the current effective population size (methods from Wang 2009). We also estimated N_e using a linkage disequilibrium (LD) method (Hill, 1981; Waples, 2006; Waples and Do, 2008, 2010) implemented in NeEstimator v2 (Do et al., 2014), at three levels of genetic structure: 1) all data, 2) regional gene pools, and 3) the nine populations.

Results

A. Protocol improvement

The improved protocols are described in Laboratory Methods above, and the Appendices to this report. They will be included in published papers that derive from this project.

B. Effects of environmental exposure

The quality of microsatellite genotyping was very high for all treatments in the environmental exposure experiment, with 2256 of 2175 locus-pile-treatment-day combinations successfully scored. Only two genotyping errors were encountered, and both were instances of allele dropout in the Standard Protocol after six weeks.

Q-scores representing the proportion of successful genotyping reactions are summarized in Figure 3. We found that the three treatments did not appreciably affect reproducibility of the DNA fingerprinting. Even two weeks of exposure in field conditions had no detectable effect. This curious result should be verified with additional experiments, as it contrasts with our previous assumptions (and the current literature) that DNA quality degrades rapidly within the first week. We also note that the experiment was based on only two replicate scat piles. It is possible that our current laboratory protocols have improved enough that older samples are now usable in many cases. However, we are also unsure whether the open air greenhouse at SDSU sufficiently mimics typical field conditions.

C. Minimizing genotyping error rates

The methods used to minimize error rates and the corresponding results are presented in Laboratory Methods above. The 364 scat piles we analyzed genetically were reduced to 240 for further analysis that met an extremely high reliability threshold of 99.49%. "Reliability" can be interpreted as multi-locus accuracy: 99.49% of replicate scat samples from these same individuals would be assigned identical genotypes for every microsatellite locus if they were processed and analyzed in an identical manner.

D. Identifying recaptures

The genetic markers we used were able to identify individual deer with very high precision. The probability of identity (PID) was estimated to be 1.07×10^{-7} (Table 2). This is

the probability that two deer within the study area would be identical using the 15 microsatellites from this study. Even for full siblings, the expected PID_{sib} was still very low: 5.15×10^{-5} (Waits et al., 2001).

As summarized above, the final sample set of 240 scat piles included 50 same-site/same-day duplicates, and 15 "recaptures". Thirteen individuals were recaptured once, and two were recaptured on two subsequent occasions. All recaptures occurred close to the original sample, with a mean distance of 342 m, and maximum slightly greater than 1 km (Table 3, Figure 4).

E. Delineating population boundaries, estimating relatedness, minimizing errors

The putative population boundaries identified during preliminary analyses are depicted in Figure 2 and listed in Table 1. Here, we summarize their locations and inferred connectivity for mule deer (site abbreviations correspond to Table 1):

- TP-SV (Torrey Pines-Sorrento Valley: TP, SV) comprises a small area including Torrey Pines State Reserve (TPSR), Penasquitos Marsh Natural Preserve and a small portion of Sorrento Valley (SV) just south of the I-5 / I-805 merge. This area is under continuous pressure from development. It is difficult to tell whether there is present day connectivity between TP and SV, but if movement does occur between these sites, it is likely to be along the Sorrento Valley Corridor (see below). These sites were pooled into one population based on 1) physical proximity, 2) their isolation from sites east of I-5 and I-805, and 3) preliminary spatial clustering analyses in BAPS, which found these two sampling sites to be a unique and exclusive cluster. Crooks (1997) concluded that the Sorrento Valley Corridor is the only remaining functional corridor between TPSR and habitat east of the reserve (CDPR, 1998). The SV corridor connecting Los Penasquitos Canyon Preserve (PQ) with TP corresponds with SDTT's Transect 17, The Merge Loop, and is described in the Torrey Pines Wildlife Management Plan (CDPR,1998: the southern route of the SV Corridor starts on the west side of Los Penasquitos Canyon, passes under I-805 and I-5, goes past the J & W Lumber Co. on Sorrento Valley Road, goes under Sorrento Valley Road, and ends in Los Penasquitos). SDTT reports recent deer activity in this corridor, including scat, sign, browse, and sightings (Kramer, pers. comm.).
- The PQ-PC (Penasquitos Canyon – Penasquitos Creek: PQ, PC) population is a large area spanning Los Penasquitos Canyon Preserve (PQ) and Penasquitos Creek (PC). PQ is bordered by I-805 on the west and I-15 on the east, is composed of high quality habitat, and is surrounded by development. PC located east of I-15, and is well connected to PQ via the Penasquitos Creek trail which passes under I-15.
- CC-Mir (Carrol Canyon-Miramar: CC, Mir-EM) includes samples collected in Carrol Canyon (CC) and the small Eastgate Mall parcel that is part of MCAS Miramar (Mir-EM). Because Eastgate Mall is separated from the remainder of MCAS Miramar by the heavily used six-lane Miramar Road, it probably has more opportunity for connectivity with Carrol Canyon (Black, pers. comm). CC-Mir is a highly urbanized area. There is some potential for connectivity between this population and TP-SV, or at least SV (see COLONY results below for evidence). The connectivity route would start in Penasquitos Lagoon, move south along Sorrento Valley Road and the railroad tracks, and cross under I-805 approximately 2.75 km south of the SV Corridor (see above). The San Diego Tracking Team has no transects in this area (Kramer, pers. comm.).

- MirNW (Miramar Northwest: MirA plus MirG-C) is a small population sampled at two sites along Rose Canyon, which is one of two primary east-west wildlife corridors on MCAS Miramar. This population is bordered to the north by Miramar Road and to the south by Main Station. Preliminary clustering analyses in STRUCTURE suggested this population is genetically distinct from CC-Mir samples located to the north, and San Clemente Canyon samples collected to the south.
- MT-Mir (Mission Trails-Miramar: MirD, MirG, MirE, TS, MT) is a large population that includes five sites scattered either directly along the San Clemente Canyon corridor, or in adjacent connecting canyons (Oak and Spring Canyons). This includes samples collected near Miramar Landfill (MirG), on MCAS Miramar's Parcel G (MirD, east of I-15 and south of State Road 52), in eastern MCAS Miramar (MirE), in Mission Trails Regional Park (MT) and in Tierrasanta (TS, near Admiral Baker Golf Course). Although this population is bisected by I-15 (which runs north-south) and State Route 52 (which runs east-west), preliminary clustering analysis in STRUCTURE showed these sites to be composed of genetically similar individuals.
- BC-SC (Beeler Canyon-Sycamore Canyon: BC, SC) incorporates samples collected in Upper Beeler Canyon (BC), and others that are within or near Sycamore Canyon Open Space Preserve (SC). Although BC is relatively close to PC, these sites are separated by urban development along Pomerado Road, and west of Pomerado Road.
- NW (Northwest: CP, RLC) includes samples from Calavera Preserve (CP), Rancho La Costa Habitat Conservation Area (RLC) and the Wild Animal Park (WAP). The WAP was grouped with CP and RLCHCA due to low sample sizes for all three sites, and relative proximity to each other, when compared with distance from other sites.
- The NE (Northeast: JL, CNFn, CRSP, LJ) population incorporates all samples collected north of I-8 and south of SR-78. Sampling took place in Julian, Cleveland National Forest, Cuyamaca Rancho State Park, and Lake Jennings. With the exception of Lake Jennings, which is developed on the south end, the remaining sites are primarily uninterrupted open habitat.
- SE (Southeast: CNFs, HC, RJER, SM, SoC, SP) comprises all samples collected south of I-8 and east of I-125. Sampling for this population took place at San Miguel Summit and Mother Miguel Mountain Area (SM), Sycuan Peak (SP), South Crest property (SoC), Rancho Jamul Ecological Reserve (RJER) and Hollenbeck Canyon Recreational Area (HC). This large area is largely undeveloped, with the exception of SM and SoC.

NW, NE and SE include populations that were outside of the primary sampling area for this project. It is possible (and perhaps likely) that higher density sampling in these areas would result in more localized population boundaries, similar (in terms of area) to those west of I-15.

Based on genetic similarity, the algorithm in COLONY inferred 42 dams (mothers) and 40 sires (fathers) for the 173 individuals in our data set (Table 4). Only 10 of these 82 parents were present in our samples. The other 72 are inferred based on genetic similarities or differences among the 173 individuals, when each is analyzed as an offspring. The mean family size (number of offspring) was 4.1 for dams and 4.3 for males (with ranges of 1-8 and 1-11 respectively). Because this only represents our data set, it is an underestimate of total fecundity. Within our data set, dams were inferred to have offspring with 1-6 sires each (mean = 3.5), and sires were associated with 1-8 dams each (mean = 3.6).

There were 20 individuals within the study whose parents were also sampled: 14 females, 5 males, and 1 of unknown gender. The 20 offspring were associated with 10 parents (Table 5). The distribution of parent-offspring dyad distances is presented in Figure 6 (top). The majority of parent-offspring dyads were found within 2 km of each other (and 17 / 20 within 5.1 km). Ten of the 14 females were sampled in the same site as their mother or father, and two were sampled at a different site in the same population. The twelfth female offspring was sampled in Parcel G on MCAS Miramar (population MT-Mir), 5 km from its mother at the very northwestern corner of MCAS Miramar (population CC-Mir). The thirteenth female offspring was sampled from NE, approximately 50 km from its mother in MirNW (see Figure 5). The offspring of unknown gender was also sampled in the same site as its mother. In contrast to the females, only one of five males with an identified parent was sampled in the same site or population as its parent. The second and third were sampled from populations adjacent to their mother, the fourth in PQ-PC (5 km away from its mother in MirNW), and the fifth in SE (38 km away from its mother in MirNW).

The majority (26 / 36) of full sibling (FS) dyads were found within the same site. (Note that "sites" were defined such that the average distance between two samples within a site was 2.4 km.) The distribution of FS dyad distances is presented in Figure 6 (bottom). Interestingly, nine of the ten FS dyads not within the same site were in *different* populations (Table 6). Four FS dyads were identified between SV and MirEM, supporting the probability of a functional Sorrento Valley Corridor, as described above. One dyad was identified between MirEM and MirA, suggesting deer are able to cross Miramar Rd. This is further supported by two FS dyads occurring between CC and MirG. The PQ-MirD dyad represents two males, with ~7.6 km between them. The individual in MirD (Parcel G) was a solitary male, captured in May in an area with no evidence of other deer activity, and may represent a dispersal event in progress (which typically occurs when male yearlings reach the age of 1-1.5). The PQ-HC dyad represents two females captured approximately 43 km apart, and is an unexpected result.

F. Intrapopulation diversity statistics

Genetic diversity is summarized by locus and by population in Table 7. The majority of the microsatellite markers were highly polymorphic ($H_e > 0.5$), with an average of 4.8 alleles per locus. Genetic diversity was very similar at each site, with mean H_e ranging between and 0.43 and 0.56, and mean A between 2.7 and 4.0. No strong locus- or population-specific patterns were evident in terms of heterozygosity deficit or excess. Of 126 tests conducted for Hardy-Weinberg genotype frequencies, only 6 were significant at $p < 0.05$, and none were significant after Bonferroni correction (Table 7). Similarly, only four tests for deviations from random mating patterns (null hypothesis $F_{IS} \neq 0$) were significant at $p < 0.05$, and none were significant after Bonferroni correction.

Based on these results, the low locus-specific error rates (Table 2), and the high reliability of the consensus genotypes, we infer that few if any methodological biases remain in the final data set (e.g., null alleles, allelic dropout).

