Genetic Differentiation Led by Geographical Barriers between Populations in *Peromyscus maniculatus.*

Laura Isabel Paulino

Mentored by Dr. Andrew Mahon

UNDERC East

July 19th, 2010

Abstract

Geographical barriers can disrupt mammal dispersal patterns, which may lead to stagnation of their gene pool. Over time this separation can lead to some degree of species differentiation. University of Notre Dame Environmental Research Center provides a unique scenario to test hypotheses related with mice dispersal. In this study, I focus on the small mammal, *Peromyscus maniculatus*, also know as deer mouse. Tenderfoot Creek separates mice populations and may act as a physical barrier, and preventing the dispersal of the deer mice. Mice were collected from different locations around the creek, extracted tissue samples from them, and standard molecular techniques were utilized to elucidate species separation between the different collection locations. Because we did not get results from the PCR amplifications we proceeded to analyze a previously collected series of DNA sequences from a publically available database (GenBank) obtained from different locations in North America. Using these data we found that (1) genetic similarity between individuals did not always correlate to geographic collection area, (2) several individuals from different areas were genetically closer together, (3) geographically farther groups were mostly genetically farther from each other. These results, in general, fit my hypotheses, with exception for the second observation which shows some individuals from different areas that are genetically closely related which suggest that mice may be being transported by human's activities from one area to another.

Introduction

Gene flow plays an important role in the success of natural populations. Important processes such as self-renewal rely on gene flow between groups. Homogeneous populations (not locally

differentiated) provide a large potential of genetic resources with a broad genetic base for use in local restoration activities, and recovering of the population from disturbances (Horning and Webster, 2009). Populations that are connected to each other and that maintain a high diversity within and across groups have a higher potential to be abundant in terms of individuals and to overcome ecological and environmental pressures such as predation and natural disturbances. (Inferred from Coll et al, 1994). Gene flow includes all the mechanisms resulting in the movement of genes from one population to another. Gene flow can occur due to events such as migration, movement of gametes, extinction and recolonization of entire populations, or movement of extra nuclear segments of DNA (Slatkin, 1995). It is clearly important because besides it is one of the mayor evolutionary forces (Bauduer et al, 2009) it directs essential things as restoration, fitness (diseases and external factors resistance, defense mechanisms, etc).

However gene flow may be affected and disrupted by anthropogenic actions as well as by natural geographical and/or physical barriers (mountains, streams, etc). Geographical barriers have the potential of separate population and lead to the arising of new species (Mahon et al. 2008). The effect of physical separations on the genetic diversity of populations has been previously studied (e.g., Mahon et al 2008). From studies such as these, we know that even when populations separated by physical barriers does not decline in size, and the population seems to be healthy; genetic and fitness data indicates presence of genetic drift and interbreeding depression. This may lead to further problems in terms of survival of restricted populations. (Hitchings, 1997). A disruption in the populations' gene flow means a disruption of all the processes that gene flow and genetic exchange directs (resistance, fitness, evolution, etc), so these previous findings help us to understand how crucial a gene flow disturbance can be for populations.

Mammals usually move from their home range dispersing their alleles for specific genes and leading to gene flow across different populations. Barriers on their dispersal may cause stagnation of their gene pools in one area and over time result in species differentiation and speciation (Gering et al. 2009). Small mammals are more prone to be affected by these geographical barriers due to their small home range, and limited capacity to overcome physical obstacles (Kurta, 1995). Hence, in this study I focused on the small mammal, *Peromyscus maniculatus*, commonly known as deer mouse. *Peromyscus maniculatus* is a very small mammal (length: 4.7-6.5 inches). It is widely found in North America, having the most extensive range; it occupies a diverse assortment of habitats across the continent (open areas, forested habitats, shrubby areas, regenerating clear cuts, recent burns, etc) (Kurta, 1952).This small rodent has a variable home range, but most of them travel less than 152 m from the natal area to establish their own home range Stickel 1968). Deer mouse have been already suffering differentiation, therefore there are four subspecies of *P.maniculatus* that inhabit North America.

The University of Notre Dame Environmental Research Center (UNDERC) provides a unique site to test hypotheses related to deer mouse dispersal as related to geographical barriers. Tenderfoot Creek at UNDERC runs from the south end of property to the north, dividing the property in half. This separation could potentially separate mouse populations and may constitute a physical barrier for the mice. However, it is also known that mice can swim, but under what pressures and conditions they do swim is not well known yet. Therefore little is known about how Tenderfoot Creek acts as a barrier for the mice dispersal. In this project I hypothesize that Tenderfoot creek acts an effective barrier in disrupting gene flow between populations around the creek.

