Genetic analyses of Eastern Coyote (*Canis latrans*) in Nova Scotia

Don Stewart and Soren Bondrup-Nielsen
Department of Biology ,
Acadia University

Report on 2016 NSHCF Grant.

Project goals and objectives

This project focused on the following objectives:
Objective 1 - To optimize amplification of nuclear DNA markers for coyotes and related canids (8 microsatellite primer pairs) (Honours student project);
Objective 2 - To determine the percentage of each of the three canids (*Canis lupus, Canis lycaon,* and *Canis latrans*) present in the gene pool of Nova Scotia compared to that of Atlantic Canada generally (using the microsatellite markers);
Objective 3 - To correlate this information with data on mtDNA and body size already collected by former members of the research team (see Power et al. 2015);

Work completed

A recent study by Power et al. (2015) demonstrated a trend among male Coyotes in Atlantic Canada that contained an Eastern Wolf (*Canis lycaon*) haplotype and a larger overall body size as measured by principal components analysis (PCA). The purpose of this study was to determine if there is a significant correlation between Wolf mitochondrial DNA (mtDNA) haplotypes or nuclear DNA (nuDNA) (as assessed using microsatellite markers) and larger body size among male Coyotes within NS. To investigate this potential trend exclusively male samples were studied, and nuDNA was utilized to complement the mtDNA data. DNA was isolated from the Coyote skin tissue (ear), and specific haplotypes of the mtDNA control region were amplified and sequenced. After sequencing, the mtDNA was identified to contain one of two Coyote haplotypes (Cla28, Cla29), or an Eastern Wolf haplotype (GL20). Each Coyote sample was assigned a haplotype using phylogenetic analysis, and linked to its associated morphological meta-data.

The mtDNA haplotype diversity remained consistent with Power et al. (2015) within NS; however, to determine if a size correlation among haplotypes exists Tukey's Test must be completed on the PCA results. Non-coding nuDNA was amplified using eight microsatellite primers, and allele lengths were used to genotype individuals. Three genetic clusters were found within NS male Coyotes using STRUCTURE, which are speculated to be of Coyote, Gray Wolf (*Canis lupus*), and Eastern Wolf origin. All three genetic clusters were present in each Coyote supporting genetic admixture among *Canis* types. Assortative mating may be occurring among NS populations, as all clusters existed in high proportions within certain individuals. Furthermore, nuDNA microsatellites revealed a decrease in heterozygosity within NS Coyotes, as a possible result of inbreeding. This study was the first to explore microsatellites within NS Coyotes, thus the data will provide essential data for future analyses, possibly explaining the larger body size among NS and Northeastern Coyotes.

Results in 2016 (From Honours Thesis of Colton Burke with analytical assistance from former MSc candidate Natalie LeBlanc).

The NS male Coyotes were sorted by haplotype (*i.e*., either the Western Coyote haplotype Cla28 or the Eastern Wolf haplotype GL20). The mean ($\bar{x}$) was the measure of central tendency used to represent all nine morphological measurements taken. The mean of each measurement (± one standard deviation), sorted by haplotype, was taken for the following variables: skinned mass (kg), chest girth (cm), body length (cm), tail length (cm), front paw width and length (mm), rear paw width and length (mm) and shoulder height (cm) (Table 1). When comparing individual measurements by haplotype, all the means were quite comparable and the Eastern Wolf haplotype (GL20) did not appear to contribute to a larger body size (Table 1).

Table 1. Means of various NS male Coyote measurements ± 1 S.D., grouped per mtDNA haplotype.

