SAMTOOLS

Latest version on Fourierseq: 0.1.18 installed 9/26/11 by SPHS

Useful samtools utilies:

1. samtools idxstats : This tool will provide statistics about how many reads have aligned to each sequence/chromosome in the reference genome. The input bam file must be sorted and indexed.

```
samtools idxstats <in.bam>
```

2. samtools flagstat : Simple stats about how many reads mapped to the reference, how many reads were paired properly etc. The input bam file must be sorted and indexed.

```
samtools flagstat <in.bam>
```

Example:

1. *samtools mpileup -Euf reference.fna aln1.bam aln2.bam | bcftools view -bvcg - > var.raw.bcf*

where reference.fna : reference, in fasta format

aln1.bam, aln2.bam : BAM files containing alignment results. You can use 1 or more alignment flies at a time. Note that as of late 2011, the new BAQ filter seems to aggressively remove SNPs unless you "extend" it with the "-E" option.

2. *bcftools view var.raw.bcf | vcfutils.pl varFilter -D10 > var.filtered.vcf*

BCFtools does the actual calling of SNPS and the SNP information is stored in var.filtered.vcf. -D option is used to filter by depth of coverage at the SNP location.

Information about VCF file and other filter options at : http://samtools.sourceforge.net/mpileup.shtml

OLD VERSION: Commands to use samtools with a bam file, input.bam,

1. Use samtools pileup to call SNPs

samtools pileup -vcf reference.fna input.bam > out.pileup 2>out.log &

where reference.fna : reference file, in fasta format

input.bam : BAM file containing alignment results

2. Filter the results further by snp quality:

samtools.pl out.pileup||awk '$6>=20' > out.final.pileup