# Supplementary Methods

Sample Processing. Surgical specimens were preserved in RNALater prior to freezing, matched blood samples were collected in parallel at the time of surgery. All cases were from treatment naive patients. Tissue samples were macro-dissected to remove excess non-tumour tissue, and all tissues were confirmed to be ≥80% tumour content by a consultant pathologist. Tumour and matched whole blood DNA (normal germline) was extracted using DNeasy kit (Qiagen) according to standard protocols, as previously described([Feber et al., 2015](#_ENREF_12)). DNA quantification was carried by Qubit (Invitrogen), and DNA integrity assessed by agarose gel electrophoresis.

Whole-exome sequencing. PCR duplicates were removed and coverage metrics were calculated using Picard-tools (v.1.48, <http://picard.sourceforge.net/>). Reads which were unmapped or as a result of PCR duplication were excluded from further analysis.

A mean coverage of 60X was achieved across each of 27 paired samples, with 70% of target bases covered at least 30X. Validation samples were subjected to targeted sequencing

candidate genes (*CSN1/GPS1, FAT1* and *TP53*) using Fluidigm custom amplification. A mean coverage of 156x was achieved across the validation cohort.

Epigenetic alterations. DNA methylation data were derived from the Infinium 450K Human methylation array and analyzed are previously described[[7](#_ENREF_7)]. 500ng of genomic DNA was bisulphite converted, using the Zymo Gold Bisulphite conversion kit (Zymo), and hybridized to the Infinium 450K HumanMethylation array (Illumina), and processed in accordance with the manufacturer's recommendations. DNA R statistical software (version 2.14.0 ) was used for the subsequent data analysis. The ChAMP pipeline (version 1.6) was used to extract and analyse data from iDat files, samples were normalised using BMIQ ([Morris et al., 2014](#_ENREF_31)). Raw **β** values (methylation value) were subjected to a stringent quality-control analysis as follows: samples showing reduced coverage were removed and only probes with detection levels above background across all samples were retained (detection P< 0.01). A linear regression model was used were used to define MVPs (Methylation Variable Positions) from normalized **β** values (methylation value), DMRs (differentially methylated regions) were called using the Probe Lasso algorithm with default parameters with the exception of applying a minimum DMR size of 100bp ([Butcher and Beck, 2015](#_ENREF_7)). As a result, all DMRs identified have a minimum of 3 significant probes, are at least 1Kb from a neighboring DMR, and have a minimum size of 100bp. Maximum DMR size is effectively unbounded but is dependent the genomic separation between contiguous CpG probes, which itself is contingent on the local underlying genomic and epigenomic features with larger DMRs more likely to occur in probe-poor regions ([Butcher and Beck, 2015](#_ENREF_7)),([Morris et al., 2014](#_ENREF_31))).

Detection of copy number alterations. Data was derived from the Infinium 450K HumanMethylation array using the ChAMP or conumee algorithm[35[\_ENREF\_35](#_ENREF_35), [36](#_ENREF_36)] and GISTIC, to confirm the significance of alterations.