Abstract
We introduce some genetic statistics, mainly the range expansion statistic $\psi$, efficiently computed from the allele frequency files of populations, to explore some perspectives of the origin of a range expansion, and leverage these genetic statistics to construct a potential path of migration for Polynesian settlement of Hawaii.

Genetic Statistics

- $F_2$ Genetic Drift Statistic
  \[ F_2(A, B) = \left( \hat{\beta}_A - \hat{\beta}_B \right)^2 \frac{\hat{\beta}_A(1 - \hat{\beta}_A) \cdot \hat{\beta}_B(1 - \hat{\beta}_B)}{\hat{n}_A - 1} \cdot \frac{\hat{n}_B - 1}{\hat{n}_A} \]

- $F_3$ Shared Drift Statistic
  \[ F_3(A, B; O) = (\hat{\beta}_A - \hat{\beta}_O)(\hat{\beta}_B - \hat{\beta}_O) - \hat{\beta}_O(1 - \hat{\beta}_O) \cdot \frac{\hat{n}_O - 1}{\hat{n}_O} \]

- Average Number of Pairwise Differences (Nucleotide Diversity)
  \[ \pi(A, B) = \hat{\beta}_A(1 - \hat{\beta}_B) + \hat{\beta}_B(1 - \hat{\beta}_A) \]

- Range Expansion Statistic (Directionality Index)
  \[ \psi(A, B; O) = \hat{\beta}_A^{(O)} - \hat{\beta}_B^{(O)} \]

Method

- For all genetic computations, the correlation of nearby SNPs (single nucleotide polymorphisms) is induced by chromosomal crossover. We use a block bootstrap to estimate standard error by dividing the data into consecutive blocks and resampling the entire blocks in the bootstrap.

- $F_3$ Only We need to select one/more outgroup populations for computation.

- $\psi$ Only When two populations have unequal sample sizes at locus $j$, we usually downsample the large sample to the smaller sample size, which follows a hypergeometric distribution. In reality, we pick a global downsampling size such that both populations can downsample to a fixed size at all loci.

Future Work

- We will test with different hyperparameters on outgroup populations, derived allele frequency, and downsampling size to see whether the directionality output/signs ($\psi$) is robust to changes.

- We will convert the code from R to Python via ARCH for block bootstrap and Ray for parallelism.

Reference
