Variant Call Format

VCF files

A common format for NGS data is the VCF file. This tool is designed to work with version 4.0 or 4.1 VCF files such as those found in the 1000 genomes project. (http://www.1000genomes.org/data#DataAccess)

VCF File Assumptions:

The header line is as follows and lists sample IDs.

#CHROM   POS   ID   REF   ALT   QUAL FILTER   INFO
        FORMAT   NA00001   NA00002   NA00003

In the example above, the sample IDs are NA00001, NA00002 and NA00003.