G. Divergence among populations

The operational definition of a biological population (and related concepts such as management unit or "stock" in marine fisheries) varies depending on the research questions or

management goals (e.g., Waples and Gaggiotti, 2006). For interpreting intraspecific genetic data, we regard a population as a geographic unit that is statistically unique based on:

1. population-based analyses in which the null hypothesis of panmixia is rejected (e.g., traditional pairwise F-statistics, contingency tables). These analyses are conducted on all pairwise combinations of putative populations.
2. individual-based Bayesian clustering algorithms (e.g., STRUCTURE, BAPS). These analyses are conducted without predefining populations, collection sites, or other groups.

For the mule deer samples, all 36 contingency table tests between population pairs were significant at $p < 0.05$ (Table 8). Thirty-one were significant even after Bonferroni correction for 36 tests ($p < 0.0014$). After Bonferroni correction, CC-Mir was not significantly different from TP-SV and MirNW; PQ-PC was not significantly different from MirNW; SE was not significantly different from the NE and NW populations. F_{ST} analyses showed similar patterns (Table 8). All pairwise F_{ST} estimates were significantly greater than zero ($p < 0.05$) except NW vs. SE, and 28 of 36 were significant after Bonferroni correction. Divergence was highest ($F_{ST} > 0.1$) between TP-SV at the western edge of the study area, and the three furthest populations to the north and east (NW, NE, SE).

The optimal model for the individual-based clustering method in STRUCTURE included two gene pools, regardless of whether sampling site was included as a prior in the Bayesian analysis. The relative contribution of these two gene pools to individual genetic composition varied dramatically between populations in the southwestern part of the study area, and those to the north and east (Figure 7). All but one individual from the five western populations had >50% membership to the "western gene pool" (black portion of bars in Figure 7). Individuals from TP-SV had a genetic background that was ≥ 97% derived from the western gene pool. PQ-PC, CC-Mir, and Mir-NW were very similar, with western gene pool ancestry varying between 67%-96%, 78%-97%, and 54%-88% respectively. As a transitional zone, MT-Mir showed the widest range of individual genetic compositions (41%-93%). To the north and east of MT-Mir, all but three BC-SC individuals all had genetic ancestry that was > 50% derived from an "eastern gene pool". Individuals from NW, NE and SE also had genetic backgrounds that were primarily associated with the eastern gene pool (50%-97%). Overall, these results demonstrate:

1. General concordance between the STRUCTURE results, contingency table and F_{ST}-based analyses.
2. A clinal change from east to west in individual genetic composition.
3. MT-Mir as a transitional zone in the east-west genetic cline.
4. An abrupt break in the cline between BC-SC, and populations to the west.
5. Pronounced isolation of populations in the Northwest (NW) from those 20 km to the south in PQ-PC.

H. Fine-scale population structure, including relatives

The spatial autocorrelation analysis showed that individuals are more closely related than one would expect based on random spatial associations, up to a distance of 3-8 km (depending on the gender and analysis; Figure 8). In the western portion of the study area where sampling was the most dense, 3-8 km corresponds to the approximate north-south distance between

canyons (which run parallel to each other from west to east; Figure 1). It also corresponds to the approximate distance between I-5/I-805 at the west end of each canyon, and I-15 to the east. Beyond this distance, mule deer show approximately random relatedness, except perhaps for 14-21 km. In this distance range, individuals in both the Eastern and Western gene pools may be less related than one would expect from random spatial associations (Figure 8).

We used three statistical tests to interpret the spatial autocorrelation analyses in terms of sex-biased dispersal. Each indicated that over very short distances, females are more closely related to one another than males. (However, different subsets of the analyses were statistically significant in each case.)

1. The presence of non-overlapping CIs in the first distance class is the most conservative measure for detecting sex-biased dispersal (Banks and Peakall, 2012; Figure 8, top). This first test was only significant for our analysis of the Western gene pool when the first distance class was set to either 0.1 km or 0.6 km.
2. The T2 test showed statistically significant differences between male and female autocorrelations in the Western gene pool when first distance class was set to 0.6 km ($p = 0.005$).
3. The Omega test only detected significant differences when the entire data set was analyzed as one unit, but not for the individual gene pools.

I. Effective population size (Ne)

Across the study area, estimates of N_e were 74 (95% CI = 53-104) for a population mating at random and 68 if the random mating assumption is relaxed (95% CI =48-95). Although this number seems low, N_e in natural populations is only a fraction of the *census* population size that you would estimate in a demographic or mark-recapture study (typically 50% - 90% less; Frankham, 1995). It is also possible that our N_e estimate is biased downward, because our samples represent multiple generations of mule deer (rather than a single cohort). Because multiple generations are represented, some of the individuals inferred to be full siblings could actually represent parent-offspring pairs.

The LD method for estimating N_e suggested a similar value of 70.8 (95% CI = 59.1 - 85.7) for the entire data set. However, Neel et al. (2013) showed that the LD method provides a good approximation only when the scale of sampling corresponds to the scale of local breeding. Given the strong signal of spatial autocorrelation in our data and the presence of two gene pools, we also estimated N_e using the LD method for each gene pool, and for each of the nine local populations. N_e for the Western gene pool is 59.1 when it is analyzed as one unit (95% CI 46.5 - 76.9), or 65.9 for the sum of its five populations analyzed separately. The overall estimate of N_e for the Eastern gene pool is 41.7 (95% CI = 31.4 - 57.9), or 108.4 for the sum of its four populations from separate analyses. However, we assume that the Eastern estimates are highly inaccurate due to sparse sampling throughout that portion of the study area.

In summary, effective population size for the area of this study is estimated to be 68-174, depending on how samples are pooled for analysis.

Management recommendations

A. Prevent regional declines in genetic diversity

Corroborating the results of Pease et al. (2009), we found lower genetic diversity in the southern mule deer than other subspecies of mule deer in California. Pease et al. reported statewide genetic diversity metrics of A = 7.5 and H_e = 0.66. In contrast, we found A = 4.8 and H_e = 0.58 in San Diego County. This indicates that southern mule deer in San Diego County have lower effective population sizes than other subspecies to the north. We estimate effective population size to be approximately 100 or fewer individuals in each of two regional gene pools. A more detailed simulation analyses (using Approximate Bayesian Computation or similar approaches) might provide insight for how extreme and recent any past population bottleneck may have been. It is possible that this subspecies has always had relatively small population sizes -- even prior to modern urbanization.

Regardless of how small pre-development *Odocoileus hemionus fuliginatus* populations were, we expect future habitat loss and population size reduction in San Diego County to have a proportionally larger effect than it would for deer elsewhere in the state. Inbreeding effects will increase if population size continues to decrease. We also note that 1) there has been a recent decline in southern mule deer sign and sightings in urban parks (Martin, pers. comm.), 2) the U.S. distribution of this subspecies is limited for the most part to San Diego County, and 3) even without the pressures of urbanization, southern mule deer are predicted to undergo a range contraction due to climate change and encroachment from other subspecies (Pease et al., 2009). Southern mule deer are non-migratory, so that unlike migratory deer that may divide their time between urbanized and rural/undeveloped areas, this subspecies experiences pressures from development year-round (Sommer et al., 2007). Urbanization can prevent mule deer from using all of the available good quality habitat, and lower fawn recruitment (McClure et al., 2005).

- We recommend that further urban encroachment be minimized, and that connectivity and local population sizes be maintained at least at their current levels, or increased where feasible.

B. Define multiple management units for southern mule deer

Our analyses demonstrate that southern mule deer are relatively sedentary in comparison with most other large mammals. In the absence of dramatic alterations to their habitat, individuals remain in the same local areas for many years (perhaps their entire adult life), and offspring tend to remain close to one or both parents. The data suggest that multiple gene pools exist in the study area. However, gene pools may not necessarily be correspond to units for management. In addition to "gene pool" and "population", a variety of concepts are associated with spatial units that require individualized conservation or management. These include stock, distinct population segment, and evolutionarily significant unit (e.g., Ryder, 1986; Moritz, 1994; USFWS and NMFS, 1996). Moritz (1994) attempted to focus the diversity of opinions about these definitions on management units (MUs): "populations connected by such low levels of gene flow that they are functionally independent". Diagnosis of MUs is usually based on statistically significant divergence for allele frequencies in a population genetic data analysis. However, MU designation based on genetic data must be mindful that:

1. Different population genetic analyses implicitly integrate gene flow over different temporal scales.

2. Contemporary gene flow is unlikely to equal historic gene flow, especially in species of concern for conservation or management.
3. The movement of individuals between populations does not always result in interbreeding and the transfer of gene copies (Bohonak, 1999; Whitlock and McCauley, 1999; Palsbøll et al., 2007).

Based on the traditional MU approach, our genetic data suggest that southern mule deer be managed as at least two units (a Western gene pool and an Eastern gene pool), and possibly as many as nine (Table 1). Even adjacent canyons may harbor genetically unique sets of individuals. Analysis of additional samples collected from new areas in future work would undoubtedly add additional MUs to the subspecies.

We acknowledge that there are other approaches to delineating management units. Palsbøll et al. (2007) made the reasonable suggestion that MUs be defined using a specific rate of gene flow or dispersal as a threshold, rather than statistical rejection of a null hypothesis. (Failure to reject the null hypothesis of no population structure could be because the data set has insufficient statistical power.) MUs defined by rates of gene exchange are likely to be more objective, but would require an independent estimate of effective population size, and probably generation time as well.

- We recommend that southern mule deer in an area that includes Torrey Pines, Sorrento Valley, Penasquitos Canyon, Penasquitos Creek, Carrol Canyon, MCAS Miramar and Mission Trails be considered as a separate management unit from those elsewhere in the subspecies range. This corresponds to the "Western gene pool" from this study.
- We recommend that management units more spatially restricted than these two gene pools also be considered.
- We recommend that genetic analyses be conducted in other areas of the species range under threat of habitat conversion, comparable in spatial and temporal coverage to this study.
- We recommend that the existing literature be combined with new demographic studies to provide accurate estimates of population size (both N_e and census size), and generation time. These estimates will improve interpretations of current and future genetic data.

C. <u>Maintain high levels of connectivity.</u>

Southern mule deer have relatively small territories and home ranges. Consistent with expectations from Porter et al.'s (1991) rose petal hypothesis, females show a pattern of intense fine-scale spatial autocorrelation up to approximately 8 km. Male spatial autocorrelation is also statistically significant, although less strong than in females. This suggests that male gene flow is limited within the study area, even though males are the dispersing gender in this polygamous mating system (see also Figure 6 and Table 5).

Another test of the rose petal hypothesis comes from the locations in which relatives were found. The term "rose petal" was invoked because offspring (particularly female) tend to set up territories immediately adjacent to their mother, like petals surrounding the center of a flower. We found that 12 of 20 offspring from known parents were sampled in the same local population as their parent (Table 5), as expected under this hypothesis.

Consistent with the autocorrelation and relatedness analyses, the mean distance for 16 recapture events was only 342 m, and one female in Sorrento Valley from the original 2006 field

season was recaptured in 2013 less than 100 m from the original location (Table 3). (Qualitatively similar results were noted by Valero, 2004 and Mitelberg, 2010.) This suggests that southern mule deer do not readily relocate territories.

