Filtering variants in a clinical setting

MSc in Genomic Medicine

Lucy Crooks

26/1/2016

SDGS Gene panels

- 4 Panels of 300 kb (CTD, Hereditary cancers, NGD, IEM)
- 70-100 genes
- 5 million paired-end reads up to 1000X coverage
- 96 samples per run
- Confidence that no variants are missed

SDGS variant analysis

Alignment

BWA Picard GATK

Variant calling

GATK HaplotypeCaller

Base quality≥10
Mapping quality≥20

Quality filtering

Remove variants with QUAL<30

Identify key variant

VCFtools (filter to target)

Remove polymorphisms (in-house list)

SNPEff

ANNOVAR

Sanger sequence gaps with coverage < 30X

Pipeline Validation

- 100% detection rate for 300 variants found by Sanger sequencing
  - 31 patients
  - 28 genes
  - 3 panels