**Lesson 2 Exercises**

1. Write a short paragraph describing what we know about the genetic ancestry in Europe leading into the Haak et al. (2015) paper.
2. Using the IND file (look in /public/adna/student/data/), determine how many samples were separated from their group, and note why they were separated? Hint: Look at the population ID label in the ‘ind’ file.
3. For your previously published individual assigned at the end of class, read the corresponding paper. Add a row to the INFO file we started in class, with information for that individual/set included as well.
4. Using the info file developed in class, write a python function **getinfo\_anc** to retrieve the information as needed. Ideally, this script can be used to quickly find the set of individuals belonging to a certain category. (Example: Can this function in one line list all sets found from Middle Bronze Age Germany?)
5. Write a python function **makepar** to make a CONVERTF PAR file quickly, when given arguments for the filename and datatype to convert to (and whatever other options you might want).
6. Store **makepar** and **getinfo\_anc** in a single python file, titled “popgen\_funcs.py”. In jupyter notebook, type ‘import popgen\_funcs’ and try to use your functions without keeping them inside your script.

BONUS: What present-day populations are included in the dataset? Divide the **present-day populations** into different groups by region, and store them in a text file that will make these groups easy to retrieve.