There is high connectivity within the nine population boundaries we identified (Figure 2). As discussed elsewhere in the report, these populations correspond broadly to known corridors, preserves, or contiguous areas of open space. Beyond those affiliations, two lines of evidence suggest that Carrol Canyon (at the west end, near I-805) is a key hub in the regional gene pool. First, the full sibling (FS) location data (Table 6) show four FS dyads between Sorrento Valley just west of Carrol Canyon, and Miramar / Eastgate Mall south of Carrol Canyon. There is not an alternate route for movement between SV and MirEM, since the area west of MirEm is completely developed. Second, there are FS dyads between Carrol Canyon and other areas of Miramar (MirE and MirG). If mule deer move between SV and CC under I-805 just south of Mira Mesa Boulevard, this would provide a critical corridor between coastal populations such as Torrey Pines, and the large amount of open space on MCAS Miramar. As far as we know, the San Diego Tracking Team does not currently conduct surveys in this area. In addition to providing a large amount of undeveloped habitat, Miramar seems to be a well connected set of source populations (see parent-offspring locations in Figure 5). Although managed as one regulatory unit, Miramar seems to provide sufficient habitat for multiple MUs (management units).

- We recommend that existing levels of habitat connectivity in western San Diego County be maintained, in light of limited lifetime movement that appears to be typical of southern mule deer.
- We recommend that additional non-genetic studies be conducted to quantify mule deer movement between SV and CC.
- We recommend that additional genetic studies of mule deer be conducted on MCAS Miramar, to clarify its role in regional conservation of this species.

Conclusions

With regards to microsatellite markers, we have shown in this study that:

1. DNA of sufficient quality for reliable, high resolution genetic analyses can be obtained from mule deer scat.
2. The microsatellites we have optimized provide a very accurate individual DNA fingerprint. Even for full siblings (same mother and same father), the probability of an identical fingerprint is only 5×10^{-5}.
3. Gender, parent-offspring relationships, within-generation movement and cross-generation movement can be inferred from the microsatellite markers.
4. Laboratory costs for microsatellite analysis have been reduced over the course of this project. Future costs to process scat or tissue samples in batches of 96 would be approximately $50-$75 per sample for an academic lab. (We are unsure of costs in a high throughput commercial laboratory). This includes supplies (with larger projects achieving economies of scale), outsourced genotyping services, personnel time and associated fringe. This estimate

does not include overhead (indirect), and it does not include personnel costs associated with field collections.

We have inferred from analyses of these markers that:

5. Southern mule deer have less overall genetic diversity than subspecies elsewhere in the state. This is consistent with an effective population size that is less than 200 individuals for the region we studied (up to 500 km2), and perhaps less than 100.
6. Southern mule deer are relatively sedentary/territorial over many years. Offspring are often found very close to one or both parents.
7. Females in close proximity tend to be more closely related than males in close proximity.
8. There is statistical justification for dividing the area covered by this study into 2-9 management units. In the western part of our study area, where sampling was the most dense, populations generally correspond to existing reserves and canyons.
9. As in prior studies, the isolating effects of I-5 north of the I-5/805 merge are apparent, as well as the isolating effects of I-805 south of the merge.

Suggestions for future genetic monitoring

1. Fine-scale inferences about movement patterns could be made elsewhere in the species range with new field samples that have spatial and temporal coverage comparable to this study.
2. The goals for future genetic studies must be defined before sampling, and should link to specific parameters that inform management directly. The task of identifying exactly what these parameters are, and what level of precision is needed, requires consultation with population ecologists and conservation biologists. For example, what quantitative criteria could be used to establish whether a corridor is "functional"? What parameter should be estimated in order to assess functionality, and what threshold value would be required? Genetic data will be useful for many, but not all of the parameters that could be used to answer these questions.
3. Any future monitoring with genetic markers should be tied to a primary goal of estimating one population parameter with very high precision. Examples might be:
   - estimating N_e within a very specific geographic area
   - estimating seasonal individual movement patterns within a limited number of corridors
   - complete genetic sampling of all deer in a particular area in order to understand familial relationships
   - accurately estimating the mean and variance of family size in different parts of the county.

   Data from completed projects could be used in a power analysis to guide sample sizes for future studies.
4. Regardless of what future priorities may be important for southern mule deer, we recommend that a low-cost (or volunteer) labor force be employed to annually archive scat or tissue samples in a permanent genetic resource bank. The supply costs to extract, preserve and archive DNA for future studies would be relatively low.

Acknowledgements

We are grateful to the following members of the Tracking Team, who provided or assisted in the collection of the samples collected from 2005-present: Barry Martin, Uli Burgin, Carol Crafts, Anna Gateley-Stanton, Bill Grabb, Sarah Haas, Denise Harter, Juliet Hendershott, John Kiss, Fred Kramer, Karen Larsen-Gordon, Mac MacNair, Karen Merrill, Lani Noreke, Don Rowe, Frank Santana, Doug Schanzenbach, Gary Seiser, Shankar Shivappa, Phoenix Von Hendy, Bryan Webber, Bryan and Diane Wold, Jerry Woodworth. T. Pillsbury and Steven Thomas provided samples from the Wild Animal Park. Kathleen McNamara Schroeder assisted with extraction protocol development. Chuck Black (MCAS Miramar), Bill Bretz, John Martin (FWS), Tracie Nelson (DFG), and Stephen Rice guided and assisted with sample collection. Gi Hye Im, Paymaun Shakiba, Jacques Calame, Justin McKetney, Stefania Veneri, Stephen Rice and Paul Maier assisted in the laboratory.

References

Alcaide, M., Serrano, D., Negro, J.J., Tella, J.L., Laaksonen, T., Müller, C., Gal, A., Korpimäki, E. 2009. Population fragmentation leads to isolation by distance by not genetic impoverishment in the philopatric Lesser Kestrel: a comparison with the widespread and sympatric Eurasian Kestrel. Heredity 102, 190-198.

Banks, S.C., Peakall, R. 2012. Genetic spatial autocorrelation can readily detect sex-biased dispersal. Molecular Ecology 21, 2092-2105.

Beck, N.R., Peakall, R., Heinsohn, R 2008. Social constraint and an absence of sex-biased dispersal drive fine-scale genetic structure in white-winged choughs. Molecular Ecology 17, 4346-4358.

Bohonak, A.J. 1999. Dispersal, gene flow, and population structure. Quarterly Review of Biology 74, 21-45.

Broquet, T., Petit, E.J. 2004. Quantifying genotyping errors in noninvasive population genetics. Molecular Ecology 13, 3601-3608.

California Department of Parks and Recreation. 1998. Wildlife Management Plan for Torrey Pines State Reserve. 47 pages.

Conservation Biology Institute (CBI). 2002. Wildlife corridor monitoring study for the Multiple Species Conservation Program. Prepared for City of Poway, City of San Diego, and California Department of Fish and Game. January. www.consbio.org

Conservation Biology Institute (CBI). 2003. Review of regional habitat linkage monitoring locations. Prepared for California Department of Fish and Game. January. www.consbio.org

Corander, J., Sirén, J., Arjas, E. 2008. Bayesian spatial modeling of genetic population structure. Computational Statistics 23, 111-129.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillet, B.J., Ovenden, J.R. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Molecular Ecology Resources 14, 209-214.

Doerner, K.C., Braden, W., Cork, J., Cunningham, T., Rice, A., Furman, B.J., McElroy, D. 2005. Population genetics of resurgence: white-tailed deer in Kentucky. Journal of Wildlife Management 69, 345-355.

Earl, D.A., vonHoldt, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4, 359-361.

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.

Falush, D., Stephens, M., and Pritchard, J. K. 2003. Inference of population structure: extensions to linked loci and correlated allele frequencies. Genetics 164, 1567-1587.

Frankham, R. 1995. Effective population size/adult population size ratios in wildlife: a review. Genetical Research 66, 95-107.

Gilson, A., Syvanen, M., Levine, K., Banks, J. 1998. Deer gender determination by polymerase chain reaction: validation study and application to tissues, blood stains and hair forensic samples from California. California Fish and Game 84, 159-169.

Goudet, J. 2001. FSTAT (v.1.2): a computer program to calculate F-statistics. Journal of Heredity 86, 485-486.

Hayden, S.K. 2002. The effects of habitat fragmentation on large mammals in a San Diego county regional corridor. 2002. Master's thesis, San Diego State University, San Diego, California.

Hill, W.G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genetical Research 38, 209-216.

Hubisz, M., Falush, D., Stephens, M., and Pritchard, J. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9, 1322-1332.

Jones, O.R., Wang, J. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10, 551-555.

Kalinowski, S.T., Taper, M.L., Marshall, T.C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16, 1099-1006.

Leberg, P.L., Ellsworth, D.L. 1999. Further evaluation of the genetic consequences of translocations on southeastern white-tailed deer populations. The Journal of Wildlife Management 63, 327-334.

Markovchick-Nicholls, L., Regan, H.M., Deutschman, D.H., Widyanata, A., Martin, B., Noreke, L., Hunt, T.A. 2008. Relationships between human disturbance and wildlife land use in urban habitat fragments. Conservation Biology 22, 99-109.

McClure, M.F., Bissonette, J.A., Conover, M.R. 2005. Migratory strategies, fawn recruitment, and winter habitat use by urban and rural mule deer (*Odocoileus hemionus*). European Journal of Wildlife Research 51, 170-177.

Mech, S.G., Hallett, J.G. 2001. Evaluating the effectiveness of corridors: a genetic approach. Conservation Biology 15, 467–474.

Miller, C.R., Joyce, P., Waits, L.P. 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160, 357-366.

Miller, M.P. 1997. Tools for population genetic analyses (TFPGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author.

Miquel, C., Bellemain, E., Poillot, C., Bessiere, J., Durand, A., Taberlet, P. 2006. Quality indexes to assess the reliability of genotypes in studies using non-invasive sampling and multiple-tube approach. Molecular Ecology Notes 6, 985-988.

Mitelberg, A. 2010. Social structure and genetic connectivity in the San Diego southern mule deer. Master's thesis, San Diego State University. San Diego, California.

Moritz, C. 1994. Defining 'evolutionary significant units' for conservation. Trends in Ecology and Evolution 9, 373-375.

Neel, M.C., McKelvey, K., Ryman, N., Lloyd, M.W., Short Bull, R., Allendorf, F.W., Schwartz, M.K., Waples, R.S. 2013. Estimation of effective population size in continuously distributed populations: there goes the neighborhood. Heredity 111, 189-199.

Palsbøll, P. J., Bérubé, M., Allendorf, F.W. 2007. Identification of management units using population genetic data. Trends in Ecology and Evolution 22, 11-16.

Peakall, R., Smouse, P.E. 2006. GenAlEx 6: genetic analysis in Excel; population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295.

Peakall, R., Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel; population genetic software for teaching and research-an update. Bioinformatics 28, 2537-2539.

Pearse, D.E., Crandall, K.A. 2004. Beyond F_{ST}: analysis of population genetic data for conservation. Conservation Genetics 5, 585-602.

Pease, K.M., Freedman, A.H., Pollinger, J.P., McCormack, J.E., Buermann, W., Rodzen, J., Banks, J., Meredith, E., Bleich, V., Schaefer, R.J., Jones, K., Wayne, R.K. 2009. Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation. Molecular Ecology 18, 1848-1862.

Peles, J.D., Weathersbee, F.W., Johns, P.E., Griess, J., Baker, D.L., Smith, M.H. 1999. Genetic variation in a recently isolated population of mule deer (*Odocoileus hemionus*). The Southwestern Naturalist 44, 236-240.

Porter, W.F., Mathews, N.E., Underwood, H.B., Sage, R.W., Behrend, D.F. 1991. Social organization in deer – implications for localized management. Envrionmenal Management 15, 809-814.

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Puttker, T., Meyer-Lucht, Y., Sommer, S. 2008. Fragmentation effects on population density of three rodent species in secondary Atlantic Rainforest, Brazil. Studies on Neotropical Fauna and Environment 43, 11-18.

