

GlbPSs 1.0 Manual Part 7: Documentation of indpoploc_06.0.pl

5/22/2015 Andreas Hapke,

Institute of Anthropology, Johannes Gutenberg University Mainz, Germany, ahapke2@gmail.com

This program links individual loci and alleles to the catalog of loci and alleles produced by **poploc**. It produces outfiles that are the basis for data analyses across individuals and for the export of genotypes. It further identifies split loci and prints data about them to a file `split_loci.txt`.

Split loci

It can happen that **poploc** assigns different alleles of a given individual locus to different population loci. Population loci that contain alleles from split individual loci are erroneous constructs that must/should be excluded from downstream analyses. The programs **pair_finder** and **depth analyzer** automatically exclude split loci. The program **data_selector** enables you to select/unselect split loci when creating subsets of data for export or analysis.

Usage

```
indpoploc_06.0.pl  
or  
perl indpoploc_06.0.pl
```

Infiles

The infiles for this program are produced by **indloc** and **poploc**:

`individuals.txt`

Outfile of **indloc**: text file containing the IDs of all individuals in the database.

`popall.txt`

Outfile of **poploc**: text file containing a catalog of loci and alleles at the level of the study population.

`*_alleles.txt`

Outfiles of **indloc**, one file per individual, file name starts with the individual ID, e.g. "ind_01_alleles.txt". These files contain data about individual loci and alleles.

Outfiles

This program produces two outfiles for each individual, `ind_01_indpopall.txt` and `ind_01_indpoploc.txt` (where "ind_01" is the individual ID) and one additional file `split_loci.txt`.

ind_01_indpopall.txt

Format: tab-delimited text table with header line, one line per individual allele:

indlocID: individual locus ID

indall_ID: individual allele ID

poplocID: population locus ID

popall_ID: population allele ID

popallvar: characters of this allele at variable positions of the locus or “consensus” when the locus has only one allele

ind_01_indpoploc.txt

Format: tab-delimited text table with header line, one line per individual locus:

indlocID: individual locus ID

poplocID: population locus ID

Each individual locus ID appears only once in the file. A population locus ID can appear several times when several individual loci have been tied into one population locus by **poploc**.

split_loci.txt

Format: tab-delimited text table with header line:

poplocID: population locus ID

ind: individual ID

indlocID: individual locus ID