Supplementary Methods

Whole Exome Sequencing

A minimum of 500 ng of DNA was used to generate DNA libraries using Agilent's SureSelect protocol as per manufacturer's instructions. Neo42 had very low quantities of DNA extracted and libraries were therefore generated using the Nextera DNA library protocol (Illumina Inc.). We sequenced DNA from matched lymphocytes to use as a control for germline variants. The sequencing was subsequently performed on Illumina HiSeq2000 platforms with 100 base paired-end reads. High-quality trimmed reads were aligned to the human reference genome (UCSC hg19) using the Burrows-Wheeler Alignment tool (BWA 0.7.12). Insertions/deletions (indels) were re-aligned using Genome Analysis Tool Kit (GATK). PCR duplicates were marked with Picard. Single nucleotide variants (SNVs) and indels were called by means of the SAMtools software. Variants with minimum of 10 reads were considered. At the end, variants were annotated with ANNOVAR and custom in-house scripts.

Final allele frequencies (AFs) were calculated by dividing the generated AF by the tumor cellularity percentage for each sample.