

Export Impute2 Genotype Probabilities

Author: Greta Peterson, Golden Helix, Inc.

Overview

From a spreadsheet containing marker-mapped genotypic columns, this script saves the spreadsheet in a series of Impute2 chr*.gen files with the corresponding chr*.sample files to the specified directory. For more information about the Impute2 format, see:

http://www.stats.ox.ac.uk/~marchini/software/gwas/file_format.html

Note: One sample file is created per chromosome to capture the missing data rate for each sample per chromosome. This is a required column for programs that use this data format for analysis and imputation.

Recommended Directory Location

Save the script to the following directory:

*..\Application Data\Golden Helix SVS\UserScripts\Spreadsheet\File\

Note: The **Application Data** folder is a hidden folder on Windows operating systems and its location varies between versions. The easiest way to locate this directory on your computer is to open SVS and select **Tools >Open Folder > User Scripts Folder**. If saved to the proper folder, this script will be accessible from the spreadsheet **File** menu.

Using the Script

1. Open a spreadsheet that contains genotype information in columns.
2. Select **File >Export Impute2 Genotype Probabilities**.
3. Choose the SNP ID and RS ID fields from the marker map (including column headers). Click **OK**.

Upon completion, a separate *.gen and *.sample file is created for each chromosome with data in the folder you indicated.