**Methods**

Data set

X samples in 29 pops

Collection methods and bioinformatics pipeline

*Phylogenetic analysis*

All analyses were conducted in R version 4.0.2. Final pipeline files in .vcf format were imported into R and converted first to a genind object, then to a genpop object with population annotations using the package **adegenet** (citation). Distance measures were calculated using the dist.genpop function in the package **adegenet** with options as appropriate for Nei’s and Reynolds’ distance measures. To test if different distances measures ranked pairwise site distances similarly, I conducted a Kendall concordance test using the kendall.global function in **vegan** (citation), where the two distance measures were the judges ranking average pairwise distances for each site. The null hypothesis for a Kendall’s concordance test is that the ranking of sites by the two judges is no more concordant than would be expected by random chance (citation).

Geographic distances between sites were calculated based on latitude and longitude recorded using a Trimble Geo 7x handheld GPS unit during field work. Geodesic distances were calculated using the geodist function in the package **geodist** (citation). To test if the distance matrices were correlated with each other, I used a Mantel test from the R package **vegan** (citation). Several papers in recent years have argued that the Mantel test is not appropriate for assessing correlation between two distance matrices (citations), however this is still standard practice and therefore I use it here. The null hypothesis for a Mantel test is that the correlation between two matrices is zero (citation).

Trees were built using the nj function in the **ape** package for unrooted neighbor-joining trees, and hclust from **stats** and as.phylo from **ape** for UPGMA trees (citations). Cophenetic distances were calculated using cophenetic from **stats**. Finally, bootstrap analyses were conducted with 1,000 replicates on unrooted trees using boot.phylo in the **ape** package.

**Results**

*Distance measures*

A comparison of pairwise distances between populations using Nei’s distance and Reynolds’s distance formulas suggest several similarities and some differences. Both distance measures indicate that the populations with the largest pairwise distances are Bearpaw Lake Intersection, Lozier Road, and Buffalo Fork (Figure heatmap). Kendall’s coefficient of concordance indicates that the ranking of sites based on average pairwise distances is more concordant than would be expected by random chance (W stat = 0.092, p-value = 0.0001 on 9,999 permutations). While the two distance measures rank the sites similarly, the difference between the largest distances and the smallest distances is larger for the Nei’s distance calculations than for the Reynolds’. We expect trees based on the Nei’s distance matrix, therefore, to differentiate between sites more clearly.

Neither Nei’s distance nor Reynolds’ distance were associated with geodesic distance between sites, however both measures of genetic distance were closely associated (Figure distance comparisons). Mantel test results comparing 3 distance matrices as follows:

* Nei to Reynolds: z-statistic = 4.53, p-value = 0.0001
* Nei to geodesic: z-statistic = 201936, p-value 0.445
* Reynolds to geodesic: z-statistic = 2430009, p-value = 0.808

*Tree algorithms*

Cophenetic plots depict the population distance as calculated in the distance matrix against the tree distance produced by the final dendrogram. Cophenetic plots for the neighbor-joining tree and the UPGMA tree for both Nei’s distance and Reynolds’ distance indicate that the neighbor-joining trees better capture the distances in the population distance matrices (Figure cophenetics). The variability in the spread of the plots is slightly larger in the Nei’s NJ tree versus the Reynolds’ NJ tree, but both UPGMA trees assign similar tree distances to very different population distances. The inability of the UPGMA trees to capture population distances increases as the population distances increase.

This result is not surprising, for two reasons. The UPGMA tree is a rooted tree, and the algorithm is therefore more constrained than the neighbor-joining algorithm that doesn’t assume a particular root. The UPGMA procedure also produces a tree in which every leaf is equidistant from the root. Because the NJ tree is not constrained in these ways, the resulting tree is better able to capture the distances in the original distance matrix.

*Phylogenetic trees*

Because the UPGMA trees were not able to capture the population distances very well, I report only the results of the neighbor-joining trees. Bootstrap support for most nodes on both trees was very weak. For the tree constructed using Nei’s distance measure, of the 27 nodes in the dendrogram fully 17 had zero bootstrap support. On 1000 permutations, support for other nodes ranged from 5 to 38. For the tree constructed using Reynolds’ distance measure, 13 nodes (of 27) had bootstrap support of zero while support for other nodes ranged from 1 to 38.

Keeping in mind that the bootstrap support was very weak for all clades in both trees, for those clades with non-zero bootstrap support five were identified in both trees (eleven total populations), one was identified in the Reynolds’s tree but not the Nei’s tree (2 populations), and six populations were involved in clades that differed topographically between the trees (Figure both trees).

**Discussion**