

File types

| .gz | Appended to files that are compressed |
|--------|---------------------------------------------------------------------------------------------------------------------|
| .fasta | Simple format for storing sequence information. |
| .fastq | Stores sequence and quality information |
| .gff | General Feature Format: a list of genes and other genomic features, and their location in a particular genome. |
| .sam | Sequence Alignment/Map format. Links sequences (as from reads) to a position in a reference genome. |
| .bam | The compressed version of a .sam file. |
| .vcf | Variant Call Format; stores information about variation between sequences, as between reads and a reference genome. |