Materials and Methods

Study site

I set three trapping grids at three different locations: (1) Southeast Tenderfoot Creek, (2) Southwest Tenderfoot Creek, and (3) Northeast Tenderfoot Creek (Appendix figure 1). The grids were located so they were more than 100 meters away from the road and more than 100 meters away from the creek; insuring the complete home range of the populations was sampled, and that the populations were far enough away from the road that the mice could not be using the road to cross the creek.

Trapping procedure

Each trapping grid consisted of five parallel transects 15m apart from one another. Each transec thad five sites separated 15m apart. At each site one Sherman trap was placed for a total of 25 traps at each grid. I trapped each grid for 6 days until five mice were caught from each grid. Each mouse caught was tagged with ear tags and sample of ear tissue was taken. The individuals were released immediately after sampling. Ear tissues from the ear punches were preserved in 95% ethanol until the DNA could be extracted from them.

DNA extraction &PCR amplification

I extracted the DNA from the mice tissue using a QiagenDNeasy® Tissue Kit (Qiagen Inc., Valencia, CA) following the manufacturer's instructions. A fragment (~1100bp) of the mitochondrial cytochrome b (cytB) gene was amplified using the Cytochrome-b 15334L and L14724 (Gering et al, 2009) primers and a reaction cocktail consisting of 0.75U Taq Polymerase

and 10X PCR buffer, 2.5 mM Mg(OAc)2, 10 nmol of each dNTP, DNA template, primers, and water to 25µl. The PCR cycling program included an initial incubation at 95°C for 2 min and 40 cycles of 94°C for 30 s, 50°C for 1 min, and 72°C for 1 min. This was followed by a final extension at 72°C for 5 min.

Unfortunately, I did not get any positive results from the PCR from the samples around Tenderfoot creek. This could be due to mistakes in the DNA extraction procedure or the inability of the PCR primers to amplify the targeted sequences of the cytochrome B gene. To test hypotheses regarding *Peromyscus maniculatus* dispersal and molecular speciation, I analyzed DNA sequences from a previously published paper (Gering et al, 2009). Forty six cytB sequences, taken from GenBank (http://www.ncbi.nlm.nih.gov); Accession numbers listed in supplementary table (Appendix figure 2) , were from mice collected from different locations in North America.

Data analysis

Using these data, I performed a suite of statistical analyses to analyze haplotype connectivity, and to investigate genetic similarity and gene flow between the samples. I analyzed 46 DNA sequences from Gering et al (2009) paper, each sequence comprised of a total of 1011 base pairs (sites). Sequences were aligned using BioEdit 7.0.1 (Hall 1999) and COI nucleotide sequences were translated to protein sequences to check for sequencing errors. Having aligned the sequences I performed a Tajima's D test and Fu's Fs test to determine if those DNA sequences were evolving neutrally (randomly) or under not random processes (directional selection, balancing selection, etc). To get a visual representation of the individuals analyzed I used a maximum parsimony coalescent calculation in the program TCS (Clement et al. 2000) with a

95% confidence interval, from which the haplotype connectivity between individuals were visualized; the individuals genetically closer related were grouped in the same haplotype networks) the networks were connected in such way that the lines show the relative genetic distances between individuals, the separation between circles represents genetic distances; dots along the lines are the relative number of base pairs that the two circles or individuals differ from each other, and I used different colors to refer to the geographical regions from were the samples were taken.

Finally I used the program Mega 4 (Kumar et al 2008) which calculated the average difference within group, and between groups to obtain the genetic similarity between the individuals. Excel was used to organize the data and convert the uncorrected p genetic distances to percentages.

Results

(1) The Tajima's and Fu's Fs neutrality tests showed that the DNA sequences changes and mutations are evolving randomly (neutrally), not under not random processes such as directional selection, stabilizing selection, demographic contraction, etc. No excess in rare polymorphism were found, hence things as an increase in population or positive selection are not occurring.

Table 1. Fu's Fs and Tajima's D neutrality tests results.

