

Allele Frequencies

Author: Josh Forsythe, Golden Helix, Inc.

Overview

This script calculates the number of alleles and their respective frequencies for each SNP. The result is a 2x1, 2x2 or 3x3 table for each SNP depending on the selected output. Inactive columns and inactive rows will not be included in the allele counts and frequency calculations. Only the rows and columns that are active are included in the calculations.

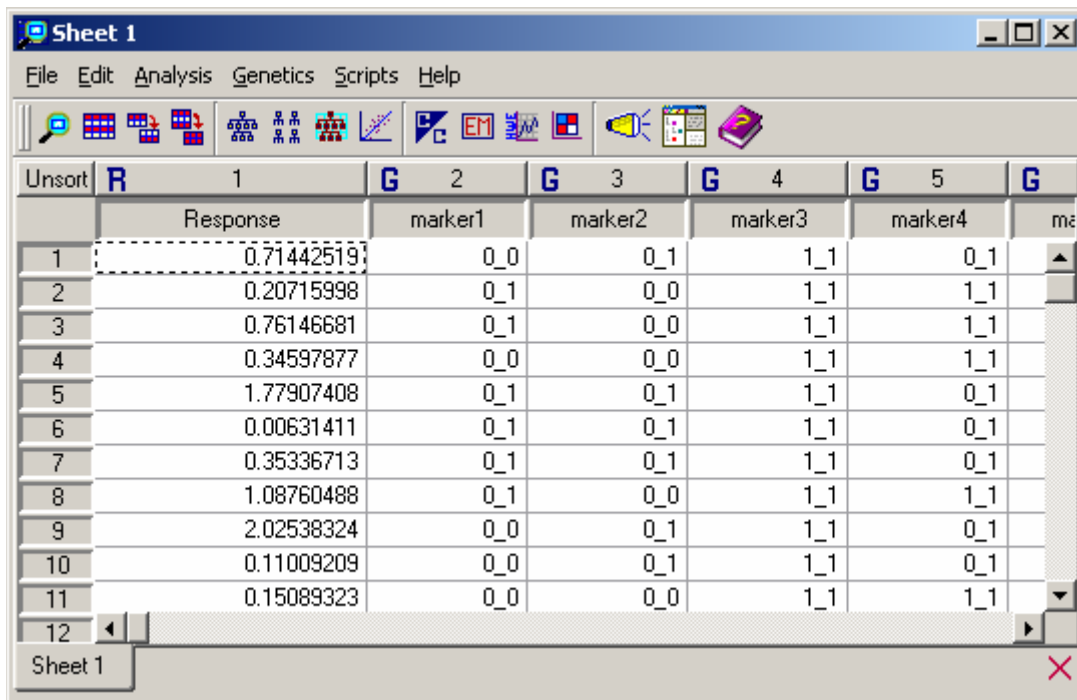
Recommended Directory Location

A good directory location for saving this script would be:
C:/HelixTree/scriptsHT/user/Spreadsheet/Genetics/.

When saved to this folder, the script will be accessible from the spreadsheet Genetics menu.

Using the Script

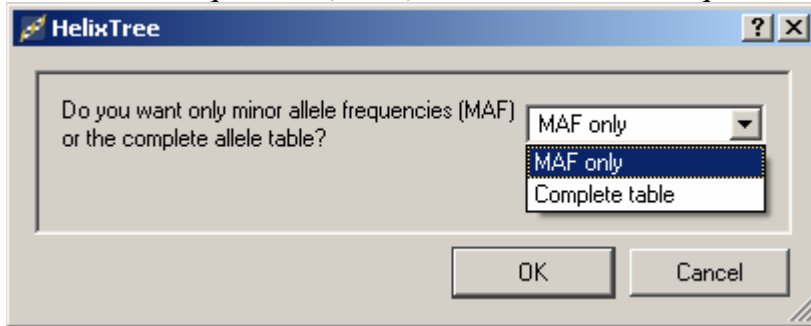
Open a spreadsheet that contains genotype information. The spreadsheet may also contain non-genetic columns. Inactivate any SNPs or samples that are not to be included in the calculations.



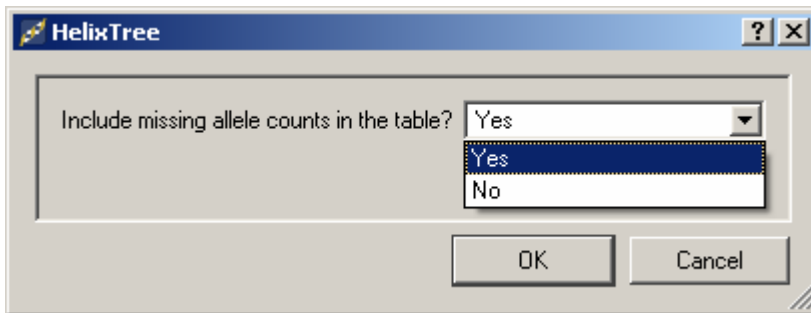
	R	G	G	G	G	G
	1	2	3	4	5	
	Response	marker1	marker2	marker3	marker4	ma
1	0.71442519	0_0	0_1	1_1	0_1	
2	0.20715998	0_1	0_0	1_1	1_1	
3	0.76146681	0_1	0_0	1_1	1_1	
4	0.34597877	0_0	0_0	1_1	1_1	
5	1.77907408	0_1	0_1	1_1	0_1	
6	0.00631411	0_1	0_1	1_1	0_1	
7	0.35336713	0_1	0_1	1_1	0_1	
8	1.08760488	0_1	0_0	1_1	1_1	
9	2.02538324	0_0	0_1	1_1	0_1	
10	0.11009209	0_0	0_1	1_1	0_1	
11	0.15089323	0_0	0_0	1_1	1_1	
12						

Select the Allele Frequencies script from the spreadsheet menu location where it was stored. Or, from the project viewer, select **Tools->Run Script** to open a dialog to browse for the script location.

When the script is run, a dialog will appear for you to select to calculate either only the minor allele frequencies (**MAF**) or all of the allele frequencies.



If you select to calculate the **Complete table**, a second dialog will appear for you to select to include missing alleles and their frequencies in the table.



When the calculations are complete, a new spreadsheet containing the allele counts and the allele frequencies will open.

Unsort	C	I	R	
1	2	3		
SNP Name	Allele	Count	Minor Allele Frequency	
1	marker1	1	54	0.27
2		0	146	0.73
3		?	0	0
4	marker2	1	58	0.29
5		0	142	0.71
6		?	0	0
7	marker3	0	2	0.01
8		1	198	0.99
9		?	0	0
10	marker4	0	59	0.295
11				

In this spreadsheet, the SNPs are arranged row-wise. If you selected to calculate the complete table and to include missing alleles, there will be 3 rows per SNP. If you selected to only calculate the minor allele frequencies or selected to calculate the complete table but to not include missings, there will be only 1 or 2 rows per SNP.

There are 3 columns in the output spreadsheet. The first column contains the allele names for each SNP. The second and third columns contain the counts and frequencies for those alleles.