Haplotype	Carcass Mass (kg)	Chest Girth (cm)	Body Length (cm)	Tail Length (cm)	Front Paw Width (mm)	Front Paw Length (mm)	Rear Paw Width (mm)	Rear Paw Length (mm)	Shoulder Height (cm)
Cla28	13.36 ± 2.60	50.47 ± 4.40	91.43 ± 4.64	37.02 ± 3.35	46.81 ± 3.98	60.68 ± 4.22	41.75 ± 5.35	55.90 ± 3.18	52.52 ± 2.90
	n=47	*n*=47	*n*=47	*n*=47	*n*=43	*n*=43	*n*=47	*n*=47	*n*=41
GL20	13.64 ± 2.44	50.09 ± 4.14	92.97 ± 4.87	37.07 ± 2.42	47.52 ± 3.83	60.47 ± 3.69	42.26 ± 4.52	55.70 ± 4.91	52.78 ± 3.33
	n=30	*n*=29	*n*=30	*n*=30	*n*=27	*n*=27	*n*=29	*n*=29	*n*=26

To determine if there was a correlation between any given morphological measurement and considering the Eastern Wolf (GL20) versus Western Coyote (Cla28) haplotype, univariate statistical analyses were performed. A student t-test was used to determine if there was a significant difference in each morphological measurement between the two groups as defined by haplotype. To determine if the variances were equal or unequal, an F-test was first completed for each measurement (Table 2). Following the student t-test no significant differences were observed between haplotype and each size measurement, as the p-values were all significantly larger than 0.05 (Table 2).

Table 2. Univariate analyses to check for a correlation between haplotype and Coyote individual body size means. An F-Test was used to determine if the variances were equal among haplotypes before completing student t-tests, which used a significance level of α=0.05. No significant differences were observed.

	Carcass Mass	Chest Girth	Body Length	Tail Length	Front Paw Width	Front Paw Length	Rear Paw Width	Rear Paw Length	Shoulder Height
F-test									
F_o	1.1363	1.0878	1.1012	1.9113	1.0619	1.2885	1.3562	2.4757	1.3257
F_α	1.7856	1.8005	1.7856	1.7856	1.8462	1.8462	1.8005	1.8005	1.8718
$v_1 \square v_2$	=	=	=	≠	=	=	=	≠	=
Student t-test									
T	-0.4620	0.3903	-0.4898	-0.0689	-0.8391	0.1705	-0.4589	0.1722	-0.3871
P	0.6454	0.6974	0.6257	0.9452	0.4044	0.8651	0.6476	0.8641	0.7
df	75	74	75	74	68	68	74	42	65

To compare the average mass of NS male Coyotes to other NE Coyote studies the skinned carcass mass had to be converted into a non-skinned Coyote mass (Table 3). This was done using the following formula: whole mass = 1.20 (skinned mass) + 0.35, $r^2 = 0.970$, $p < 0.001$ (Nelson and Lloyd, 2005). The mean ($\overline{x}$) masses of both the skinned and non-skinned Coyotes (± one standard deviation) were then calculated, as well as the local minima and maxima (Table 3). The maximum male NE Coyote mass, 24.27 kg, was notably larger than the mean of 16.32 ± 3.01 kg (Table 3).