Raymond, M., Rousset, F. 1995. An exact test for population differentiation. Evolution 49, 1283-1286.

Rodriguez-Ramilo, S.T., Wang, J. 2012. The effect of close relatives on unsupervised Bayesian clustering algorithms in population genetic structure analysis. Molecular Ecology Resources 12, 873-884.

Ryder, O.A. 1986. Species conservation and systematics: The dilemma of subspecies. Trends in Ecology and Evolution 1, 9-10.

Sommer, M.L., Barboza, R.L., Botta, R.A., Kleinfelter, E.B., Schauss, M.E., Thompson, J.R. 2007. Habitat Guidelines for Mule Deer: California Woodland Chaparral Ecoregion. Mule Deer Working Group, Western Association of Fish and Wildlife Agencies.

U.S. Fish and Wildlife Service, National Marine Fisheries Service. 1996. Interagency policy regarding the recognition of distinct vertebrate population segments under the ESA. Federal Register, pp. 4722-4725. February 7, 1996.

Valero, S. J. 2004. Estimates of genetic variation, gene flow and dispersal of the southern mule deer, *Odocoileus hemionus hemionus*, using a non-invasive method. Master's thesis, San Diego State University.

Valière, N. 2002. GIMLET: a computer program for analyzing genetic individual identification. Molecular Ecology Notes 2, 377-379.

van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P. 2004. Microchecker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4, 535-538.

Waits, L.P., Gordon, L., Taberlet, P. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Molecular Ecology 10, 249-256.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166, 1963-1979.

Wang, J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. Molecular Ecology 18, 2148-2164.

Wang, J., Santure, A.W. 2009. Parentage and sibship inference from multi-locus genotype data under polygamy. Genetics 181, 1579-1594.

Waples, R.S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics 7, 167-184.

Waples, R.S., Do, C. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8, 753-756.

Waples, R.S., Do, C. 2010. Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evolutionary Applications 3, 244-262.

Waples, R.S., 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.

Whitlock, M.C., McCauley, D.E. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. Heredity 82, 117-25.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16, 97-159.

Table 1. Population (as determined from genetic analysis), and the sampling sites that comprise each population (Figure 2). N refers to the number of unique individuals (not scat samples collected).

Population (total N)	Sampling site	N
Western Populations		
TP-SV (23)	TPSR (Torrey Pines State Reserve)	13
	SV (Sorrento Valley)	10
PQ-PC (42)	PQ (Penasquitos Canyon)	33
	PC (Penasquitos Creek)	9
CC-Mir (11)	CC (Carrol Canyon)	8
	Mir-EM (Miramar, Eastgate Mall)	3
MirNW (14)	MirA (Miramar, Rose Canyon)	10
	Mir-GC (Miramar Rd, across Golf Course)	4
MT-Mir (26)	MirD (Miramar, "Parcel G")	1
	MirE (Miramar, East)	3
	MirG (Miramar, Landfill)	8
	MT (Mission Trails Regional Park)	10
	TS (Tierrasanta)	4
Eastern Populations		
BC-SC (22)	BC (Beeler Canyon)	3
	SC (Sycamore Canyon)	19
NW (8)	CP (Calavera Preserve)	1
	RLC (Rancho La Costa Habitat Conservation Area)	2
	WAP (Wild Animal Park)	5
NE (11)	CRSP (Cuyamaca Reserve)	2
	CNFn (Cleveland National Forest, north)	1
	JL (Julian)	5
	LJ (Lake Jennings)	3
SE (16)	CNFs (Cleveland National Forest, south)	1
	HC (Hollenbeck Canyon)	5
	RJER (Rancho Jamul Ecological Reserve)	2
	SM (San Miguel Mountain Open Space)	1
	SoC (South Crest)	4
	SP (Sycuan Peak)	3
	Total	**173**

Table 2. Microsatellite diversity statistics, and single PCR error rates. For genetic diversity based on 173 unique individuals: N_s = number of samples genotyped per locus, H_O = observed heterozygosity, H_e = expected heterozygosity, A_{Tot} = number of alleles at each locus, PID = probability of identity, and PID_{sib} = probability of identity for full siblings. The single PCR error rates are based on analysis of 240 consensus genotypes: %ADO = percent allele dropout, and %FA = percent false alleles. Error rates do not represent uncertainty in the final data set (where consensus genotypes were constructed from replicate PCRs).

Locus	Diversity statistics						Single PCR error rates	
	N_s	H_O	H_e	A_{Tot}	PID	PID_{sib}	%ADO	%FA
B	160	0.275	0.501	2	0.375	0.594	4.17	0.16
C	171	0.228	0.238	4	0.596	0.78	8.82	0.17
D	173	0.711	0.694	6	0.148	0.441	0.98	0.00
F	172	0.262	0.275	3	0.548	0.75	0.57	0.15
G	173	0.555	0.588	3	0.253	0.52	4.42	0.67
H	171	0.368	0.413	2	0.431	0.652	7.30	0.78
J	173	0.387	0.48	3	0.384	0.607	1.85	0.68
K	172	0.674	0.703	5	0.147	0.436	2.21	0.00
L	171	0.509	0.567	3	0.253	0.531	8.28	1.04
M	173	0.595	0.693	4	0.158	0.444	0.88	0.80
N	170	0.806	0.879	11	0.028	0.319	4.06	1.48
P	173	0.468	0.592	5	0.252	0.518	4.44	0.17
R	172	0.587	0.614	6	0.214	0.498	4.80	0.17
S	169	0.698	0.82	9	0.059	0.356	4.09	0.80
V	173	0.480	0.554	4	0.276	0.543	0.37	0.34
Average			0.574	4.67			3.82	0.49
Combined					1.07×10^{-10}	5.15×10^{-5}		

Table 3. Collection and movement information on recaptured individuals (mapped in Figure 4). All were recaptured at the same site. A second number, when present, represents the time and distance between the second and third dates.

Individual (gender)	Population (site)	First Date	Second Date	Third Date	Days Apart	Distance (m)
1 (f)	TP-SV (TPSR)	4/19/06	7/22/06	3/25/07	93, 243	384, 1086
2 (f)	TP-SV (TPSR)	7/16/06	4/8/07		262	203
3 (f)	TP-SV (SV)	6/18/06	2/24/13		2406	78
4 (f)	MirNW (MirA)	5/29/12	6/20/12		21	935
5 (m)	PQ-PC (PQ)	7/10/06	3/25/07		255	269
6 (m)	PQ-PC (PQ)	6/11/06	7/9/06		28	68
7 (f)	PQ-PC (PC)	5/8/06	4/8/07		330	195
8 (f)	PQ-PC (PC)	5/8/06	4/8/07		330	101
9 (f)	NW (WAP)	12/20/12	6/3/13		163	n/a
10 (f)	BC-SC (SC)	7/17/06	9/26/06	12/6/06	69, 70	184, 20
11 (f)	BC-SC (SC)	6/6/06	12/6/06		180	1010
12 (f)	BC-SC (SC)	8/19/13	8/21/13		2	318
13 (f)	MT-Mir (TS)	10/9/12	10/15/12		6	453
14 (m)	MT-Mir (MT)	9/3/12	10/8/12		35	79
15 (f)	MT-Mir (MT)	9/3/12	10/8/12		35	95

Table 4. Summary of family sizes, using full siblings inferred by genetic similarity.

Sires	
Number of sires inferred	40
Number of inferred sires that were sampled	3
Mean # offspring / sire (range)	4.3 (1 - 11)
Mean dams / sire	3.6 (1 - 8)
Dams	
Number of dams inferred	42
Number of inferred dams that were sampled	7
Mean # offspring / dams (range)	4.1 (1 - 8)
Mean sires / dam	3.5 (1 - 6)

Table 5. Location of the 3 sires and 7 dams inferred from genetic relatedness that were actually sampled, and their 20 associated offspring. Eight offspring were sampled from a different site than their parent (see map in Figure 5). Single-underlined offspring were sampled from a different sampling site (but same population) as the parent, double-underlined offspring were sampled from a different population, and bold offspring were sampled from a different gene pool.

Parent			Offspring	
Gender	**Sample number**	**Population (site)**	**Gender**	**Population (site)**
Male	D359	TP-SV (SV)	female (3)	TP-SV (SV) (3)
Male	D545	MirNW (MirA)	**male (1)**, female (1)	**SE (SP),** MirNW (MirA)
Male	D577	NW (WAP)	female (1)	NW (WAP)
Female	D177	TP-SV (TPSR)	female (1)	TP-SV (TPSR)
Female	D507	CC-Mir (Mir-EM)	female (1), female (1), male (1)	CC-Mir (Mir-EM), MT-Mir (MirG), MirNW (MirA)
Female	D518	MirNW (Mir-GC)	unknown (1**), female (1)**, female (1)	MirNW (Mir-GC), **NE (JL),** MirNW (MirA)
Female	D523	MirNW (Mir-GC)	male (1), male (1)	MirNW (Mir-GC), MT-Mir (MT)
Female	D528	MirNW (MirA)	female (1)	MirNW (MirA)
Female	D531	MirNW (MirA)	female (1), female (1), male (1)	MirNW (MirA), MirNW (Mir-GC), PQ-PC (PQ)
Female	D680	NE (JL)	female (1)	NE (JL)

Table 6. Locations of 36 full sibling (FS) dyad pairs (p > 0.97). Only the sites with dyads are shown. Ten dyad pairs from different sites are underlined.

		TP-SV		PQ-PC		CC-Mir		MirNW	MT-Mir					BC-SC		NW	SE
		SV	TPSR	PC	PQ	CC	Mir-EM	MirA	MirD	MirE	MirG	MT	TS	BC	SC	WAP	HC
TP-SV	SV	1															
	TPSR		5														
PQ-PC	PC			2													
	PQ				6												
CC-Mir	CC					1											
	MirEM	4					1										
MirNW	MirA						1	3									
MT-Mir	MirD				1												
	MirE					1											
	MirG					2					1						
	MT											1					
	TS												1				
BC-SC	BC													1			
	SC														2		
NW	WAP															1	
SE	HC				1												

Table 7. Microsatellite diversity statistics, based on the set of 133 unrelated individuals. A = number of alleles, A_R = allelic richness rarefied to six individuals, H_O = observed heterozygosity, H_e = expected heterozygosity, F_{IS} = inbreeding coefficient. Asterisk indicates $p < 0.05$. None of the 126 tests (9 sets of tests per locus) were significant at the Bonferroni-corrected level of $p < 0.0004$.

