

Common Conventions

After this module you should be able to explain the following terms and ideas.

KEY TERMS

Reference genome, Position, Reference Allele, Alternate Allele, Derived Allele, Ancestral Allele, Major Allele, Minor Allele, Frequency, Haplotype Matrix, Genotype Matrix, VCF, rsID, vcftools and bcftools

CONCEPT QUESTIONS

What is a reference genome? How is it used? How do you know which one to use?

How is variant position defined?

What does it mean for an allele to be the reference or alternate allele? How are these designations different from being the ancestral or derived allele? What about a major or minor allele?

What is an rsID? Where can you find them?

What is allele frequency?

What does VCF stand for? What kind of information does it hold?

What is a BCF? How is it different from a VCF?

What are vcftools and bcftools used for?

What is a haplotype matrix? How is it different from a genotype matrix?

How do you convert a haplotype matrix into a genotype matrix?

How do you convert a genotype matrix into a haplotype matrix?