

Genetic diversity of chimpanzees' species

Here we study the genetic diversity of a subset of data about chimpanzees and bonobos from Manuel et al. (2016 Science). The data are in the form of a genotypic data matrix where each entry is either 0, 1 or 2 for the number of alternative allele per site.

The expected mean heterozygosity is calculated across the SNPs for each species as the probability to take two sequences with different alleles based on the allele frequency of the population

The observed mean heterozygosity is the actual proportion of heterozygotes sites (i.e. number of 1's) per individual in a population.

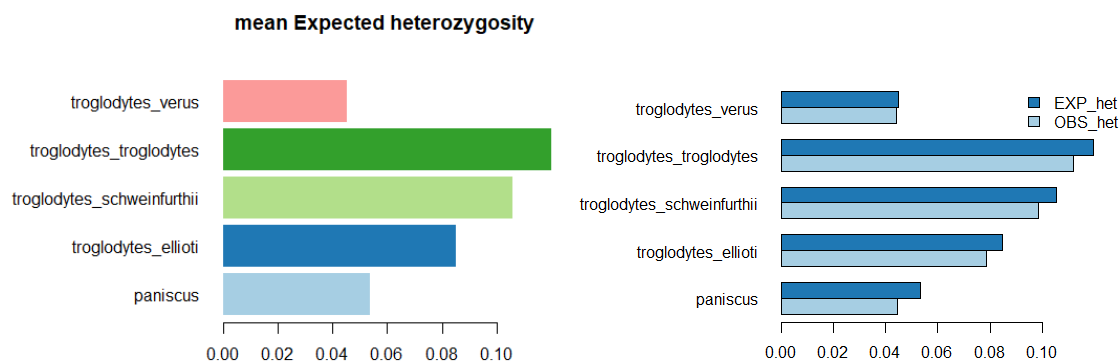


Figure 1 (left) Mean expected Heterozygosity (right) Expected vs observed heterozygosity

The species with the higher Expected and Observed heterozygosity is *P. t. troglodytes*, chimps from Gabon and Congo, meaning they have the highest genetic diversity. The observed heterozygosity is lower than the expected one in all the population because the expectation is made under Hardy-Weinberg assumption which does not reflect the reality.

The two population with the least diversity, *P. t. venus* and *P. paniscus* (bonobos) might be explained by isolation on the west for *P. t. venus* and by sympatric speciation for the *P. paniscus*. As mentioned by Manuel et al. the two species can hybridize in captivity but in the wild it remains unclear.