Wildlife Diseases

Using Genetics to Assess Differentiation Among Raccoons in an Area with Variable Rabies Status in Alabama

SHYLO R. JOHNSON, USDA, APHIS, Wildlife Services, National Wildlife Research Center, Fort Collins, CO, USA
ANTOINETTE J. PIAGGIO, USDA, APHIS, Wildlife Services, National Wildlife Research Center, Fort Collins, CO, USA
MELISSA A. NEUBAUM, USDA, APHIS, Wildlife Services, National Wildlife Research Center, Fort Collins, CO, USA
MIKE R. DUNBAR, USDA, APHIS, Wildlife Services, National Wildlife Research Center, Fort Collins, CO, USA

ABSTRACT The western spread of raccoon rabies in Alabama has been slow and even appears to regress eastward periodically. While the disease has been present in the state for over 30 years, areas in northwest Alabama are devoid of raccoon rabies. This variation resulting in an enzootic area of raccoon rabies primarily in southeastern Alabama may be due to landscape features that hinder the movement of raccoons (i.e., gene flow) among different locations. We used 11 raccoon-specific microsatellite markers to obtain individual genotypes to examine gene flow among areas that were rabies free, enzootic with rabies, or had only sporadic reports of the disease. Samples from 70 individuals were collected from 5 sampling localities in 3 counties. The landscape feature data were collected from geographic information system (GIS) data. We inferred gene flow by estimating \( F_{ST} \) and by using Bayesian tests to identify genetic clusters. Estimates of pairwise \( F_{ST} \) indicated genetic differentiation and restricted gene flow between some sites, and an uneven distribution of genetic clusters was observed. Of the landscape features examined (i.e., land cover, elevation, slope, roads, and hydrology), only land cover had an association with genetic differentiation, suggesting this landscape variable may affect gene flow among raccoon populations and thus the spread of raccoon variant of rabies in Alabama.

KEY WORDS Alabama, genetics, landscape, *Procyon lotor*, raccoon rabies, raccoon

Rabies is a lethal zoonosis resulting in over 55,000 human deaths annually throughout the world (Knobel et al. 2005). In the United States, less than 3 deaths per year are due to rabies infection (Childs 2002). Two main reasons for this low death rate are administration of post-exposure prophylaxis (PEP) treatments and pre-exposure vaccination of animals. Through vaccination programs, the U.S. has effectively reduced rabies cases in domestic animals to less than 7% (i.e., 482) of the reported cases (Blanton et al. 2008). The other 93% (i.e., 6,776) of the cases are in wildlife, primarily raccoons (*Procyon lotor*; 36.6%), bats (*Chiroptera* spp.; 27.2%), and skunks (*Mephitis* and *Spilogale* spp.; 20.4%).

While bat rabies is reported throughout the U.S., cases from terrestrial wildlife are often geographically limited to regions where rabies exists as species-specific variants (Childs 2002). Costs associated with prevention, control, and treatment dramatically increase when the disease spreads into a new region and an epizootic develops. For example, PEP treatments increased from pre-epizootic rates of less than 4/100,000 residents to epizootic rates of 45/100,000 residents in Massachusetts (Kreindel et al. 1998) and to 66/100,000 residents in New Jersey (Uhaa et al. 1992) after an expansion of rabies into these states. The median PEP cost was $2,376 per patient while the total health care charges related to PEP in Massachusetts during the epizootic were estimated at $2.4 to $6.4 million...
The cause of the epizootic in these locations was the raccoon rabies virus variant. The raccoon rabies variant is currently enzootic in eastern U.S. Prior to the 1970s the variant was confined to Florida and Georgia. Positive raccoon cases began occurring in Alabama in the 1970s, and the previously localized focus of the disease changed dramatically in 1977 when a raccoon-specific rabies virus variant was detected in West Virginia. The disease radiated outward from this new focus at a rate of 30–50 km/yr (Rupprecht and Smith 1994) and by 1996 had spread into eastern Ohio (United States Department of Agriculture 2003). Due to the increase in animal cases and human treatments associated with the disease radiating into these new areas, a cooperative effort led by the United States Department of Agriculture, Animal and Plant Health Inspection Service, Wildlife Service (WS) was developed to control the spread of rabies. The main focus of the WS National Rabies Management Program (NRMP) was implementation of an oral rabies vaccination (ORV) program. The ORV program distributes baits filled with rabies vaccine across the landscape. This distribution is focused on the western front of the raccoon rabies variant in the eastern U.S. to create an immune barrier. However, in a few areas where prevalence of the disease (i.e., Alabama) is low or where natural barriers (e.g., high elevations) are thought to inhibit the movement of raccoons, an ORV barrier is not maintained (United States Department of Agriculture 2003).

