Take Home Exam 3/22/17

For this exam, you will need to use the file Take\_Home\_Exam\_Table located in /home/patrick/. Once you have this located this file, your task will be to calculate the allele frequency for each of the 100 SNPs in the table. The table format is as follows:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| CHR# | LOCATION | SNP ID | REF ALLELE | ALTALLELE | Ind 11st Allele | Ind 12nd Allele | ……. | Ind 2071st Allele | Ind 2072nd Allele |

“1” indicates the alternative allele, while “0” indicates the reference allele. Your goal is to calculate the allele frequency $num\_zeros/($num\_zeros + $num\_ones). For example:

2 293842937 rs373822 A G 1 0 1 1 0 0

Would equal an allele frequency of 0.5 (50%) because 3/6 = 0.5. You can use regular expressions or arrays to solve this problem. If you feeling ambitious, try both!!

Your output should look as follows:

Rs23974293 0.23

Rs23847345 0.05

Rs112823221 0.56

Rs23499000 0.77

………

If you have any questions, please email your TA.

Please email your TA your completed script and output by 3/29.