Export LongGen File

Author: Alison Figueira, Golden Helix, Inc.

Overview

This script will export genotypic spreadsheet data into Plink's LongGen file format. For more information on this format, please see:
http://pngu.mgh.harvard.edu/~purcell/plink/data.shtml#long

Recommended Directory Location

Save the script to the following directory:
***..\AppData\Local\Golden Helix SVS\UserScripts\Spreadsheet\File\Save As**

Note: The **AppData (or Application Data)** folder is a hidden folder on Windows operating systems and its location varies between various versions. The easiest way to locate this directory on your computer is to open SVS and select the **Tools > Open Folder > UserScripts Folder** menu option and save the script in the **\Spreadsheet\File\Save As** folder. If saved to the proper folder, this script will be accessible from the spreadsheet menu.

Using the Script

1. Open the spreadsheet containing the data to be exported. The data should be by column, as the example below. The Family ID column is optional and if not present, there is be no Family ID column in the resultant text file. In addition, the sample IDs can be in the row labels.

Figure 1: Example Spreadsheet

2. While in the spreadsheet window, select **File > Save As… > Export LongGen File**.

3. In the prompt window, type in the desired file name, select the directory the file should be saved in, and select the sample ID column. If a family ID column is available, it can be chosen here too.
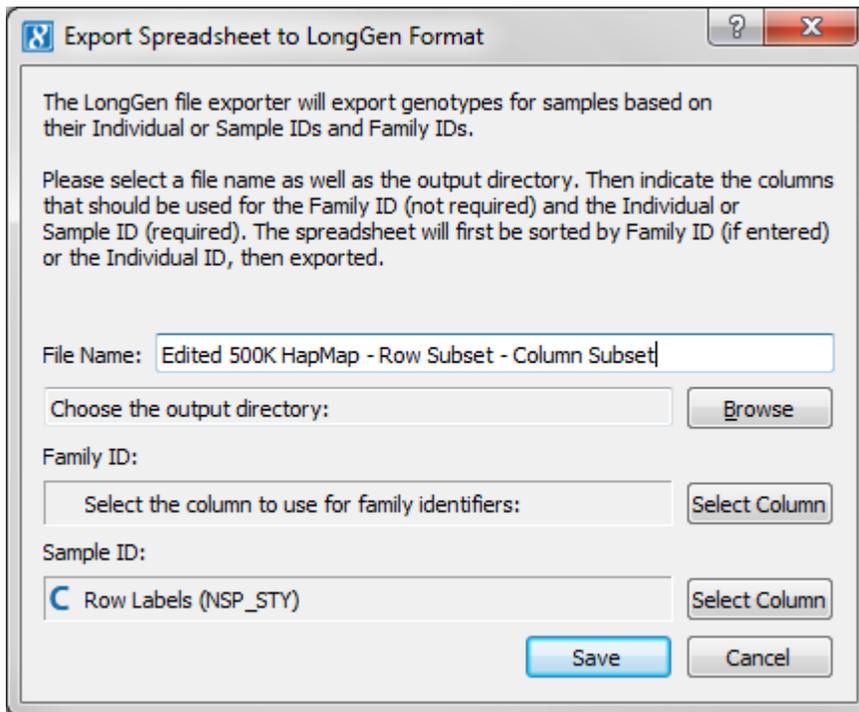

Figure 2: Prompt Dialog

4. Click **Save** to export the spreadsheet.

5. For exporting the spreadsheet will be sorted by family ID first (if available) then sample ID.

6. The exported text file will look similar to the example below. The columns are: Family ID, Sample ID, SNP Name, Allele 1, Allele 2. If no family ID column was chosen this field will be left out.

```
Edited Edited 500K HapMap - Row Subset - Column Subset - Sheet 1.lgen  ⊠
  1    1 1 SNP_A-1909444 B B
  2    1 1 SNP_A-2237149 B B
  3    1 1 SNP_A-4303947 A A
  4    1 1 SNP_A-1886933 B B
  5    1 1 SNP_A-2236359 B B
  6    1 1 SNP_A-2205441 A A
  7    1 1 SNP_A-2116190 A A
  8    1 1 SNP_A-4291020 A A
  9    1 1 SNP_A-1902458 A A
 10    1 1 SNP_A-2131660 A B
 11    1 2 SNP_A-1909444 B B
 12    1 2 SNP_A-2237149 B B
 13    1 2 SNP_A-4303947 A A
 14    1 2 SNP_A-1886933 B B
 15    1 2 SNP_A-2236359 B B
 16    1 2 SNP_A-2205441 A A
 17    1 2 SNP_A-2116190 A A
 18    1 2 SNP_A-4291020 A A
 19    1 2 SNP_A-1902458 A A
 20    1 2 SNP_A-2131660 A B
 21    1 3 SNP_A-1909444 B B
 22    1 3 SNP_A-2237149 B B
```

Figure 3: Example Text File