While raccoon rabies cases were detected over 30 years ago in Alabama, the raccoon variant is only enzootic in southeast Alabama with a few cases occurring in central part of the state. Had the spread of the raccoon rabies variant in Alabama matched the unrestricted spread in the mid-Atlantic States, it would have reached Louisiana, Arkansas and eastern Texas, which have not reported the raccoon variant (Rupprecht and Smith 1994). Landscape features in Alabama may be hindering the movement of raccoons and the spread of the raccoon rabies variant. Identifying these features may allow for more effective baiting strategies to control the spread of the disease.

Using gene flow as an inference of raccoon movement, our primary objective was to examine if gene flow occurred between raccoon populations with different rabies status: enzootic, sporadic cases, and rabies free, by testing the hypothesis that raccoon populations in our sampling area were panmictic (i.e., unrestricted gene flow). If not panmictic, our second objective was to assess the role of dominate landscape features in restricting gene flow.

**METHODS**

Tissue samples for DNA analysis were taken from the ears of raccoons captured in Alabama by WS personnel or from a concurrent telemetry study (Fisher 2007, Arjo et al. 2008). Seventy individuals were sampled from five sites in three counties. Two sites, ANE (n=9) and ASW (n=14), were in Autauga County, which only reported sporadic cases of rabies. One site was in Hale County (n=20), which had no reports of rabies and the remaining two sites, LNW (n=14) and LSE (n=13), were in Lowndes County, which consistently reports cases of rabies and was considered enzootic for the disease (Fig. 1).

Extractions of DNA were conducted with DNeasy® Tissue Kits (QIAGEN Inc., Valencia, CA) following the manufacturer’s protocol. Using the extracted DNA, we amplified 11 raccoon microsatellite loci, as developed and characterized by Cullingham et al. (2006), to obtain unique genotypes for each individual raccoon. Genotypes were visualized using an Applied Biosystems
3130 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA). Alleles were scored using ABI GeneMapper Software v4.0. Scored data were exported from GeneMapper using Excel and GMConvert (Faircloth 2006). CONVERT (Glaubitz 2004) was used to transform this file into input files for software packages used in statistical analyses. Microsatellite loci were tested for departures from Hardy-Weinberg equilibrium (HWE) with 9000 steps of a Markov Chain in ARLEQUIN 3.1 software (Excoffier and Schneider 2005). Sequential Bonferroni corrections were used to compute critical significance levels for multiple HWE tests (Rice 1989). Tests for evidence of significant pairwise linkage disequilibrium between loci were conducted with 3060 permutations using FSTAT 2.9.3 (Goudet 2001).

We estimated population differentiation among the five sampling localities by first estimating traditional pairwise $F_{ST}$ (Weir and Cockerham 1984) in ARLEQUIN. Population differentiation was further tested using a coalescent genetic clustering approach employed in the program STRUCTURE (Pritchard et al. 2000). This analysis makes no a priori assumptions of the individual’s population or sampling localities of origin. We assumed both admixture and correlated allelic frequencies among populations. A burn-in period was chosen based on preliminary analyses, and data was collected for $10^2$ iterations after burn-in. Each test for the number of genetic clusters (K) present in the sampling area was repeated multiple times to determine constancy of likelihood scores. The best estimate of K was selected based on the highest likelihood score.

Another test of panmixia was performed using GENECLASS 2 (Piry et al. 2004). We tested assignment of individuals to each population and for first-generation migrants. If assignment rates per individual were low then we concluded panmixia existed. If population differentiation was present, then level of gene flow was determined by evaluating the rate of migration.