		Locus														
Western Populations		**C**	**D**	**F**	**G**	**H**	**J**	**K**	**L**	**M**	**N**	**P**	**R**	**S**	**V**	**Average**
TP-SV	A	2	3	3	3	2	2	4	3	3	7	2	3	4	2	3.1
n=17	A_R	1.936	2.871	2.194	2.637	1.997	1.999	3.340	2.769	2.895	5.121	2.000	2.762	3.604	1.999	2.7
	H_e	0.327	0.569	0.337	0.504	0.472	0.484	0.638	0.597	0.590	0.791	0.498	0.569	0.704	0.493	0.541
	H_O	0.412	0.765	0.294	0.588	0.412	0.471	0.824	0.412	0.706	0.875	0.471	0.588	0.765	0.529	0.579
	F_{IS}	-0.231	-0.316	0.158	-0.139	0.158	0.059	-0.262	0.337	-0.167	-0.074	0.086	-0.003	-0.056	-0.043	-0.040
PQ-PC	A	2	5	2	3	2	2	4	3	4	9	3	4	7	3	3.8
n=32	A_R	1.502	3.456	1.765	2.731	1.405	1.946	3.669	2.922	3.116	5.822	2.397	2.867	5.380	2.401	3.0
	H_e	0.117	0.669	0.219	0.543	0.089	0.359	0.719	0.632	0.675	0.853	0.516	0.539	0.833	0.534	0.521
	H_O	0.125	0.813	0.188	0.563	0.094	0.406	0.688	0.594	0.594	0.800	0.375	0.594	0.774	0.563	0.512
	F_{IS}	-0.051	-0.199	0.158	-0.021	-0.033*	-0.116	0.060	0.077	0.136	0.079	0.288	-0.087	0.086	-0.038	0.034
CC-Mir	A	2	3	2	2	2	2	4	3	3	5	3	2	3	2	2.7
n=6	A_R	1.833	2.985	1.833	1.833	2.000	1.985	3.667	3.000	2.818	4.803	2.833	2.000	2.818	2.000	2.6
	H_e	0.153	0.611	0.153	0.153	0.486	0.278	0.639	0.625	0.403	0.764	0.486	0.444	0.403	0.375	0.427
	H_O	0.167	0.667	0.167	0.167	0.833	0.333	0.500	0.500	0.500	1.000	0.500	0.333	0.167	0.167	0.429
	F_{IS}	0.000	0.000	0.000	0.000	-0.667	-0.111	0.302	0.286	-0.154	-0.224	0.063	0.333	0.643	0.615	0.086
MirNW	A	3	4	1	3	2	2	3	3	3	6	2	3	6	2	3.1
n=8	A_R	2.250	3.581	1.000	2.963	1.992	1.875	2.963	2.956	2.839	5.323	2.000	2.990	5.076	2.000	2.8
	H_e	0.227	0.648	0.000	0.617	0.375	0.219	0.617	0.586	0.477	0.805	0.469	0.648	0.755	0.469	0.494
	H_O	0.250	0.750	0.000	0.625	0.500	0.250	0.625	0.875	0.625	0.625	0.500	0.500	0.857	0.500	0.534
	F_{IS}	-0.037	-0.091	NA	0.054	-0.273	-0.077	0.054	-0.441	-0.250	0.286	0.000	0.291	-0.059	0.000	-0.015
MT-Mir	A	3	4	3	3	2	2	3	3	4	8	2	5	7	3	3.7
n=21	A_R	1.662	3.185	2.136	2.418	1.999	1.999	2.967	2.505	3.128	5.369	1.986	3.603	4.849	2.562	2.9
	H_e	0.135	0.670	0.255	0.516	0.496	0.490	0.662	0.439	0.632	0.824*	0.427	0.655	0.800*	0.541	0.539
	H_O	0.143	0.524	0.286	0.524	0.619	0.381	0.714	0.429	0.571	0.667*	0.524	0.714	0.650*	0.524	0.519
	F_{IS}	-0.034	0.241	-0.096	0.009	-0.226	0.245	-0.054	0.048	0.119	0.215*	-0.202	-0.066	0.212	0.056	0.061

Table 7. (continued).

		Locus														
Eastern Populations		**C**	**D**	**F**	**G**	**H**	**J**	**K**	**L**	**M**	**N**	**P**	**R**	**S**	**V**	**Average**
BC-SC	A	3	5	3	3	2	2	4	3	4	8	4	3	5	3	3.7
n=18	A_R	1.762	3.912	2.667	2.635	1.948	2.000	3.496	2.597	3.052	5.721	2.893	2.861	3.867	2.451	3.0
	H_e	0.156	0.685	0.468	0.573	0.346	0.499	0.665	0.452	0.557	0.841	0.542	0.586	0.654	0.452*	0.534
	H_O	0.111	0.667	0.556	0.500	0.333	0.611	0.611	0.471	0.667	0.944	0.444	0.500	0.611	0.389*	0.530
	F_{IS}	0.313	0.056	-0.160	0.155	0.064	-0.199	0.110	-0.012	-0.169	-0.095	0.207	0.175	0.094	0.168	0.037
NW	A	2	3	3	3	2	2	4	3	4	7	4	4	6	2	3.5
n=6	A_R	1.833	3.000	2.818	2.833	2.000	2.000	3.667	3.000	3.818	6.167	3.667	3.985	5.500	2.000	3.3
	H_e	0.153	0.653	0.403	0.569	0.420	0.375	0.583	0.653	0.681	0.792	0.625	0.736	0.792	0.444	0.563
	H_O	0.167	0.833	0.167	0.333	0.200	0.167	0.500	1.000	1.000	0.833	0.667	1.000	0.667	0.667	0.586
	F_{IS}	0.000	-0.190	0.643	0.487	0.600	0.615	0.231	-0.463	-0.395	0.038	0.024	-0.277	0.245	-0.429	0.052
NE	A	3	4	2	3	2	2	5	3	4	7	4	3	5	4	3.6
n=10	A_R	2.457	3.844	1.763	2.989	1.993	1.895	4.246	2.658	3.526	5.245	3.394	2.815	4.115	3.773	3.2
	H_e	0.395	0.725	0.180	0.660*	0.401	0.255	0.710	0.405	0.660	0.785	0.655	0.568	0.675	0.690	0.555
	H_O	0.500	0.900	0.200	0.600*	0.556	0.100	0.800	0.400	0.500	0.700	0.700	0.667	0.600	0.800	0.573
	F_{IS}	-0.216	-0.191	-0.059	0.143	-0.333	0.640	-0.075	0.065	0.291	0.160	-0.016	-0.116	0.163	-0.108	0.020
SE	A	3	6	3	3	2	3	4	3	4	8	3	4	7	3	4.0
n=15	A_R	2.613	3.332	2.282	2.758	1.823	2.333	3.641	2.763	3.634	5.631	2.785	2.665	4.673	2.714	3.1
	H_e	0.494	0.618	0.291	0.498	0.231	0.527	0.700	0.487*	0.687	0.824	0.527	0.553	0.762*	0.464	0.547
	H_O	0.308	0.667	0.333	0.467	0.267	0.667	0.533	0.214*	0.533	0.733	0.400	0.667	0.600*	0.533	0.494
	F_{IS}	0.411	-0.045	-0.111	0.097	-0.120	-0.233	0.270	0.585*	0.256	0.144	0.273	-0.172	0.246	-0.114	0.131*
All individuals	A	3	6	3	3	2	3	5	3	4	11	5	6	9	4	4.8
n=133	A_R	2.006	3.620	2.092	2.718	1.966	2.032	3.516	2.806	3.330	6.070	2.813	3.098	5.012	2.828	3.1
	H_e	0.240	0.698	0.282	0.585	0.405	0.485	0.706	0.573	0.692	0.871	0.590	0.623	0.820	0.557	0.581

Table 8. Divergence among all pairs of populations, quantified as F_{ST}. All tests except NW vs. SE are significant at $p < 0.05$; values in bold are significant after Bonferroni correction for 36 tests ($p < 0.00139$).

	TP-SV	PQ-PC	CC-Mir	MirNW	MT-Mir	BC-SC	NW	NE
PQ-PC	**0.065**							
CC-Mir	**0.083**	**0.075**						
MirNW	**0.075**	**0.046**	0.096					
MT-Mir	**0.033**	**0.034**	**0.078**	**0.070**				
BC-SC	**0.081**	**0.035**	**0.094**	**0.090**	**0.032**			
NW	**0.089**	**0.070**	0.118	0.087	0.053	0.051		
NE	**0.122**	**0.109**	**0.165**	**0.109**	**0.084**	**0.054**	0.056	
SE	**0.115**	**0.080**	**0.116**	**0.056**	**0.084**	**0.050**	0.037	0.046

Figures

Figure 1. Map of the study area, showing individual sample locations. Dark gray polygons represent all protected lands (city, county, state and federal) in CPAD (California Protected Areas Database, acc. April 2, 2014). Open space on military lands (such as MCAS Miramar immediately north of state route 52) is not included in this classification.

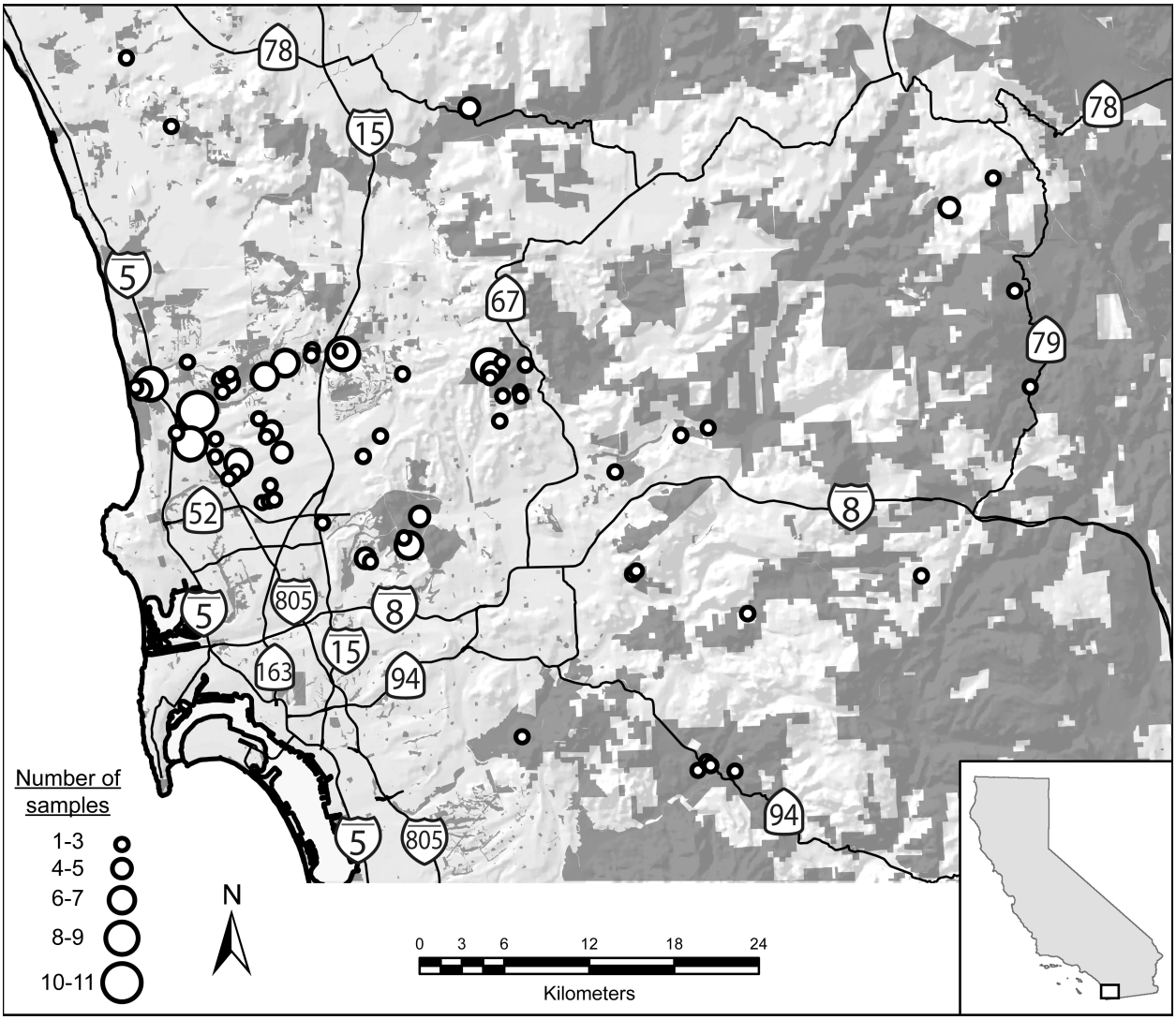


Figure 2. Map of the local populations and regional gene pools inferred from the genetic analyses, with CPAD protected lands shaded.