Fu's Fs	-1.30554
Tajima's D	-0.35247

Not significant, P>0.10

(2) The TCS network produced 7 number of groupings (figure 1) .I found that: Individuals genetically more similar are mostly from one area or areas that are close between each other, geographically farther groups were mostly genetically farther from each other. Anyhow the genetic similarity between individuals did not always correlate to geographic collection area; several individuals from different areas were genetically closer together.
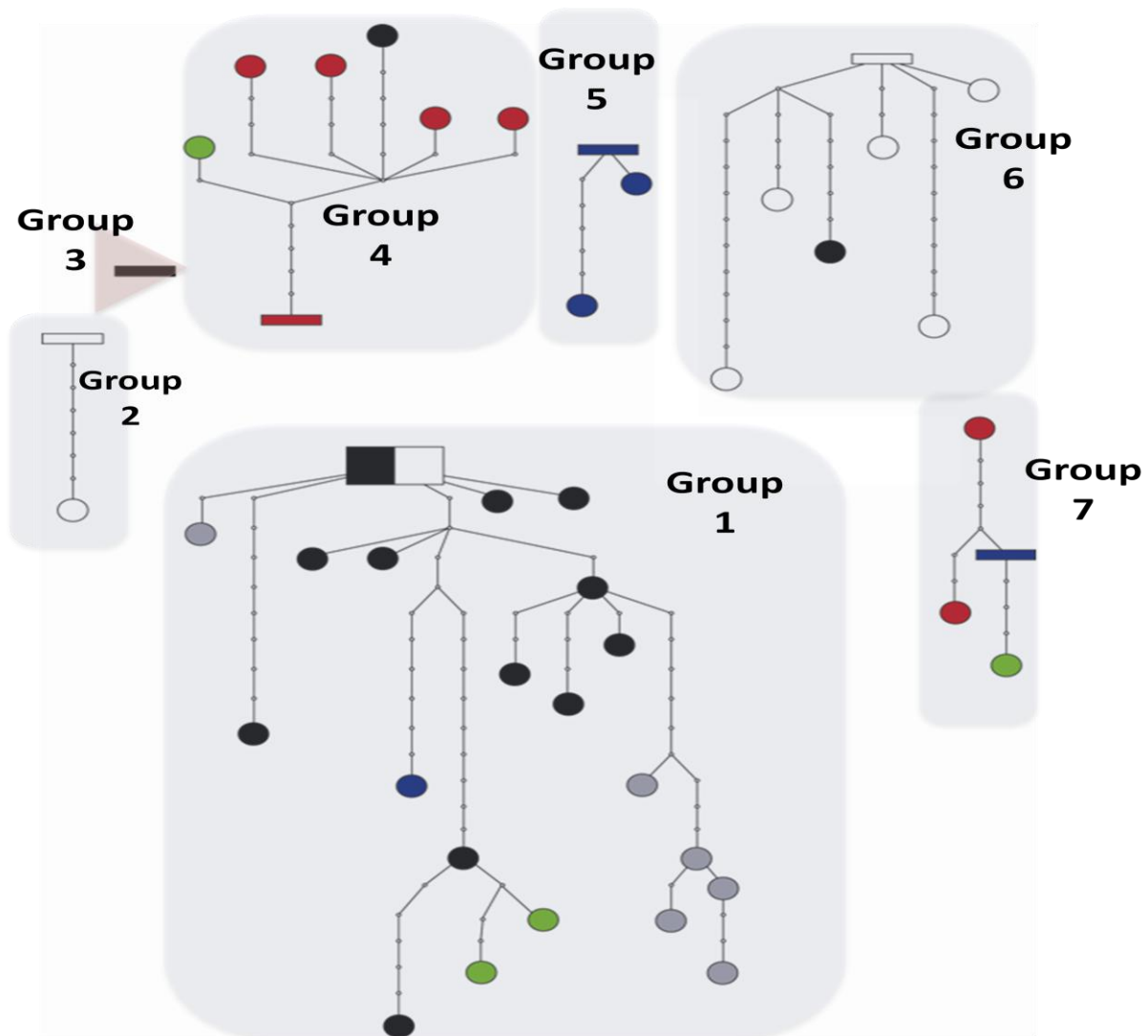

Figure 1. Haplotype connectivity of the analyzed individual. Individuals genetically closer are grouped in the same group or haplotype network. The color represents a geographic region: Red-East Coast US, Blue-North Central

US, Green-Central US, Black-Rocky Mountains, White-California, and Grey-Pacific North West.