Finally, distance in kilometers from the center of each site and Slatkin’s linearized $F_{ST}$ were tested for correlation using a Mantel test in ARLEQUIN to test for the influence of isolation-by-distance among populations. If genetic differentiation is detected and isolation-by-distance is found, then distance alone may explain the differentiation. With significant differentiation and no isolation-by-distance, then some other factor must be a barrier to gene flow. Distance between sites was calculated using Euclidian distance in ArcGIS (ESRI ArcGIS Version 9.2). Location coordinates were provided with each sample and used to plot capture locations. Using ArcGIS, the center of capture locations at each site was determined and used for calculating distance between sites.

Landscape feature data were examined after the completion of the genetic analyses. We identified site and pathway characteristics using 5 types of land feature data (Table 1). Radiating out 3 km from the center point, a circle (28.3 km$^2$) was
established to determine the site characteristics. Straight lines connecting the center of the sites were created and then buffered to create a 3 km wide straight pathway between sites. This pathway was used to identify landscape features the raccoons may travel through or cross in the process of migrating to another site. These circles and buffered lines were used as a mask to extract data from the land feature types. The distance of 3 km was selected as most raccoons do not disperse more than 3 km from their natal home (Cullingham et al. 2008). Additionally, 3 km encompassed all capture locations for a site.

The 21 land cover classes listed in the National Land Cover Dataset (NLCD) were grouped into 8 land cover classes (i.e., water, developed, forest, shrub, grassland, agriculture, woody wetlands, emergent herbaceous wetlands) for the purposes of our analyses. We determined the percent cover of the 8 classes for each site and pathway. We looked at the minimum, maximum, and mean for each site and pathway for elevation and slope data. Slope was derived from elevation data and provided insight into rapid elevation changes. We analyzed roads by counting the number of pathways or sites crossing the three most used or complex roads. The three road types in descending order of use or complexity were primary highway with limited access (A1), primary highway without limited access (A2), and secondary and connecting roads (A3). For hydrology, we identified rivers and large bodies of water that crossed sites or pathways. Finally, we graphically compared the two dominate land cover classes for the pathways with $F_{ST}$ values from the corresponding pairwise comparisons.

RESULTS
Population Structure
Genotypes were obtained for all 70 individuals. There were no significant deviations from HWE and no significant evidence of linkage disequilibrium. Thus, data from all 11 microsatellite loci were used in downstream analyses.

No population differentiation occurred between LSE and LNW ($F_{ST} = 0.010$, $P = 0.083$) or between LSE and ASW ($F_{ST} = 0.008$, $P=0.122$). Population differentiation was observed between LSE and Hale ($F_{ST} = 0.016$, $P=0.015$) and between LSE and ANE ($F_{ST} = 0.025$, $P = 0.008$). All other pairwise comparisons (i.e., Hale-ANE, Hale-ASW, Hale-LNW, LNW-ANE, LNW-ASW, and ASW-ANE) also showed differentiation ($F_{ST} \geq 0.03$, $P < 0.005$). STRUCTURE analysis indicated 4 genetic clusters over the study area. These genetic clusters (i.e., A, B, C, D) were shared among the sites in

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roads</td>
<td>Processed TIGER 2002 Roads</td>
<td><a href="http://www.census.gov/geo/www/tiger/">http://www.census.gov/geo/www/tiger/</a></td>
</tr>
<tr>
<td>Elevation</td>
<td>Digital Elevation Model (DEM) 30 m</td>
<td><a href="http://www.aces.edu/waterquality/gis_data/index.php">http://www.aces.edu/waterquality/gis_data/index.php</a></td>
</tr>
<tr>
<td>Slope</td>
<td>Digital Elevation Model (DEM) 30 m</td>
<td><a href="http://www.aces.edu/waterquality/gis_data/index.php">http://www.aces.edu/waterquality/gis_data/index.php</a></td>
</tr>
</tbody>
</table>

Table 1: Metadata for land features used for analysis.
different proportions. The cluster that dominated Hale (A=48.8%) was rare (< 5%) in the other four sampling localities. The dominate cluster in LNW (B=45.0%) was in a low proportion (≤ 10%) in the other sampling localities. Clusters C and D were present in all the sites ranging from 15.4% to 64.7% of the total distribution of clusters for each site. For LSE, clusters C and D represented 38.5% and 47.9% of the total distribution, respectively. However, cluster C consisted of only 29.2% for ANE, while cluster D was 61.5%. For ASW clusters C and D were 64.7% and 30.0% of the total distribution, respectively.

