**From BAMs to VCF**

*Script name:* rnaseq\_bam\_to\_vcf.pl

What it does: it takes one or more bam files and call SNPs with HaplotypeCaller or UnifiedGenotyper from GATK

*Synopsis:*

***perl rnaseq\_bam\_to\_vcf.pl -r reference\_fasta -p output\_prefix -c HC|UG -x extension\_file\_to\_treat***

/!\ BAM files have to be present in CWD.

*Parameters:*

-p (string): VCF output prefix

-c: HC for HaplotypeCaller or UG for UnifiedGenotyper from GATK

-x (string): file extension (bam)

*All the steps in detail*

* Create a VCF file with HaplotypeCaller or UnifiedGenotyper (deprecated) from GATK

<https://software.broadinstitute.org/gatk/documentation/tooldocs/3.8-0/org_broadinstitute_gatk_tools_walkers_haplotypecaller_HaplotypeCaller.php>

* Prefiltering with VariantFiltration and SelectVariants from GATK. Annotating variants with:
  + SnpCluster: remove clusters of SNPs (more 2 SNPs in 10bp).
  + QualitybyDepth: remove SNPs with a QD < 1.5 (<https://software.broadinstitute.org/gatk/documentation/tooldocs/3.8-0/org_broadinstitute_gatk_tools_walkers_annotator_QualByDepth.php>)