Using Mega 4 program I obtained the percent of differentiation between the individuals within each group. The mean within group sequence divergence (uncorrected p distance) was 0.67 %(From Table 1). The net between group genetic distances was found to be 3.4 % (From Table 2).

Table 2. Percent of differentiation between the individuals within each group. Darker colors highlight the highest values.

Group	% within group differentiation	Amount of Individuals
1	1.237	22
2	0.692	2
3	0	1
4	0.725	7
5	0.462	3
6	1.027	7
7	0.577	4

Table 3.Percent of differentiation between the different groups.

Groups	1	2	3	4	5	6	7
1							
2	0.95						
3	4.102	4.006					
4	3.993	4.35	3.834				
5	4.021	4.402	4.088	1.338			
6	3.167	3.535	3.175	3.886	3.839		
7	3.375	3.891	3.742	1.832	2.011	3.836	

Discussion

The genetic-geographic distances relationships indicate that geographical barriers are indeed causing differentiation between populations analyzed in this study (i.e. from Gering et al 2009). The distribution of geographical groups was found predominantly together in the same haplo-group (Figure 1). This suggests that the populations are being effectively separated from each

other by a physical barrier (including distance as a geographic impediment). However, the genetic similarity between individuals did not always correlate to geographic collection area; several individuals from different areas were genetically closer together (Results Figure 1 and Appendix Figure 3).

Because of the small home range of the mice and limited moving abilities I infer that events independent from the mice behaviors and migrations may be leading to that finding, such as transportation of individuals by human activities from one population to another or deeper reasons as convergent evolution that caused by ecological necessities and/or conditions may lead to the evolution of similar traits on distant populations. These suggested explanations should be later tested for a better understanding of the distribution of mice populations.

This work found that the percent sequence differentiation within groups (Table 2) is much lower in value than the differentiation between groups (Table 3). The higher value found between groups indicates potential subspecies differentiation, but as the percent values found were less than 5%, this does not constitute a value that previous researchers would consider distinct genetic species.

There is a notable differentiation between populations (mostly, but not perfectly distributed within the same haplotype networks). This finding reinforces our initial hypothesis and shows that geographical barriers are in fact impeding genetic flow across populations and causing certain degree of divergence. Geographic distance as an effective barrier to dispersal has been limiting the probabilities of mice populations to have contact between each other. We observe that the genetic differences increase as the geographical distances increases. Nevertheless genetic differentiation showed not always to be correlated with geographic distance, i.e. group two and

three are geographically close to each other, but they appear to be highly differentiated. This differentiation may be due to more specific physical barriers as topography and land obstacles. The individual that composes group three may be located in a very inaccessible point in the mountains, enough like to prevent it from having contact with other populations. These differentiations between populations are important because besides that they reinforce my hypothesis about geographical barriers; they show values as high as 4.402. Considering that the percent of differentiation between different species of chipmunks is ~5-10 % for the cytB gene (A. Mahon, personal communication), it is worth to direct our attention to values as 4.402% differentiation between populations of the same species. Speciation may be occurring led by geographical barriers, and we may be able to detect it using molecular tools.

According to my results from the Tajima's D and Fu's Fs test the changes and mutations occurring in these different populations are occurring under random (neutral) processes; which means that no environmental pressures are being exerted to cause the change or evolution of these populations. This is important information in order to understand that these populations are showing patterns of evolution not forced by not random selection.

The results bring us some light in terms of geographical barriers acting as effective obstacles in disrupting the gene flow between populations and causing differentiation. They do not provide a clear understanding of the population dynamics of *Peromyscus maniculatus* here at UNDERC because of the experimental failures experienced during this project. Further studies need to be done in order to test the gene flow and disruption of it at a smaller and more local scale.

Literature cited

Bauduer, F. , Degioanni, A. , and Dutour, O., 2009. Medical genetic polymorphisms as markers of evolutionary forces within the human genome: Hypotheses focusing on natural selection in the basque population. *Human Biology*. pp 23.

Bergek S., and Olsson J., 2009. Spatiotemporal analysis shows stable genetic differentiation and barriers to dispersal in the Eurasian perch (*Percafluviatilis*L.). *Evolutionary Ecology Research*. pp 827–840.