The GeneClass analysis revealed that 98.6% of individuals were assigned to the same population from which they were sampled. Five possible migrants were identified; three in Hale from LSE, ASW, and LNW, and two in ANE from LSE and ASW. Distance between sites ranged from 6.2 to 99.5 km. Hale was greater than 80 km from the other sites and the remaining four sites were within 35 km. There was no significant correlation (P = 0.26) between population differentiation and distance, therefore isolation-by-distance did not occur.

Landscape Features

Three sites, ANE, ASW, and Hale, were dominated with the forest land cover class. The dominate class for LNW was woody wetlands and for LSE it was agriculture. Of the 5 sites, Hale was the most homogenous with forest covering 88% while LNW was the most heterozygous with the dominate land class less than 30% of the area (Fig. 2). Each pathway was either dominated by forest or by agriculture. All pathways to and from Hale and the ASW-ANE pathway had greater than 50% of area covered by forest. The other classes for these pathways were less than 21%. The LSE-LNW pathway consisted of 42.1% agriculture followed by woody wetlands at 22.8%. The remaining pathways were a mix of the different classes with the percent coverage for forest and agriculture between 21% and 37%.

Elevation changed less that 200 m over sites and pathways. The highest point, 217 m, was in the Hale-ANE pathway and the lowest point, 24 m, was contained in the ASW site and Hale-ASW pathway. ANE had the highest mean elevation (133.8 m ± 17.7 m SD), while LNW was the site with the lowest mean elevation (47.9 m ± 9.5 m SD). The Hale-ANE pathway mean elevation was 120.6 m (± 27.2 m SD) and was the highest of the pathways, while the LNW-ASW had the lowest mean elevation (46.7 m ± 11.8 m SD). The steepest location was in the Hale-LSE pathway and was 40.4°. All the sites and pathways consisted of flat areas with a slope of 0°. LNW had the lowest mean slope at 1.3° (±2.1° SD), but the largest range (0–30.9°). The LSE-LNW pathway also had a mean slope of 1.3° (±1.7° SD). Hale had the steepest mean slope (6.7° ± 3.6° SD) and the pathway with the steepest mean slope was the Hale-LNW pathway (5.6° ± 4.3° SD).

The number of roads crossing a pathway increased with distance. None of the pathways or sites had A1 roads (i.e., interstates). At least one A2 or A3 road crossed all the pathways with the exception of the LSE-LNW pathway, which had no A2 or A3 road crossings. Pathways to and from Hale had the most (n = 6–7) A2 roads crossings and had a few (n = 2–4) A3 crossings. The ASW-ANE pathway had the most A3 crossings (n=5), but only 2 A2 roads. All other pathways had ≤ 2 A2 crossings and ≤ 3 A3 crossings.

All pathways to and from Hale crossed the Cahaba River and all pathways to and from Lowndes country crossed the Alabama River. Only LNW was dissected by the Alabama River and all animals were captured on the south side of this river. Hale
was the closest site to the Cahaba River and all other sites were near the Alabama River. The Alabama River ranges in width from 400 to 700 feet and the Cahaba River averages 400 feet wide (Owen 1921). The width of the Alabama River near LNW varies from approximately 210 m to 290 m (United States Army Corps of Engineers [USACE] 2008a) and the flow rate averages 566 m$^3$/second (USACE 2008b).

Graphical comparison of the 2 dominant feature classes, forest and agriculture, to the pairwise $F_{st}$ values indicated a trend associated with percent forest cover for the pathways to and from Hale. As the percent forest decreased from 69.4% to 61.7%, the $F_{st}$ values decreased from 0.037 to 0.016. LSE was the only agriculture dominated site and pairwise $F_{st}$ values with LSE (0.008 to 0.02) were the lowest of all the comparisons. The pathways (LSE-ASW, LSE-ANE, LNW-ANE, LNW-ASW) with an almost even mix of forest (21.5–33.2%) and agriculture (25.6%–36.4%) did not show any trend with $F_{st}$ values (Fig. 3).

**DISCUSSION**

The hypothesis of panmixia was rejected and the differentiation that occurred was not found to be correlated to distance alone, suggesting that landscape was influencing gene flow. However, examination of the site and pathway characteristics did not indicate any single landscape feature that explained the observed population differentiation.

