## **Convert Binary and Integer Values to Genotypes**

**Author:** Greta Peterson, Golden Helix, Inc.

## Overview

This script recodes binary and integer genotypes to the standard genotype format of A\_A, A\_B, and B\_B. Prompts for value of A\_A, A\_B, and B\_B. All other numbers are encoded as missing. Thus if there is multi-allelic data in the spreadsheet, all numbers other than those specified will be encoded as "?".

\*Note: If the spreadsheet is marker-mapped, assumes that only marker-mapped binary and/or integer columns are to be recoded. If the spreadsheet is not marker-mapped, assumes that any integer (and binary, if selected) columns found in the spreadsheet should be recoded. In any case, all other column types will be added to the new spreadsheet as is.

If the spreadsheet is marker-mapped, the marker map will be reapplied to the new spreadsheet.

## **Recommended Directory Location**

Save the script to the following directory:

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

**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 select **Tools >Open Folder > UserScripts Folder**. If saved to the proper folder, this script will be accessible from the spreadsheet **Edit** menu.

## **Using the Script**

1. Open a spreadsheet containing binary or integer values representing genotypes, such as genotypes numerically encoded in an additive model, see **Figure 1**.

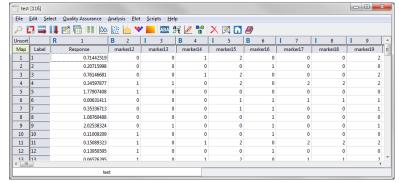


Figure 1: Spreadsheet with genotypes numerically encoded with binary and integer values

- 2. Select Edit > Convert Binary and Integer Values to Genotypes.
- 3. Enter in the values for the corresponding genotypes, see Figure 2.

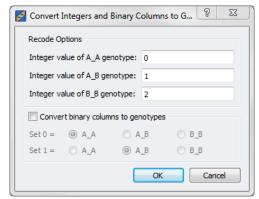


Figure 2: Genotypes to be assigned to the specified integers and binary values

The script will then create a new spreadsheet with the (mapped, if applicable) binary and integer columns from the original spreadsheet converted to genotypic containing only A\_A, A\_B, B\_B and ?\_? genotypes. Any integer values from these columns that do not correspond to the specified values are replaced with missing. Missing values remain missing. See **Figure 3**.

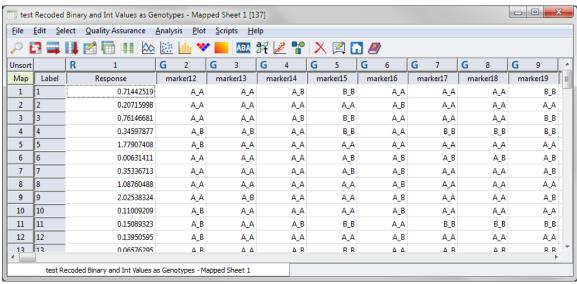


Figure 3: Recoded integers as genotypes