

### Morning Exercises

1. Based on Mal'ta1's relationship with ancient Europeans (Kostenki14, Loschbour) and present-day East Asians (Han, Dai), would you argue Mal'ta1 is an early European, early Asian, or neither?
2. In the Haak et al (2015) study, the ancestry related to the Yamnaya from the steppe region plays a role in Western Europe. Using sets from the Early, Middle, and Late Neolithic period, we will analyze the relationship between Yamnaya and Neolithic European populations.
  - a. Which populations show the closest relationship to Yamnaya using the outgroup  $f_3$ -statistic?
  - b. Focusing on ancient Europeans (Loschbour, Stuttgart, Spain\_MN, Corded\_Ware\_LN, and Unetice\_EBA), which of these share the closest relationship to Yamnaya, if any? Use both  $f_3$ - and D-statistics to make your hypothesis.
3. Try re-running our D-statistic analysis for Mal'ta1 using the chimp as outgroup. Are results robust to the outgroup?
4. For ancient DNA, post-mortem damage is more likely to affect transition mutations instead of transversion mutations. Thus, it is often standard practice to remove transitions and re-run results to check for consistency. What command in Admixtools allows you to filter your SNPs? Look at their documentation and use ipython to explore options for `ff.parfile()`. How in UNIX could you create the necessary files? Why might filtering for only transversions be problematic in this dataset?
5. Based on the heatmap we created in class, we find that some groups cluster together. Pick a population that has not been discussed, construct a hypothesis on whom they may form a clade with, relative to who, and calculate the D-statistics to support or reject your hypothesis.

### Afternoon Exercises

1. Middle Neolithic Europeans were found to have experienced gene flow European hunter-gathering populations (i.e. Loschbour). Can you detect evidence of such gene flow using  $f_3$ -statistics?
2. Run a two source qpAdm analysis for the Late Neolithic populations – can you determine mixture proportions represented by Early Neolithic populations and steppe populations?
3. European populations by the Late Neolithic are actually a mixture of three different ancestries – ancient European hunter-gatherer, a basal Eurasian ancestry that arrived in the Early Neolithic to Europe, and steppe-related ancestry related to the Yamnaya. Try developing a three source mixture model to test using qpWave and qpAdm. Can you find a combination that satisfies a three source mixture model? If not, what are the possible mixture models you do find?
4. For #2 and #3, the focus on European and steppe populations suggests that more outgroups can potentially be added. What are further outgroups that might be appropriate? Does this help resolve problems regarding statistical power?