Clement M, Posada D and Crandall K. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9(10): 1657-1660

Coll M, Garcia de Mendoza L, and Roderick GK. 2009. Population structure of a predatory beetle: the importance of gene flow for intertrophic level interactions. *Heredity*. pp 228-236.

Gering E., Opazo J., and Storz J., 2009. Molecular evolution of cytochrome *b* in high- and low-altitude deer mice (genus *Peromyscus*). Nature. pp 226-235.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hitchings S., and Beebee T., 1997. Genetic substructuring as a result of barriers to gene flow in urban *Rana temporaria* (common frog) populations: implications for biodiversity conservation. *Heredity.* pp 117-127.

Horning M., and Webster M., 2009. Conservation Genetics of Remnant Lilium Philadelphicum Populations in the Midwestern United States. *The American Midland Naturalist*. pp- 286.

Kamau L., Lehmann T., Hawley W., Orago A, and Collins F., 1998. MICROGEOGRAPHIC GENETIC DIFFERENTIATION OF *ANOPHELES GAMBIAE* MOSQUITOES FROM ASEMBO BAY, WESTERN KENYA: A COMPARISON WITH KILIFI IN COASTAL KENYA. *The American Society of Tropical Medicine and Hygiene*.pp 64-69.

Kumar S, Dudley J, Nei M & Tamura K (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics*. pp: 299-306.

Kurta A. 1995. MAMMALS of the Great Lakes Region.University of Michigan Press. pp 7, 153-156.

Mahon A., Arango C., Halanych K., 2008. Genetic diversity of *Nymphon*(Arthropoda: Pycnogonida: Nymphonidae) along the Antarctic Peninsula with a focus on *Nymphonaustrale* Hodgson 1902.*Marine Biology*.pp 315-323.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics pp:2496–2497. doi:10.1093/bioinformatics/btg359

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science*. Pp 787-792.

Stickel, Lucille F. 1968. Home range and travels. In: King, John Arthur, ed. Biology of Peromyscus (Rodentia). Special Publication No. 2. Stillwater, OK: The American Society of Mammalogists: 373-411. [25453]

Appendix

Figure 1. Locations of the grids around Tenderfoot creek.

Figure 2. Supplementary table with the sampled localities and Genbank accession numbers of cytochrome b (Gering, et al).

Table S1. Collection localities and Genbank accession numbers for cytochrome *b* sequences used for species-level analyses of *Peromyscus maniculatus*. Phylogenetic relationships and geographic distributions of clades are shown in Figure 1.

LOCALITY	CLADE(S)	GENBANK ACCESSION #
Marin Co., CA	1	EU006766, EU006767, EU006768, EU006769, EU006770, EU006771, EU00672
Monterey Co., CA	3	EF666142
Merced Co., CA	1,3	EF666144, EF666145, EF666146, EF666147, EF666148, EF666149, EF666150, EF666151, EF666152, EF666153
Tuolumne Co., CA	1,3	EF666154, EF666155, EF666156, EF666157, EF666158, EF666159, EF666160, EF666161, EF666162, EF666163, EF666164, EF666165, EF666166, EF666167, EF666168, EF666169, EF666170, EF666171, EF666172, EF666173, EF666174, EF666175, EF666176, EF666177, EF666178
Santa Clara Co., CA	3	EF666143
Mono Co., CA	1,3	EF666179, EF666180, EF666181, EF666182, EF666183, EF666184
White Pine Co., NV	1	EF666185, EF666186, EF666187, EF666188, EF666189, EF666190, EF666191, EF666192, EF666193, EF666194, EF666195, EF666196
Clear Creek Co., CO	1	EF666227, EF666228, EF666229, EF666230, EF666231, EF666232, EF666233, EF666234, EF666235, EF666236, EF666237, EF666238, EF666239, EF666240, EF666241
Yuma Co., CO	1,2	EF666212, EF666213, EF666214, EF666215, EF666216, EF666217, EF666218, EF666219, EF666220, EF666221, EF666222, EF666223, EF666224, EF666225, EF666226
Boulder Co., CO	1	EF666255, EF666256, EF666257, EF666258, EF666270, EF666271, EF666272, EF666273, EF666274, EF666275, EF666276, EF666277
Pawnee Co., KS	2	EF666197, EF666198, EF666199, EF666200, EF666201, EF666202, EF666203, EF666203, EF666204, EF666205, EF666206, EF666207, EF666208, EF666209, EF666210, EF666211
Weld Co., CO	1	EF666242, EF666243, EF666244, EF666245, EF666246, EF666247, EF666248, EF666249, EF666250, EF666251, EF666252
British Columbia,	1	DQ385698, DQ385699, DQ385700, DQ385701, DQ385702

