Build Sample Collated Spreadsheet

Author: Autumn Laughbaum, Golden Helix, Inc.

Overview

This tool transposes and collates several spreadsheets together. The collated spreadsheet contains a row for each intersecting column in the original spreadsheets and several columns for each original row over all spreadsheets.

This tool was designed for use on spreadsheet imported using the VCF Importer. For example, if a user imports genotype (GT) and read depth (RD) spreadsheets, the collated spreadsheet would contain a row for each marker and two columns for each sample, e.g. Sample1-GT, Sample1-RD, Sample2-GT, etc.

Recommended Directory Location

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

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 > UserScripts Folder**. If saved to the proper folder, this script will be accessible from the project navigator's **Tools** menu.

Using the Script

1. Open a project that contains several spreadsheets with intersecting column headers and row labels.
2. From the project navigator, choose **Tools> Build Sample Collated Spreadsheet**
3. Choose the appropriate spreadsheets and optionally choose a base dataset name that will precede ' – Sample Collated Spreadsheet'.
    Note: In the Advanced Options tab, the transpose memory limit may be set to something other than the default (based on OS).
4. Click **Next>**.
5. For each previously selected spreadsheet, choose the new column header suffix (GT and RD in previous example) and the column type to transpose.
    Note: The default column type will be automatically detected based on that type representing the largest number of columns in that spreadsheet.
6. Click **OK**.

The output spreadsheet contains the original columns as rows and a separate column for each sample-spreadsheet.