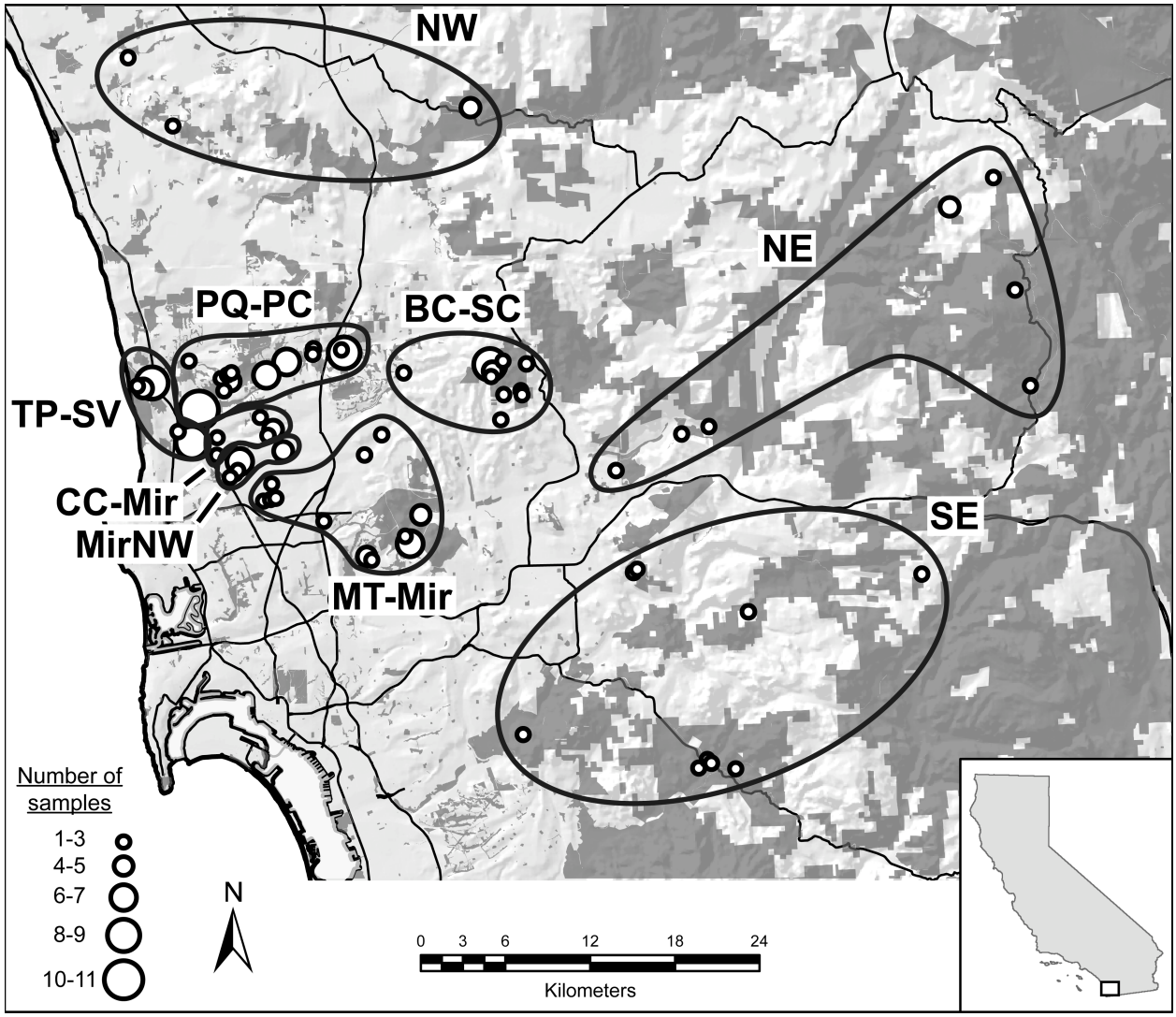


Figure 3. Mean Q-scores (quality scores) in the environmental exposure experiment. Each point represents the average of all microsatellite loci for four PCRs, from a five pellet sample, for a particular pile-treatment-day combination. The Standard Protocol treatment included one additional level of replication, with two samples taken per day, per scat pile, by different technicians. Error bars represent the minimum and maximum of the replicates.

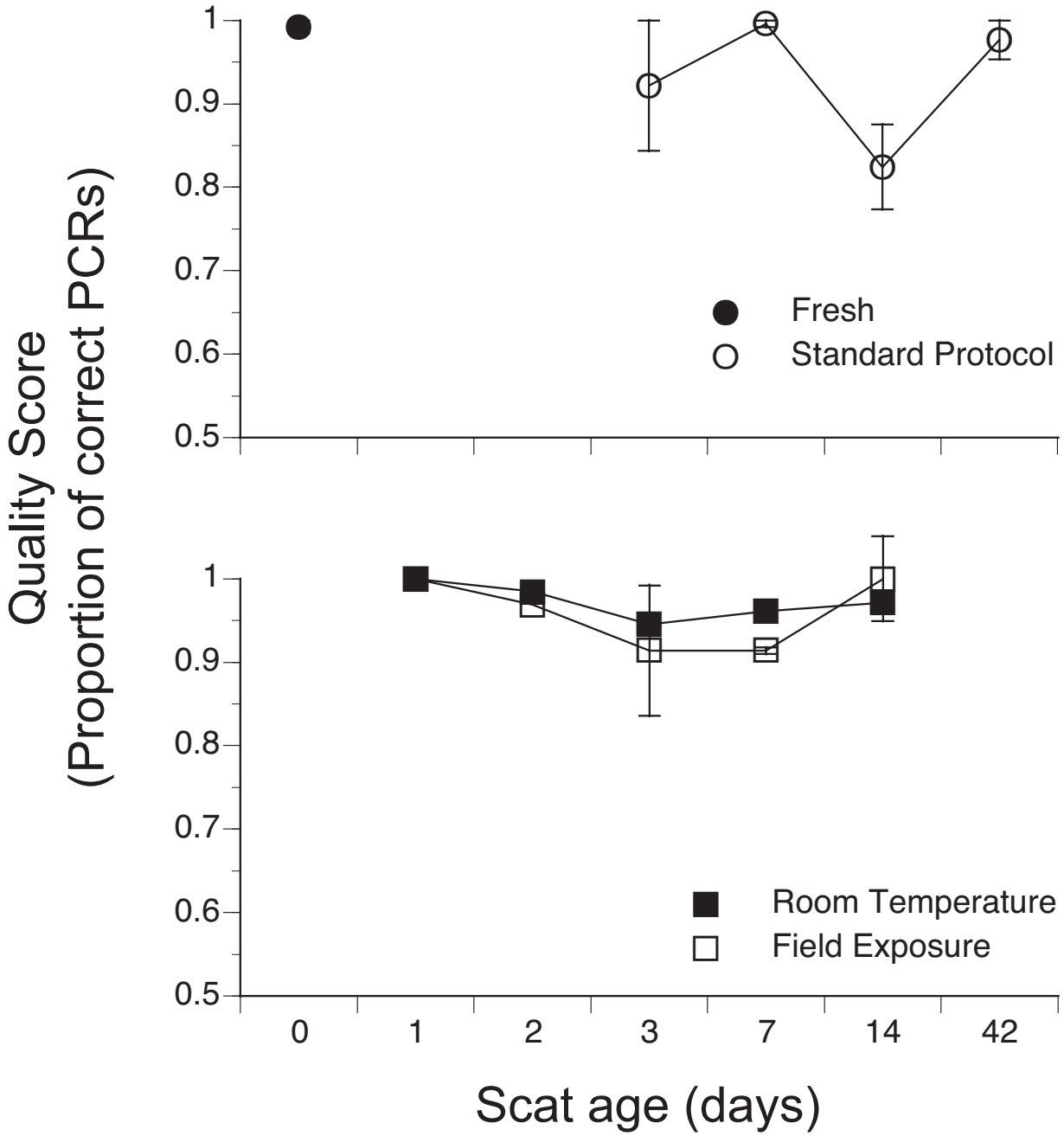


Figure 4. Locations of "recaptured" individuals (see Table 3). The maximum distance between captures and recaptures is 1.1 km; all captures and recaptures fall within the area occupied by the map symbols.

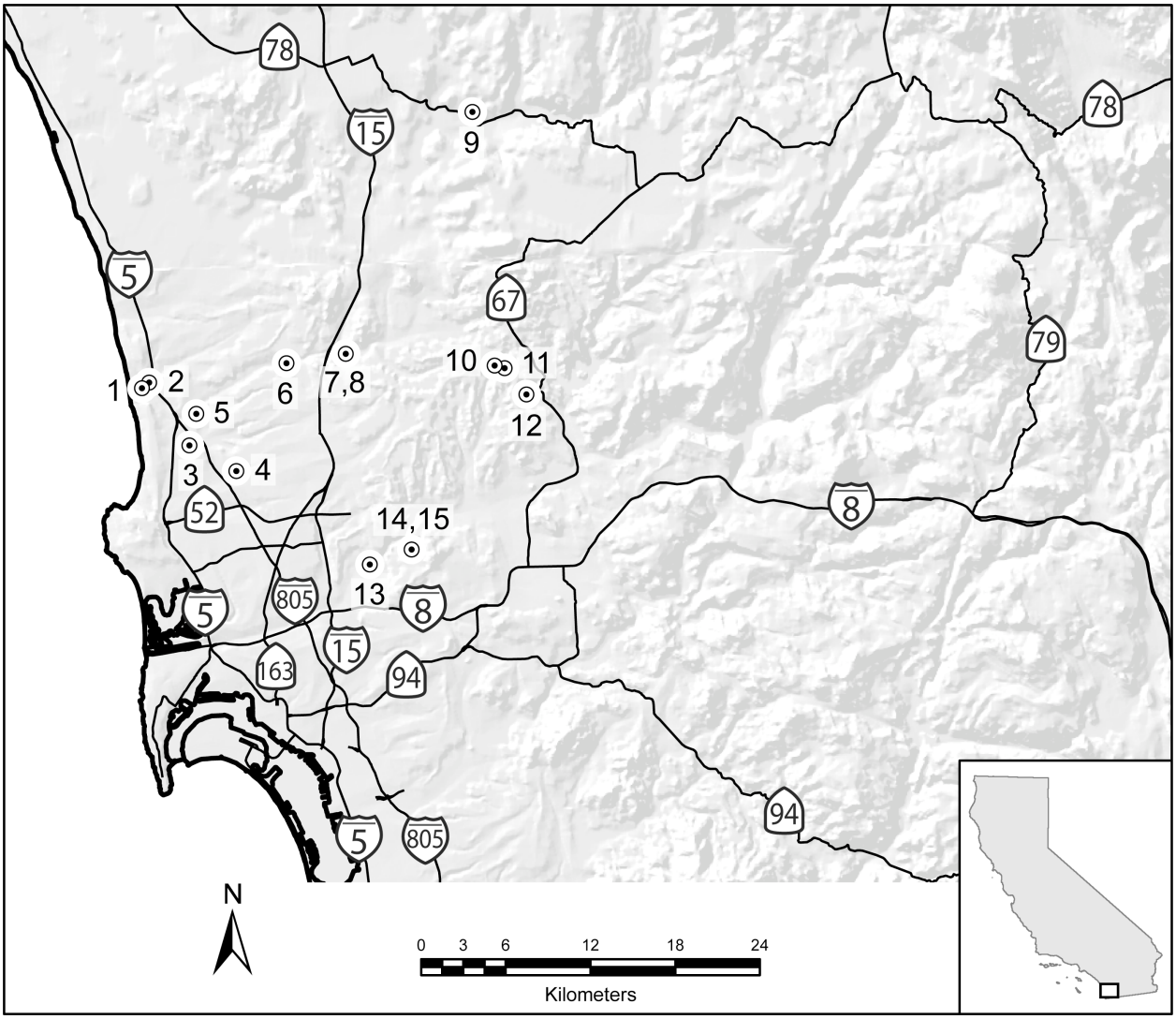


Figure 5. Locations of the eight offspring that were sampled from different sites, populations or gene pools than their parent (see Table 5). The parent is circled, and the offspring is at the destination of each arrow.