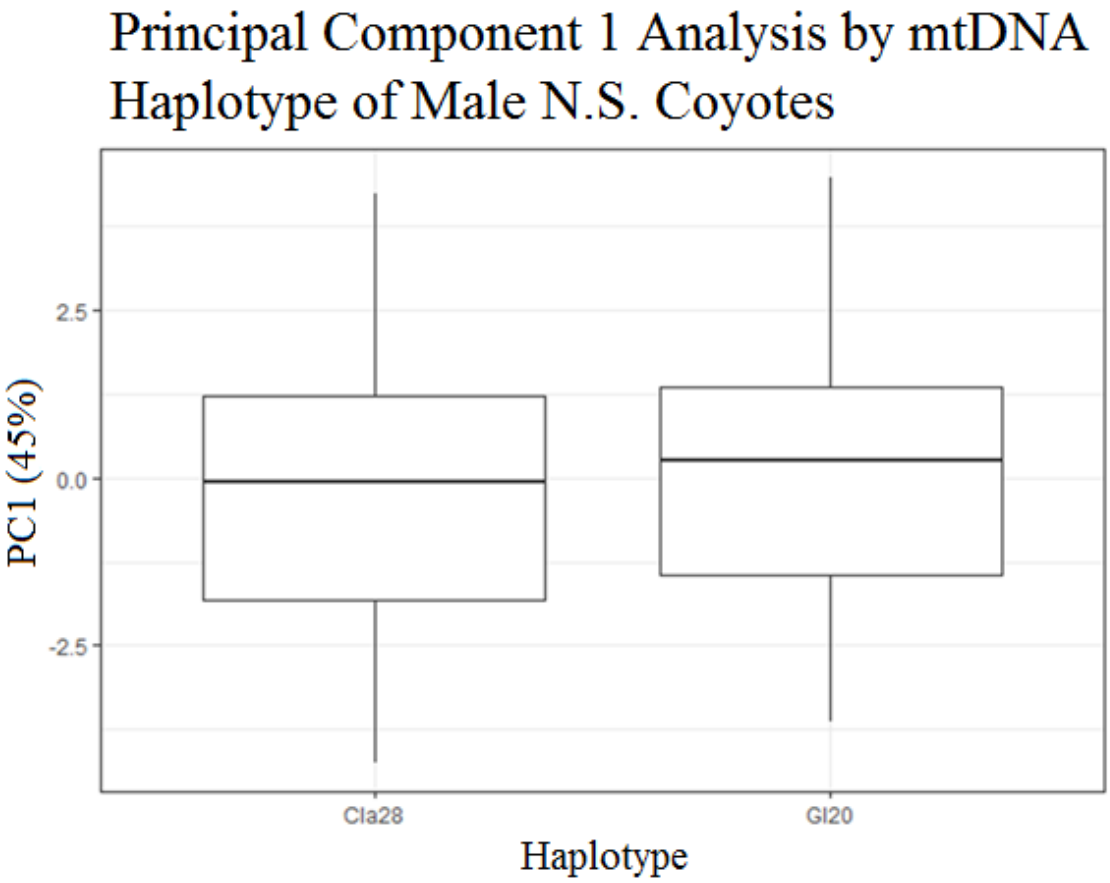
Table 3. Mean carcass and approximate non-skinned male Coyote masses, as per the formula presented by Nelson and Lloyd (2005) ± 1 S.D. All samples were from NS, *n*=78.

	Non-Skinned Coyote Mass(kg)	Skinned Carcass Mass(kg)
$\overline{x}$ ± 1 S.D.	16.32 ± 3.01	13.45 ± 2.51
Local min	9.14	7.48
Local max	24.27	20.09

To explore whether there was a slight tendency for overall larger body size of Coyotes possessing the GL20 haplotype a multivariate principal component analysis (PCA) was performed on the samples using *R*. Multivariate statistics allowed all nine morphological

measurements to represent an overall Coyote body size (e.g., Rising and Somers, 1989). Following PCA, a boxplot could then be drawn to compare general body size between haplotypes (Figure 1). *Note*- any missing data were replaced with the associated mean for that measurement. Principal component one was calculated to represent 45% of the total variation. The box (interquartile range) represents 50% of the data, whereas each whisker represents 25%. No extreme outliers were present in the data set. The Coyotes containing the GL20 mtDNA haplotype could have a larger body size (Figure 1), but whether this difference is significant will need to be evaluated using Tukey's test (e.g., Power et al., 2015).

