Identity Analysis of APCL

2018-06-07

This identity analysis was conducted by KAC on the individuals included in APCL sequencing runs 03-17 (majority of individuals captured in 2012-2015, <100 from 2016).

1. Add "pop" line to genepop in text editor. *File "SNP.DP3g95p5maf05HM_seq_17_03.LDpruned.beta.noissues.IDfin.gen" is in the id_support folder.*

2. In cervus, **convert genepop** to cervus file using the "Tools" menu. This takes seconds. *File "SNP.DP3g95p5maf05HM_seq_17_03.LDpruned.beta.noissues.IDfin.csv" is in the id_support folder*

- note the number of loci - **2253**

3. Using the "Analysis" menu, conduct an **allele frequency analysis**.
   This takes seconds.
   File "SNP.DP3g95p5maf05HM_seq_17_03.LDpruned.beta.noissues.IDfin_AF.alf" is in the id_support folder.

   i. use the file just created by converting the genepop.
   ii. make sure "header row"" and "read locus names"" are checked.
   iii. ID is in column 2
   iv. first allele is in column 3
   v. fill in the number of loci listed in the conversion step
   vi. save as input_file_name_AF
   vii. Do not do Hardy Weinberg
   viii. Do not estimate null allele frequency

4. Using the "Analysis" menu, conduct **identity analysis**. MRS notes from the past said this took 30 seconds with 809 loci, KAC notes from 2253 loci in this analysis took 1 minute, 20 seconds. *Files "SNP.DP3g95p5maf05HM_seq_17_03.LDpruned.beta.noissues.IDfin.ident.txt" and "SNP.DP3g95p5maf05HM_seq_17_03.LDpruned.beta.noissues.IDfin.ident.csv" are in the id_support folder.* i.Genotype file and allele frequency info will be automatically populated, should match what you've done for those steps.

   ii. Header should be checked
   iii. ID in column 2
   iv. First allele in column 3
   v. Do not test sexes separately
   vi. Save summary output file using the same naming scheme
   vii. Minimum number of matching loci should be 80% of the total number of loci - **1802**
   viii. Allow fuzzy matching with **225**
   ix. Do not show all comparisons

5. Run the identity_analysis.R script.