Canada		
Labrador, Canada	5	DQ385809, DQ385810, DQ385811, DQ385812, DQ385813, DQ385814, DQ385815
Manitoba, Canada	6	DQ385723, DQ385724, DQ385725, DQ385726, DQ385727, DQ385728, DQ385729, DQ385730, DQ385731, DQ385732, DQ385733
Ontario, Canada	5	DQ385734, DQ385735, DQ385770, DQ385771, DQ385772, DQ385773, DQ385774, DQ385775, DQ385776, DQ385777, DQ385778, DQ385779, DQ385780, DQ385781, DQ385782
Quebec, Canada	5	DQ385783, DQ385784, DQ385785, DQ385805, DQ385806, DQ385807, DQ385808, DQ385749, DQ385750, DQ385751, DQ385752, DQ385753, DQ385800, DQ385801, DQ385802, DQ385803, DQ385804, DQ385795, DQ385796, DQ385797, DQ385798, DQ385799, DQ385792, DQ385793, DQ385794, DQ385789, DQ385790, DQ385791, DQ385786, DQ385787, DQ385788
Baja CA, Mexico	3	DQ385706, DQ385707, DQ385708, DQ385709
Humboldt Co., CA	1	DQ385703, DQ385704, DQ385705
Santa Barbara Co., CA	3	DQ385711, DQ385712, DQ385710, DQ385713, DQ385714
Ada Co., ID	1	DQ385672, DQ385673, DQ385674, DQ385675
Des Moines Co., IA	2	DQ385633, DQ385634, DQ385635, DQ385636, DQ385637
Penobscot Co., ME	6	DQ385754, DQ385755, DQ385756
Cass Co., MN	6	DQ385736, DQ385737, DQ385738, DQ385739
Keweenaw Co., MI	1	DQ385643, DQ385644, DQ385645
Mackinac Co., MI	5	DQ385821, DQ385822, DQ385823, DQ385824, DQ385825, DQ385826, DQ385827
Gallatin Co., MT	1	DQ385676, DQ385677, DQ385678, DQ385679

Churchill Co., NV	1	DQ385715, DQ385680, DQ385681, DQ385682
Hidalgo Co., NM	2,4	DQ385718, DQ385719, DQ385720, DQ385721
Bernalillo Co., NM	2	DQ385753, DQ385754, DQ358655, DQ385756
San Juan Co., NM	2	DQ385667, DQ385668, DQ385669, DQ385670, DQ385671
Catron Co., NM	2	DQ385651, DQ385652
Otero Co., NM	2,4	DQ385658, DQ385659, DQ385660, DQ385661, DQ385662, DQ385663, DQ385664, DQ385665, DQ385666, DQ385722
Roosevelt Co., NM	2,4	DQ385628, DQ385629, DQ385630, DQ385631, DQ385632
Sullivan Co., NY	6	DQ385740, DQ385741, DQ385742
Clinton Co., NY	6	DQ385746, DQ385747, DQ385748
Somerset Co., PA	6	DQ385762, DQ385763, DQ385764, DQ385765, DQ385766
Sevier Co., TN	5	DQ385816, DQ385817, DQ385818
Unicoi Co., TN	6	DQ385767, DQ385768, DQ385769, DQ385819, DQ385820
Salt Lake Co., UT	1	DQ385684, DQ385685, DQ385686, DQ385687, DQ385688
Madison Co., VA	6	DQ385757, DQ385758, DQ385759, DQ385760, DQ385761
King Co., WA	1	DQ385696
Kittitas Co., WA	1	DQ385689
Pierce Co., WA	1	DQ385697, DQ385716
Snohomish Co., WA	1	DQ385695
Kitsap Co., WA	1	DQ385694
Cochise Co., AZ	*Peromyscus melanotis*	EU574689-EU574701
Leon Co., FL	*Peromyscus gossypinus*	DQ385624, DQ385225

Figure 3. Sampled locations. Obtained from the previous supplementary table (Appendix Figure 2) Colors match the figure 1 (results).