

## 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**.

test Recoded Binary and Int Values as Genotypes - Mapped Sheet 1 [137]										
File Edit Select Quality Assurance Analysis Plot Scripts Help										
Unsort		R 1	G 2	G 3	G 4	G 5	G 6	G 7	G 8	G 9
Map	Label	Response	marker12	marker13	marker14	marker15	marker16	marker17	marker18	marker19
1	1	0.71442519	A_A	A_A	A_B	B_B	A_A	A_A	A_A	B_B
2	2	0.20715998	A_A	A_A	A_A	A_A	A_B	A_A	A_A	A_A
3	3	0.76146681	A_A	A_A	A_B	B_B	A_A	A_A	A_A	B_B
4	4	0.34597877	A_B	A_B	A_A	B_B	A_A	B_B	B_B	B_B
5	5	1.77907408	A_B	A_A	A_A	A_A	A_A	A_A	A_A	A_A
6	6	0.00631411	A_A	A_A	A_A	A_B	A_B	A_B	A_B	A_B
7	7	0.35336713	A_A	A_A	A_A	A_B	A_B	A_A	A_A	A_B
8	8	1.08760488	A_A	A_A	A_A	A_A	A_B	A_A	A_A	A_A
9	9	2.02538324	A_A	A_B	A_A	A_A	A_B	A_A	A_A	A_B
10	10	0.11009209	A_B	A_A	A_A	A_A	A_B	A_A	A_A	A_A
11	11	0.15089323	A_B	A_A	A_B	B_B	A_A	B_B	B_B	B_B
12	12	0.13950595	A_B	A_A	A_A	A_A	A_B	A_A	A_A	A_A
13	13	0.06576295	A_R	A_A	A_R	R_R	A_A	A_R	A_R	R_R

Figure 3: Recoded integers as genotypes