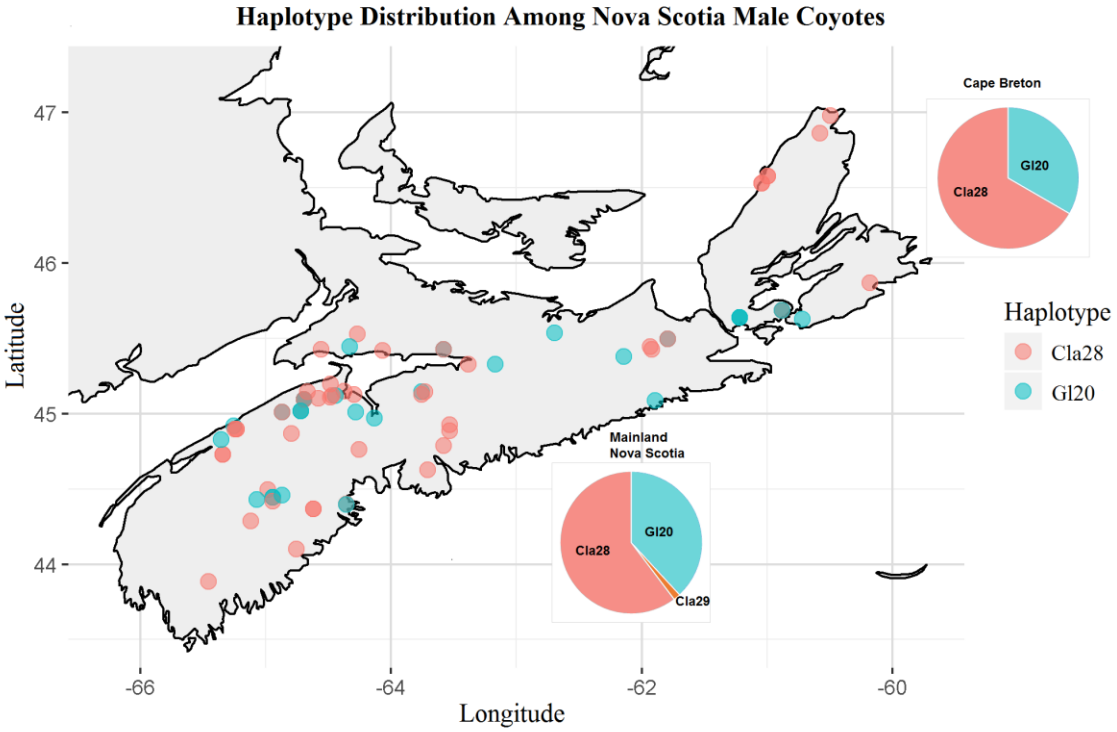
Figure 1. Multivariate boxplot demonstrating first principal component (PC1) values by haplotypes (Cla28, GL20) for NS male Coyotes. Nine morphological measurements were used (skinned mass, chest girth, body length, tail length, front paw width, front paw length, rear paw width, rear paw length, shoulder height) which represented 45% of the total measurements. *n*=78. All measurements were positive and loaded in the same direction; thus PC1 represents general differences in overall body size.



For each Coyote carcass, map location data were provided from NSDNR. The location data were converted to UTM coordinates, which allowed the individuals to be mapped by haplotype using *R* (Figure 2). Plotting the locations where each Coyote haplotype was found

allowed for comparison to recent studies (e.g., Power et al., 2015). Furthermore, pie charts were drawn to determine the haplotype distribution in both mainland NS and CB (Figure 2). The Cla29 mtDNA haplotype (*n*=1) was only observed on the mainland of NS and was not found on CB (Figure 2). The proportion Coyotes containing the GL20 haplotype remained consistent across all of NS (Figure 2).

Figure 2. Mitochondrial DNA haplotype distributions of male Coyotes of both mainland NS (*n*=63) and CB (*n*=15).



GENEPOP (Raymond and Rousette, 1995) was used to perform general probability tests on the microsatellite data. *GENEPOP* calculated allele proportions, expected number of homozygotes and heterozygotes, allele frequencies, and F_{is} (i.e., inbreeding) values for each locus (See Table 1.1 [a-g] and Appendix 1, Burke 2016). Once the allele proportions were calculated, the microsatellite sizes were then corrected via a process called "binning".

