Genotype Statistics Summary

Author: Greta Linse Peterson, Golden Helix, Inc.

Overview

This script takes a spreadsheet that contains a case/control dependent variable and SNPs and runs all of the genotype association tests as well as tests for a heterozygous advantage model (Dd vs DD, dd) and a homozygous comparison model (DD vs dd). Also calculates Chi Squared Scores, Correlation/Trend test scores and completes count tables.

Recommended Directory Location

Save the script to the following directory:
*..**\Application Data\Golden Helix SVS\UserScripts\Spreadsheet\Analysis**

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

Using the Script

1. From a spreadsheet that contains SNP data and a case/control dependent phenotype. Set the case/control column as dependent and run the script by going to **Analysis > Genotype Statistics Summary**.

2. Select the desired genotype models and statistics as well as output options in the prompt dialog. Also note, intermediate spreadsheets can either be deleted for a clean project or not.

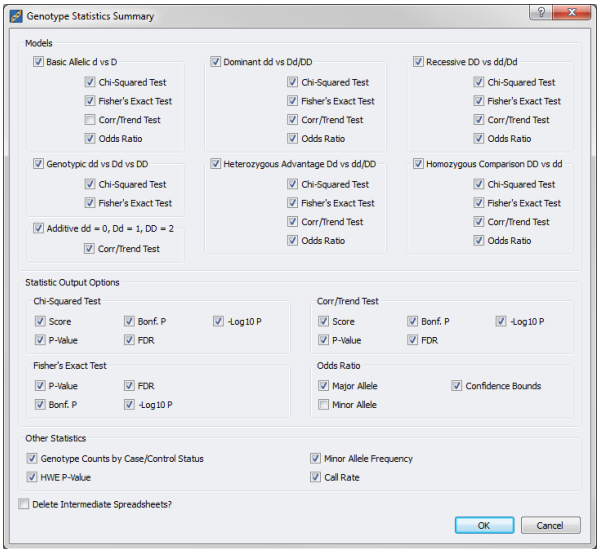

Figure 1: Genotype Statistics Summary Prompt Dialog

3. When the desired parameters are set, click **OK** and the appropriate models will be run and compiled together into one spreadsheet.

The resulting spreadsheet will have all of the active markers as row labels and the genotype statistics will be displayed in this order: model independent statistics, basic alleic (D vs d), genotypic (dd vs Dd vs DD), additive (dd = 0, Dd =1 , DD = 2), dominant (DD, Dd vs dd), heterozygous advantage (Dd vs DD, dd), recessive (DD vs Dd, dd) and homozygous comparision (DD vs dd). The start of the next group can be identified by the column header for the Chi Squared Scores columns.

If there are more output columns than desired, these can be removed by inactivating the column and creating a column subset spreadsheet. Columns can also be reordered using the Spreadsheet Editor (**Edit > Edit This Spreadsheet)** and clicking on the column to move and selecting one of the move operations.