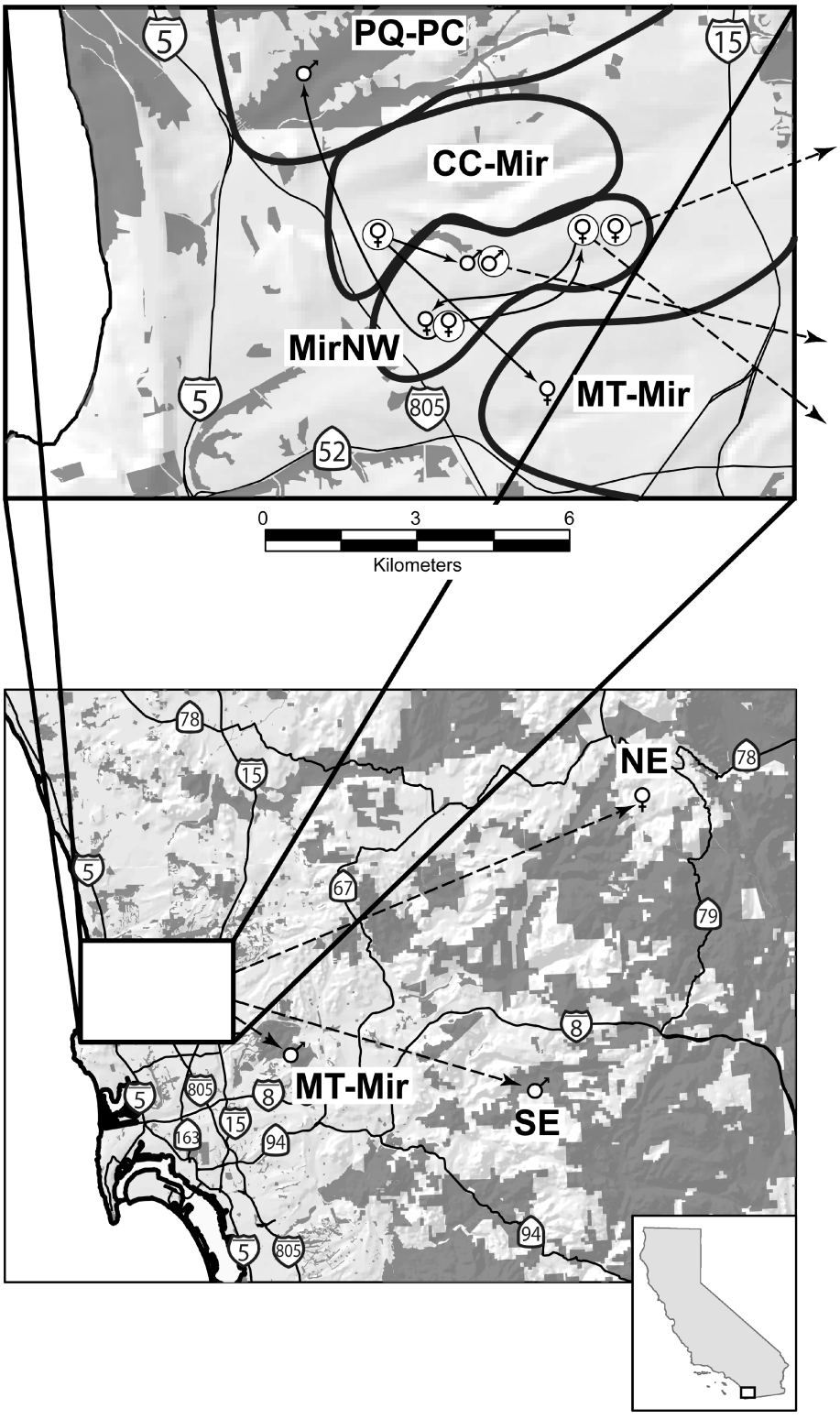


Figure 6. Distribution of geographic distances between all pairs (dyads) of parent and offspring (top), and between full siblings (bottom). Note the logarithmic scale. For the parent-offspring histogram, dyads that contain a father are in black, while dyads with a mother are open.

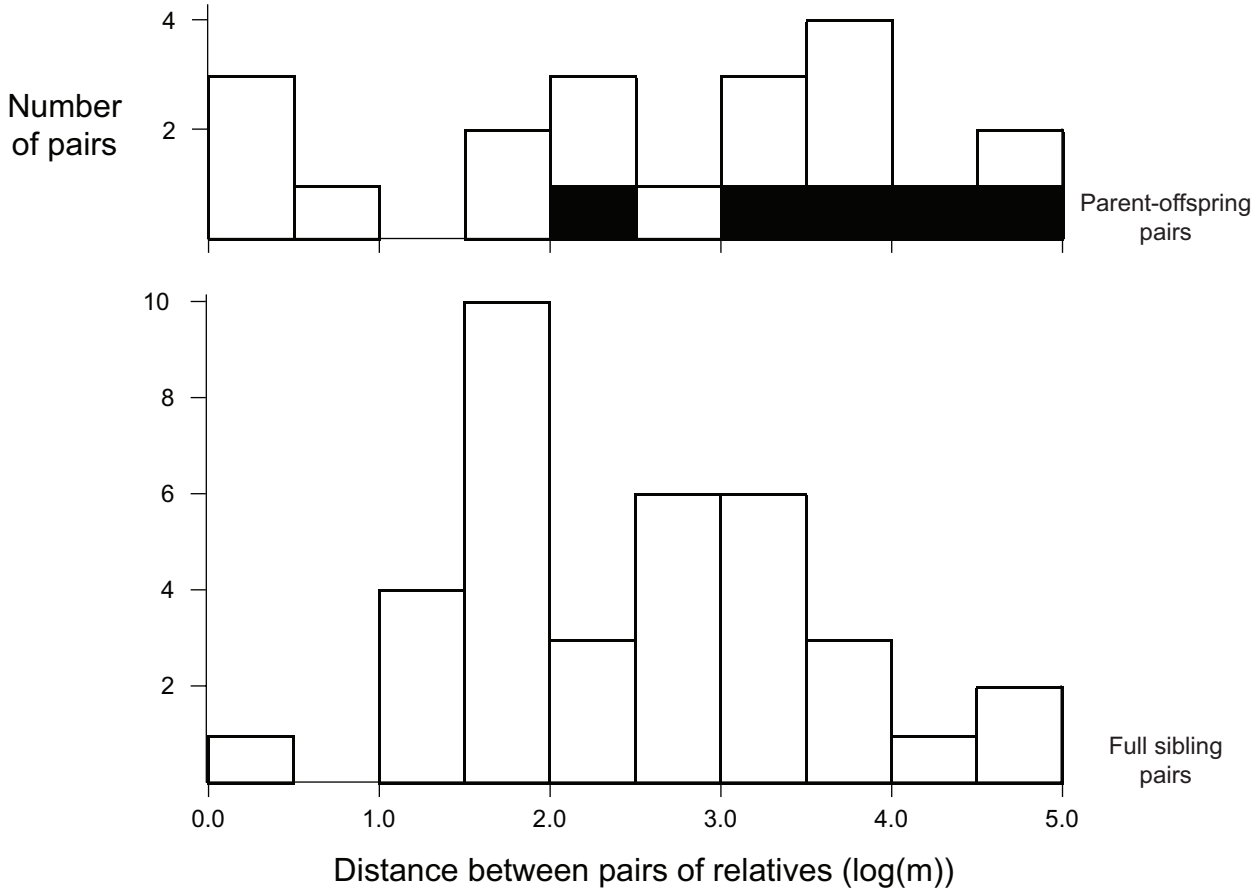


Figure 7. Results of the STRUCTURE individual clustering analysis. Each bar represents one individual in the study, with black and grey representing the relative proportions of an individual's genetic background that can be attributed to each of two regional gene pools.