The computer program *STRUCTURE* (Pritchard et al., 2000) was used to determine the number of nuDNA clusters and to compare the genetic composition and diversity among Coyotes. Two of the eight loci were omitted as they were not fully optimized and did not genotype successfully for most samples. Using *STRUCTURE*, a job was run with the following set of parameters: a 10000 "Burn-in Period" (followed by 20000 Reps for each test), k (number of clusters) values were from one to six, and each run consisted of 20 iterations. *STUCTURE Harvester* (Earl and von Holdt, 2012) was used to evaluate the *STRUCTURE* output. It was determined that the ideal k value was three clusters (Figure 3 [a-b]). This Δk method was utilized as it was deemed the most accurate by Evanno *et al*. (2005) (Figure 3.3.a). The L(k)

method was also employed (Figure 3.b). Upon selecting a k value, a bar plot could be selected within *STRUCTURE* (Figure 4). A Neighbour-Joining tree was drawn of the three clusters using *STRUCTURE* (Figure 5). Each nuDNA clusters was present in approximately equal proportions, and no individuals were genetically "pure" (Figure 4).

Figure 3.a. The most likely k was determined by *STRUCTURE Harvester* to be k=3, using the Δk Evanno *et al.* (2005) method, Δk = [x̄(|L"(K)|) / S.D. (L(K))].

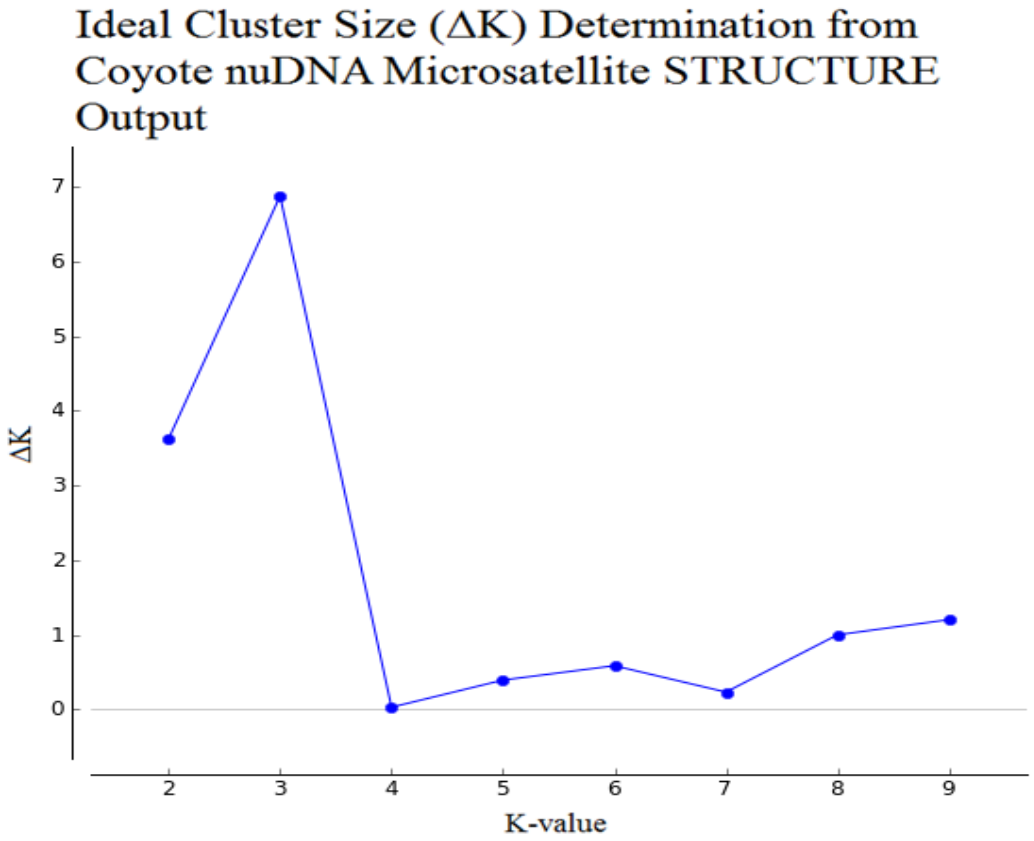


Figure 3.b. The most likely k was determined by *STRUCTURE Harvester* to be k=3, using the L(k) method, L(k) = $\bar{x}$ ± S.D.