Only land cover had an association to genetic differentiation. In a broad sense, increasing forest cover resulted in less gene flow between sites, while increasing agriculture cover increased gene flow. As percent forested area decreased for the pathways to and from Hale, the pairwise difference ($F_{st}$ values) decreased. The pathway between ASW and ANE was also dominated by forest and had a higher $F_{st}$ value than LSE-ANE comparison even though the distance between ANE and

---

**Figure 2.** Percentage coverage of land classes at each of the study sites in central Alabama.
ASW, and between LSE and ANE were similar. The LSE site was the only agriculture site, and comparisons with this site had lower pairwise $F_{st}$ values than comparisons that did not include the LSE site. The results of Root et al. (2009) supported the idea of agricultural areas having increased gene flow; panmixia was observed among the sampled sites, which were located primarily in agriculture habitat.

Besides the forest and agriculture land type relationship, our sampling approach may be an important factor contributing to significant genetic differentiation of Hale to the other populations. Even though we did not find differentiation to be correlated with distance, Hale was 80 km or more from the other sampled sites, while the remaining sites were all within 35 km of each other. Therefore, our sampling design, which did not include areas between Hale and the other sites, may have resulted in a signal of significant genetic differentiation at Hale. If individuals between Hale and the other sites shared allelic diversity with all sites, the signal of population differentiation may not have occurred.

Neither the forest and agriculture relationship nor geographical distance explained the genetic relationship of LNW with the other sites. The LNW site was closest to the centroid of all sites, but it had significant differentiation with three sites and was dominated by a unique genetic cluster that is in low proportion in the other sites. The LNW site is primarily woody wetlands and modeling by Recuenco et al. (2008) indicated that rabies cases decreased for each percent increase in wetlands. Thus, the diversity of the landscape or the higher percentage of woody wetlands may have contributed to higher site fidelity and shorter dispersal distance. However, our ability to make such inferences is limited by our small sample sizes, lack of replication, and our inability to distinguish genetic independence.

Figure 3. The percent forest and agriculture land cover class for each pathway and the corresponding $F_{st}$ values from the pairwise comparisons.
of samples (i.e., samples from different family groups).

Future research should attempt to sample raccoons from areas between Hale and the other sites to determine if the genetic differentiation we observed was an artifact of our sampling design. Also, additional agriculture and forest sites should be sampled to determine if animals originating in forested sites disperse shorter distance than animals originating from agriculture sites, as our results indicate. The influence of additional land cover types on gene flow should also be assessed.

Two other important factors related to rabies transmission that can be influenced by the landscape are density of raccoons and resource distribution. Other research has begun to examine these factors (Arjo et al. 2005, Fisher 2007, Arjo et al. 2008). In our study, where the spread of rabies is limited, raccoon density estimated for some of the sites was relatively low ($\leq 8$ raccoons/km$^2$) (Arjo et al. 2005). However, at an urban national park in Washington, D.C. where rabies spreads more rapidly than in Alabama, density was reported to be $333$ raccoons/km$^2$ (Riley et al. 1998).

Fisher (2007) documented that raccoons maintained core areas in central Alabama that are dominated by hardwood forests. This supports the hypothesis that in agriculture areas raccoons may need to disperse farther to find suitable core habitat. If land cover and habitat type influence movements of raccoons, then ORV zones will need to account for such variation to prevent breaches by infected raccoons.

Our results suggested that landscape was limiting gene flow among raccoon populations and thus the spread of raccoon rabies. This effect was likely due to a combination of landscape features, which this study only began to identify. Future research with increased sampling and the inclusion of data on raccoon density and resource distribution will help to elucidate what combination of landscape features in Alabama has contributed to the slow spread of raccoon rabies. The understanding of this system has broad implications for understanding the spread of rabies in the southeastern U.S. and how landscape features may be more usefully integrated into cost-effective rabies management strategies.

ACKNOWLEDGMENTS

We would like to thank W. Arjo for data collection facilitation. Raccoon tissue samples were provided by C. Fisher, D. Johnson, and P. Coburn. This project was supported by the United States Department of Agriculture, Animal and Health Plant Inspection Service, Wildlife Services, National Rabies Management Program.

LITERATURE CITED


