RecodeData v. 0.1

patrick.meirmans@unil.ch

What is RecodeData?

RecodeData is a simply utility to recode genetic marker data in such a way that all alleles are population specific. This is handy if you want to calculate the standardized genetic differentiation measure that was suggested by Hedrick (2005). RecodeData makes it possible to estimate a standardized measure for other F_{st}-estimators than G_{st} (though *not* R_{st}), plus that it makes it possible to estimate G_{st} with correction for sampling bias (Meirmans, 2006).

How do I calculate the standardized measure?

- Prepare a datafile in Fstat-format, and make sure that is has the extension ".dat".
- Use Fstat (Goudet, 1995), or another program, to calculate the value of your favourite F_{st}-estimator (e.g. Weir & Cockerham's (1984) Theta or Nei's (1987) G_{st}).
- Use RecodeData to recode the data in the Fstat file, the use of this program should be very straightforward. Make sure the input-file is in the same folder as the program. When you start the program you will be asked for the name of the input-file and for a name for the output-file, and that's it.
- Use Fstat again to calculate the same Fst-estimator for the recoded data file. This Fst-value is the maximum value possible given the present within-population variation.
- The standardized genetic differentiation measure can now be found by dividing the original Fst-value by the maximum value. In other words, the value obtained from the original data-file should be divided by the value obtained from the recoded data-file.
- If you want to do a hierarchical analysis, where populations are grouped e.g. based on their geographical position, the recoding has to be performed separately for every hierarchical level. So for calculating Fct (differentiation among groups of populations) the recoding should be done so that all alleles are specific to the groups of populations rather than specific to the populations themselves.

Citation:

Meirmans, P.G. 2006. Using the AMOVA framework to estimate a standardised genetic differentiation measure. Evolution, 60:2399-2402.

Known problems:

- Make sure that the data file does not have a return after the last individual, this may lead to a crash (on a PC) or to an erroneous output-file (Mac).
- If you have any other problems or have questions about the method (like doing a standardized AMOVA) do not hesitate to email me at: patrick.meirmans@unil.ch.

References:

Goudet, J. 1995. FSTAT (Version 1.2): A computer program to calculate F- statistics. Journal of Heredity 86:485-486.

Hedrick, P. W. 2005. A standardized genetic differentiation measure. Evolution 59:1633-1638.

Meirmans, P.G. 2006. Using the AMOVA framework to estimate a standardised genetic differentiation measure. Evolution, 60:2399-2402.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370.