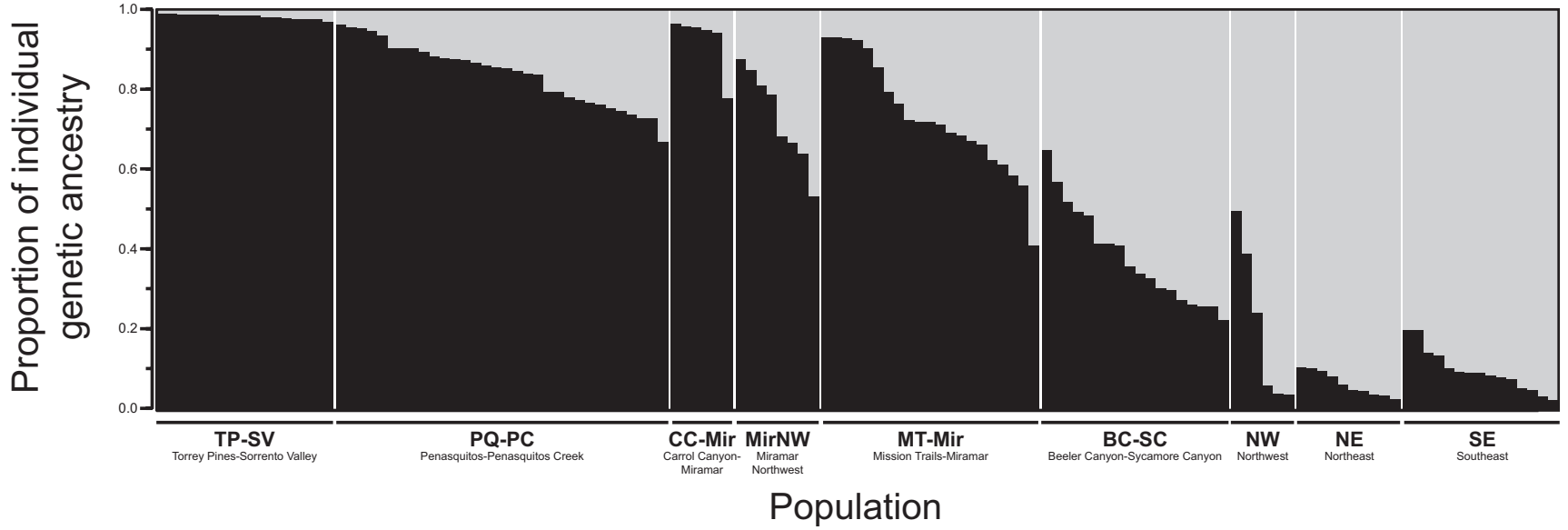
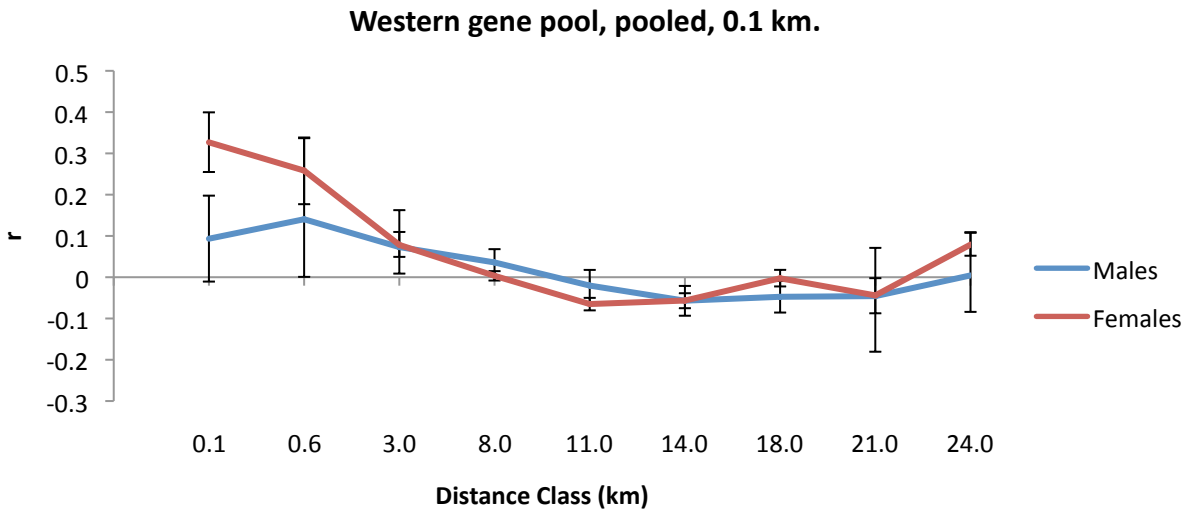
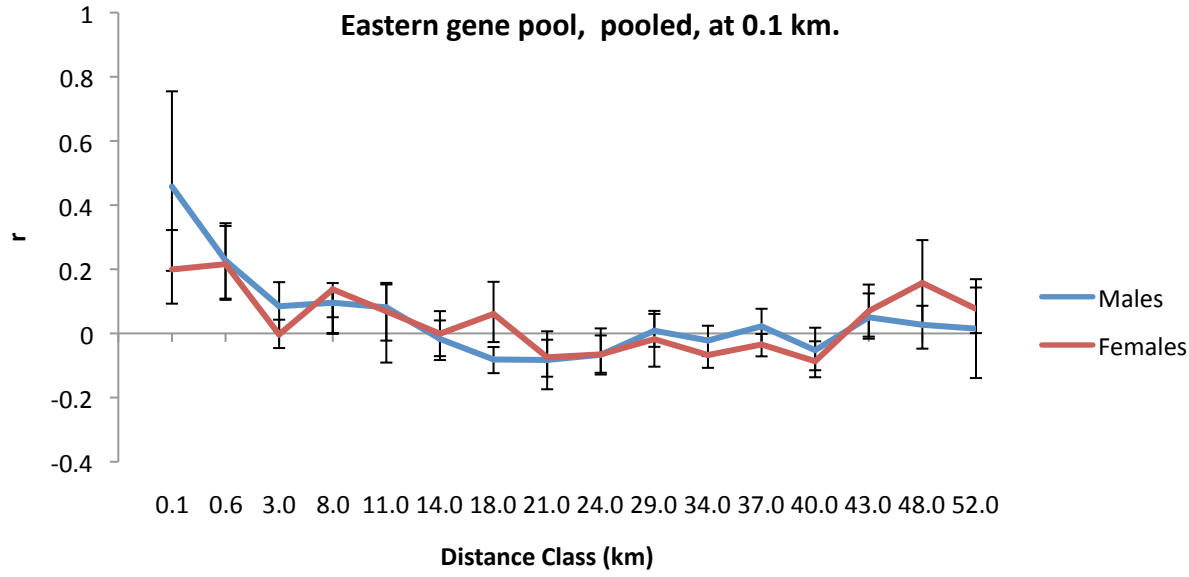


Figure 8. Individual spatial autocorrelations, plotted as a function of increasing distance classes.





Appendices

Appendix A. Mule Deer Scat Collection Protocol (revised May 2012)

Collection Protocol:

1. Using a glove, collect 20-30 pellets per pile into the provided plastic container. Seal the container either with a rubber band or with tape to avoid having the box open while in transit.
2. On the box, record GPS coordinates (in degrees), collection date, individual/team collecting the sample, as well as any notes you think may contribute (e.g. sample description at time of collection).
3. At home, open containers and leave in a shaded area, covered with some paper towels.
4. Call or email to arrange a pickup, or if you need more boxes or gloves.

To Collect or Not To Collect or Not to Collect:

1. Major factors contributing to the quality of DNA extracted for genetic analysis are: 1) time between deposition and collection, 2) exposure time to direct and indirect sunlight, 3) seasonal and site-specific environmental conditions, 4) season and site-specific diet variability. With all these factors in mind, the decision to collect a particular scat pile should be based on the following general criteria. Priority for collection should be given to samples in the following order: 1) scat piles with pellets that have limited cracks on the surface and have a sheen on the surface, 2) scat piles with pellets that cave in without crumbling when pinched between the forefinger and thumb, regardless of whether the pellet surface is cracked or dull.
2. In general, very fresh piles will have an almost green sheen to them. Very old piles will have no sheen and will crumble apart when squashed. Everything in between is fair game. In my experience, the more sheen, the more likely we'll get DNA from it. However, if all you are finding are dull piles, pick a few that look like they may be fresh and we can try them out.
3. To avoid repeatedly sampling the same group of individuals, collect no more than 20 scat piles from a site on a single day.
4. DNA quality will degrade the longer the pellets are exposed to sunlight. Try to pick the pellets on the bottom of the pile first, because they have been protected from sunlight.
5. Try to avoid clumped pellets, but if that's all that's available, go ahead and collect it.
6. Try not to pick from piles that are questionable with respect to whether one or several deer have deposited it (unless there is a clear difference in age, with the pile on the bottom being clearly "ancient" in comparison with the one on the top.)

(1) No sheen, but not cracking
(2) No sheen, with some cracking.
(3) Shiny, with some cracking.
(4) Shiny with some cracking.

I would take 3 and 4 over 1 and 2, but if 1 and 2 is all you can get, then go ahead and collect.

The sample above still had a nice sheen, though it was very dry. It was also hidden from the sun by the grass, so we collected it.

This sample was very fresh when we collected it. It was probably deposited that same morning. Notice that there are some clumps, but this sample was so fresh that I'd collect it even if it were all clumps.

Appendix B. Scat extraction protocol using DNA IQ kit (Promega, #DC6700).

1. Thaw samples, vortexing to break up pellets of sediment as best as you can. If they don't break up completely, they will once you've had them on the heat block for a couple of minutes.
2. Centrifuge tubes with PBS wash + cells for 2 min @ 14rpm.
3. Gently, remove PBS wash from the top, until you reach 200 µl.
4. Add 400 µl of Lysis Buffer + DTT (100:1 Lysis to 1M DTT) to each tube and mix by flipping upside down a couple of times.
5. Incubate @ 95C for 30 min. Make sure lids are well sealed, because they will sometimes pop off when heated. Keep an eye on the temperature, because the blocks are finicky and the temp will bounce around. If the pellets didn't break up completely in step 1, vortex for a couple of seconds to break up the pellet after a few minutes at 95C.
6. Take tubes off the block and allow to cool for a few minutes (until cool to touch).
7. Transfer approximately ½ (~300-400 µl) of the sample into a basket with filter sitting in a clean 1.7ml tube. DO NOT OVERFILL DUE TO RISK OF CONTAMINATION IN STEP 8.
8. Centrifuge for 30 seconds at 14rpm.
9. Repeat steps 7-8.
10. Throw out basket with debris.
11. Vortex bottle of beads for 20 seconds and aliquot 7 up beads into samples.
12. Vortex 3 sec.
13. Incubate for 5 min @RT, vortexing for 3 sec every minute.
14. Vortex 3 sec, set into the magnetic bead stand, and remove the liquid.
15. Add 100 µl of Lysis Buffer + DTT to each sample, vortex 3 sec, place in stand and remove liquid.
16. Add 500 µl of the Wash Buffer for each sample, vortex 3 sec, place in stand and remove liquid.
17. Add 100 µl of the Wash Buffer, vortex 3 sec, place in stand and remove liquid.
18. Repeat step 17.
19. Make sure to remove as much wash as possible in the last step. Any ethanol that remains will inhibit PCR reactions. Use a pipette tip to push down all liquid remaining and pull it out (a 200ul pipette will work best).
20. Allow beads + DNA to dry for 5 min, maximum of 20 min.
21. Add **75 µl of Elution Buffer**.
22. Vortex 3 sec, incubate at 65C for 5 min. (Keep an eye on the temperature)
23. Remove tubes from heat 2 at a time, vortex for 3 sec, place in stand and remove the 75 up of DNA into new tubes.
24. Repeat steps 21-23 into a new set of tubes, marked with *.

Appendix C. Deer primers, primers sequences, and concentration in each PCR reaction.

		Forward primer		Reverse primer	
Locus	**Size range**	**Sequence**	**Conc. (µM)**	**Sequence**	**Conc. (µM)**
B	149-153	6FAM - GCT GCT CTC CTA CTG CTC TG	0.038	CTA TTC GTC TTC TTC CTC TCT G	0.038
C	311-327	6FAM - CAA CCA TTC ATC CAT CTT G	0.203	AAA GGT AGA AAG GGT GAG C	0.203
D	162-186	PET - AGA GCC TCG TCT TTT CAT TC	0.127	TTG CTG CTT GCT TGT CTA AT	0.127
F	157-165	VIC - AAG GAG TCT TTC AGT TTT GAG A	0.025	GGT TCT GTC TTT GCT TGT TG	0.025
G	318-330	VIC - TAT GGT CAC AGC AAC ATT GT	0.038	GTT CCC TTC CTT TTT CAG G	0.038
H	349-353	PET - GCT GCC ATT GCC AGA TA	0.241	CCC CTC CTG TGC TCT CA	0.241
J	238-246	6FAM - CAC GCA ACC ACT CAT TTA CC	0.101	TGG GTG AAA GGA TTA TGT GC	0.101
K	193-209	6FAM - GCA GGA AGG AGG AGA CAG TA	0.051	GCT GGT TCG TTA TCA TTT AGC	0.051
L	260-296	PET - CCC TGT GGT CTA GCA AA	0.177	ATA GGC ACA TGC TCA TAA G	0.177
M	142-170	NED - AGG GAA ACC TCT GTT CAG GA	0.025	ACC AAG CAA AAT GCC TTA CA	0.025
N	289-330	NED - TCC AGA GAA GCA ACC AAT AG	0.127	GTG TGC CTT AAA CAA CCT GT	0.127
P	215-235	6FAM - TTT CAC TGT TTT CTC CTT CAG A	0.152	TGC CCA ATC AGA TGT TGT AG	0.152
R	264-296	VIC - GGG GTC TTC TCA ATC CA	0.127	TCA GTT TCT GGA ACT CTA AAG T	0.127
S	191-219	VIC - GCA AAG AGA CAG AAG ACA ATA G	0.101	GAC CAG GAA ACC CAG AAT	0.101
V	84-96	6FAM - GCA AAC AGA AAT AGC CAC AG	0.025	TCA GGA TGG GTT GAA TAA ATC	0.025
SRY	223	NED - CCCATGAACGCATTC ATTGTGTGG	0.101	ATTTTAGCCTTCCGA CGAGGTCGATA	0.101