



Average Allele Frequency Calculation

Created: July 7, 2005; Updated: February 18, 2014.

How does dbSNP calculate average allele frequency?

$$\text{rs allele frequency} = \frac{\text{sum of chromosome counts over all member ss for an allele}}{\text{total chromosome counts over all member ss}}$$

Average Allele frequency is calculated only when the following two conditions are met:

The rs number has at least one member ss with frequency data that include population frequency or individual genotype.

The alleles with frequency data are consistent with the assay alleles for the ss number.

We assume that populations of different ss numbers do not overlap because we don't have detailed sample sources for most populations. When we have a sample source that contains individual genotype data and pedigree information, we only include non-founders and only unique samples if one sample was used in a two-member ss within an rs.

dbSNP has a total of 754,822 rs numbers that meet the first condition at this time, but we only have the average allele frequency for 751,363 rs numbers. This means that approximately 0.004% of the rs numbers (3459 rs numbers) have problems that prevent us from computing their average allele frequency data. If you are interested in the allele frequency for the rs numbers that fall within this small group, you could look at each individual ss number within the rs cluster that interests you. Another case where we have frequency data for a ss number but don't compute the average allele frequency is when the ss number has not yet been clustered.