There is a nice introductory wiki to analysis of WES data, a bit outdated, but it is good to understand the basics.

The broad institute has developed GATK, a series of bioinformatic tools to analyze NGS data from the raw files produced by the sequencing service to the files ready for statistical analysis.

As GATK indicates in his Best Practices for Variant Discovery workflow*, WES bioinformatic workflow is divided in three main sections that are meant to be performed sequentially:

- Sequence Data pre-process: from raw DNAseq sequence reads (FASTQ files) to analysis-ready reads (BAM files)
- Variant Discovery: from reads (BAM files) to variants (VCF files)
- Preliminary analysis: callset QC, refinement and preliminary analyses
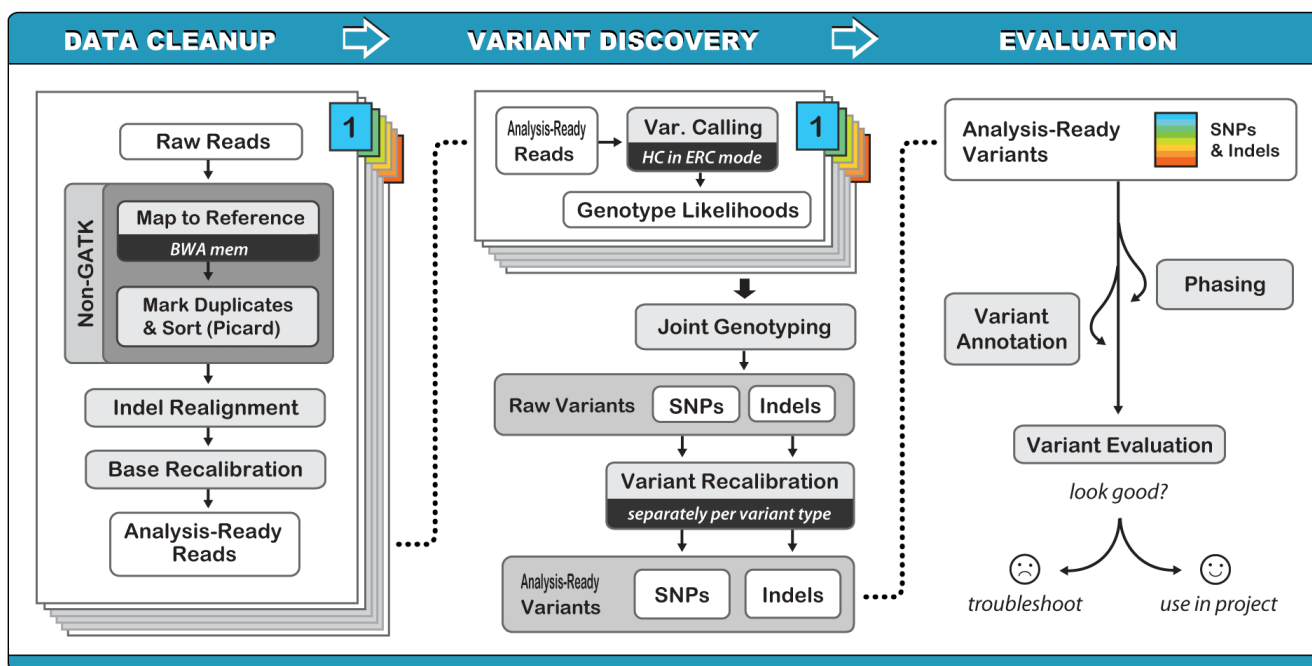
* This page displays better with Chrome than with Firefox

From:
https://cortafuegos.fundacioace.com/wiki/ - **Detritus Wiki**

Permanent link:
https://cortafuegos.fundacioace.com/wiki/doku.php?id=genetica:bioinf_process

Last update: **2020/08/04 10:58**