

UOC-BC: copy number alteration - Illumina Nextera Rapid Capture Exome, WGS

Analysis Description:

Data output is from the Caldas Laboratory at the University of Cambridge . Full details can be found in the paper Bruna et al., Cell 167, 260 - 274 (2016) .

CNA calls were performed using shallow Whole Genome Sequencing. 50 bp single-read whole-genome shallow sequencing was performed in parallel with the exome sequencing to provide a clean and accurate estimate of copy number. Alignment was performed using bwa with our custom pipeline to remove mouse contamination. Bam files were merged, sorted and indexed using samtools. Duplicates were marked using Picard tools. The data were analyzed using the Bioconductor package QDNaseq. This method divides the genome in regions of 100Kb and counts all the reads within those bins. Those reads are then corrected for mappability and GC content and segmented using DNACopy. Some additional filtering was applied to account for regions not properly mapped.

The segmented means of the tumors were corrected for normal contamination (as described in the exome pipeline) and copy numbers (HOMD, Homozygous deletions, HETD, Heterozygous deletions, NEUT, neutral copy number, GAIN, single copy gains and AMP, high-level amplifications), were called based on thresholds on the segmented mean log2-ratio (-1, -0.4, 0.25, 0.75).