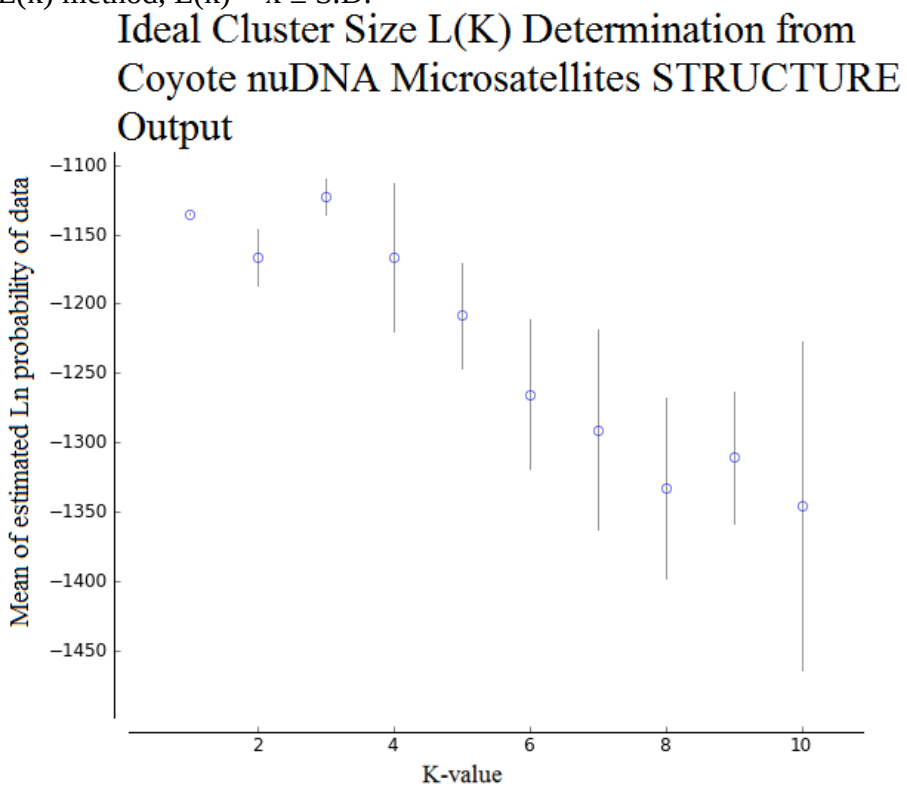


Figure 4. Bar graphs drawn using *STRUCTURE*, k=3, of NS male Coyote nuDNA assessed using six loci *n*=79, sorted by the probability of each individual belonging its cluster, or Q, (Top) and by input order (Bottom).

