NKI: mutation - Illumina Agilent V7, Targeted-Capture

Molecular Methods Description:

Genomescan prepared the samples according to the procedure for Hybridization Capture using an Agilent SureSelect custom 0.5-2.9Mb kit for the panelseq and the Agilent SureSelectXT Human All Exon V7 kit for the WES samples. The prepared libraries were sequenced with Illumina sequencing technology and prepared according to manufacturer■s protocols.

Analysis Description:

The reads were trimmed using Cutadapt61 to remove any remaining adapter sequences, filtering reads shorter than 60 bp after trimming to ensure good mappability. The trimmed reads were aligned to the human (GRCh38) and mouse (GRCm38) reference genome using BWA. The human alignment was processed for duplicate marking, indel realignment, and base recalibration using Picard Tools and GATK, as recommended by GATK best practices, and filtered to remove contaminating mouse reads using AstraZeneca■s tool disambig- uate.63 QC statistics from Fastqc.64 FastQC: a quality control tool for high throughput sequence data. Available online at: http://www. bioinformatics.babraham.ac.uk/projects/fastqc) and the above-mentioned tools were collected and summarized using Multiqc.66 Mutect2 was used for SNP calling followed by the LearnReadOrientationModel and FilterMutectCalls commands. SNPs that had a TLOD of <10, a coverage of less than 15, an alternative frequency of less than 0.2, had a different function than exonic or splicing, were classified as synonymous-SNV, and/or had a population frequency of more than 0.01 in one of the following databases down- loaded with ANNOVAR (1) (1000g, Kaviar, hrcr1, gnomad_genome, gnomad_exome, esp6500siv2, exac_03, gme) were excluded.