

1000 Genomes Exercises

Mating	Probability	Offspring Genotype		
		AA	Aa	aa
AA x AA	D^2	1	0	0
AA x Aa	2DH	$\frac{1}{2}$	$\frac{1}{2}$	0
AA x aa	2DR	0	1	0
Aa x Aa	H^2	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
Aa x aa	2HR	0	$\frac{1}{2}$	$\frac{1}{2}$
aa x aa	R^2	0	0	1
Total for Gen 1		D'	H'	R'

From this table, what are the probabilities (in p and q) for D', H' and R'?

Hint: multiply the probability by the fraction of the offspring like that

Given the following matings, each producing 2 offspring, what is the p and q for gen 1? Flip a coin if you have to make a choice between two alleles.

AA x aa; Aa x AA; aa x Aa; AA x aa; Aa x Aa

Once you have your Gen 1, choose random matings, and do it again. What is the p and q for gen 2?

Why does the higher level of variants per genome support the out of Africa hypothesis of human evolution and migration?

In the North American Ancestry paper, PC1 divides the population in what geographic dimension? PC2? Why was this the first look at the genomics clusters they used?

Where were these data acquired? How does this differ from the 1000 Genomes project? What do you think might be implications of this different type of data acquisition?