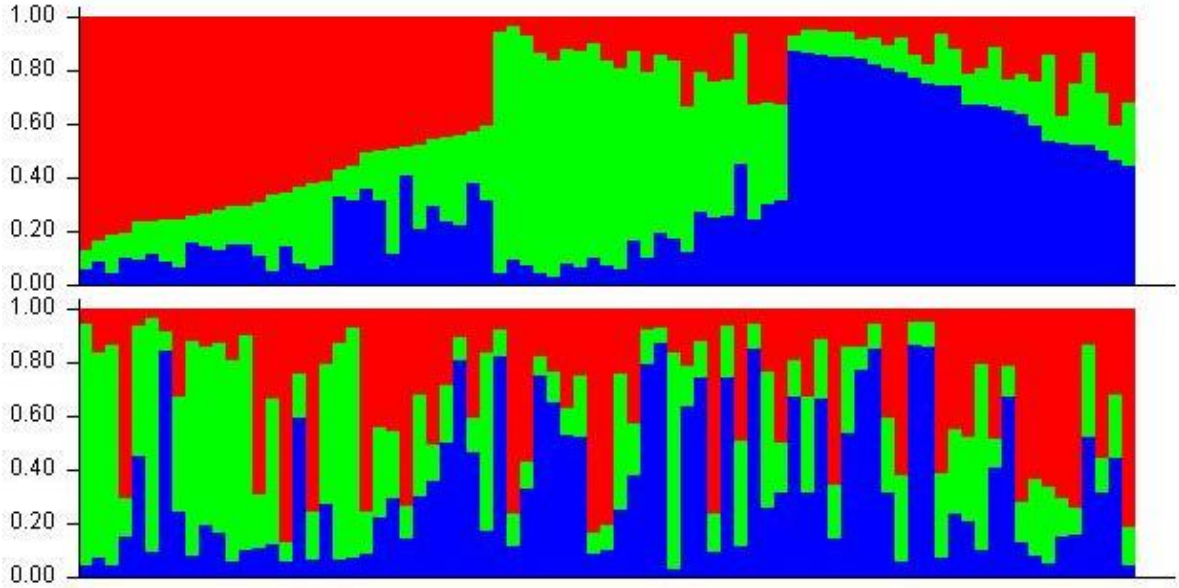
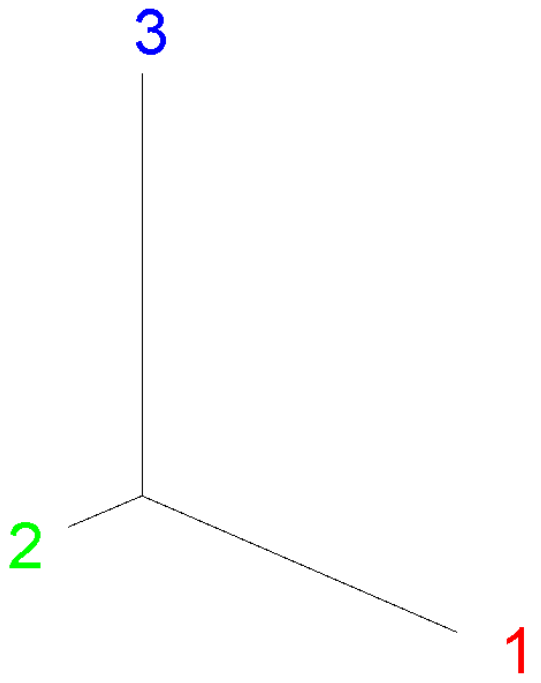


Figure 5. Neighbour-Joining Tree produced using STRUCTURE for k=3, and *n*=79, using PHYLIP (Felsenstein, 2005) and the Neighbour-Joining-algorithm (Saitou and Nei, 1987).



The trees show a representation of the genetic distances among the K Structure clusters. The trees are computed by applying the Neighbour-Joining algorithm to the matrix of allele-frequency divergence among clusters (net nucleotide distance).

The tree was estimated using the program NEIGHBOR by Mary Kuhner and John Yamato, implementing Saitou and Nei's "Neighbour-Joining Method". The plot was produced using DRAWTREE by Joe Felsenstein. Both programs are distributed by Joe Felsenstein as part of his PHYLIP phylogeny package.

Achievements and lessons learned

No significant correlation was observed between individual size measurements and a particular mtDNA haplotype, however, this may not be the case for the nuDNA. In a future study, the *STRUCTURE* bar plot should be overlaid with the carcass size measurements. This could indicate a possible trend with the proportion of a specific cluster with a larger body size. For instance, maybe Coyotes with an increased body size could be related to the proportion of Gray Wolf nuDNA assessed using *STRUCTURE*. Furthermore, PCA could also be applied to the microsatellite *STRUCTURE* results as this would allow individuals to be assigned to a specific genetic cluster (i.e., Coyote, Eastern Wolf, or Gray Wolf) or to be classified as definite admixed individuals (e.g., Kays et al., 2010; Benson et al., 2012).

Follow up recommendations

Given that the amplification of microsatellite alleles has now been optimized for Coyotes, additional analyses should be conducted in the future to determine if there are

differences in other biologically significant factors. For example, the NSDNR collects data on kidney and bone marrow fat stores, as well as data on parasite loads and uterine scars. In order to construct a more complete understanding on the contribution of genetic material from various canids to these variables, additional meta-data available from the NSDNR could be analyzed.