**Supplementary Materials**

**Materials and Methods**

**Genomic DNA preparation and whole-genome or whole-exome sequencing**

Briefly, the genomic DNA sample was fragmented to a size of ~350 bp by a Covaris sonication system. Then, DNA fragments were end-polished, A-tailed, and ligated with the full-length adapter for Illumina sequencing, followed by further PCR amplification. After PCR products were purified (AMPure XP system), libraries were analyzed for size distribution by Agilent 2100 Bioanalyzer and quantified by real-time PCR (3 nmol/L). The clustering of the index-coded samples was performed on a cBot Cluster Generation System using a HiSeq X PE Cluster Kit v2.5 (Illumina) according to the manufacturer’s instructions. After cluster generation, the DNA libraries were sequenced on the Illumina HiSeq X platform, and 150 bp paired-end reads were generated.

**Reads mapping and somatic genetic alteration detection**

For identification of the somatic SNVs in the tumor, only the highest quality data were used in order to eliminate false positives. Therefore only reads that passed the following tests were retained: (a) Mapping Quality score > 0; (b) Base quality score ≥ 5; (c) If there was an overlapping read pair, and both reads agreed, the read with the highest quality score was retained otherwise both were discarded; (d) Sum of the quality scores of the mismatches ≤ 100; (e) <30% of bases were soft-clipped; (f) Reads that were mapped by “mate rescue” of BWA (BAM XT tag ≠“M”)

All somatic indels were further filtered by the following rules: (i) All calls with a normal sample depth >3 times the chromosomal mean (to remove e.g. pericentromeric regions) (ii) Somatic indels with a reference repeat count >8 (i.e. the indel is an expansion/contraction of a homopolymer longer than 8 bases, a dinucleotide repeat longer than 16 bases, etc.) (iii) Somatic indels where >30% of basecalls were filtered out in a window extending 50 bases to either side of the indel’s call position (iv) Somatic indels overlapping ‘interrupted homopolymers’ of length >15. An interrupted homopolymer was the longest homopolymer intersecting or adjacent to the called indel when a single non-homopolymer was is allowed.

Filtering parameters for FFPE samples

Additional filters were applied to exclude artifactual mutations introduced by formalin-fixed and paraffin-embedded (FFPE) specimens. In brief, duplicates and soft clipped reads removed data were analyzed in MuTect with these parameters (align quality: 40; strand bias: 0.05; keep the mutation site with highest align quality if more than one mutation sites were examined within 11 bp; to keep the mutation sites supported by at least three different reads). Furthermore, we filtered out variants in a variant frequency peak with <5% frequency and total coverage under 30 × , and we filtered out single strand bias based on a read pair orientation of larger than 20:1.

**PCR and Sanger sequencing**

To validate somatic SNVs and InDels identified from the WGS and WES data, we used PCR to amplify genomic DNA spanning mutation sites with specific primers. Then, the PCR products were sequenced directly or cloned into TA vectors. At least 50 TA vector clones were sequenced because mutations at lower variant allele frequencies (VAFs) were difficult to detect by direct Sanger sequencing.

**Immunohistochemistry**

First, the tissue slides were baked in a 60°C oven for 2 hr, subsequently deparaffinized by dimethylbenzene twice for 10 minutes each time and rehydrated by graded ethanol. High pressure pretreatment in citrate buffer (pH 9.0) was used for antigen retrieval for NF-κB P65, NOTCH1, RPL22, and BAP1 staining, while high pressure pretreatment in citrate buffer (pH 8.0) was used for antigen retrieval for FAM135B and NCKAP5L staining. H2O2 (0.3%) was used to block endogenous peroxidase activity for 30 minutes. The samples were then washed with PBST and incubated overnight at 4°C with primary antibodies against NF-κB P65 (diluted 1:800), NOTCH1 (diluted 1:200), RPL22 (diluted 1:200), BAP1 (diluted 1:200), FAM135B (diluted 1:50), and NCKAP5L (diluted 1:100). Next, the slides were incubated with biotinylated anti-goat antibodies for 30 minutes at room temperature and then with secondary antibodies conjugated with horseradish peroxidase for 30 minutes at 37°C. The slides were then stained with 3,3ˈ-diaminobenzidine tetrahydrochloride (DAB) and counterstained with hematoxylin.

**Supplementary Figure 1. GISTIC analysis of rNPC and pNPC samples.** G-scores of genomic amplifications (AMP) and deletions (DEL) of rNPC and pNPC are plotted as curves. Genes located in the most significant regions are represented in rNPC. The genes marked with a red color are shared by rNPC and pNPC.

**Supplementary figure 2.tif**

**Supplementary Figure 2. Recurrent somatic mutations in rNPC and pNPC.** Mapping of the mutation sites of these somatic mutations from the pNPC and rNPC cohorts. Functional domains of the altered proteins are based on the UniProt database.

**Mutation_show_pNPC_rNPC.tif**

**Supplementary Figure 3. Western blot (WB) assay with the indicated antibodies in 293T cell line (A, B, C) and SUNE2 cell line (D, E, F)**

The wild-type *TRAF3*, *NFKBIA*, and *CYLD* led to lower nuclear NF-κB expressions in SUNE2 and 293T cell lines, whereas most mutant *TRAF3*, *NFKBIA*, and *CYLD* resulted in significantly higher nuclear NF-κB expression in SUNE2 and 293T cell lines.

****

**Supplementary Figure 4. Immunohistochemical (IHC) staining to examine the expression of specific SMGs in 148 rNPC samples and 122 pNPC samples** revealed similar expression levels in these specific SMGs between rNPC and pNPC (A, C, D), except for *NOTCH1* and *NCKAP5L* (B, E).

Supplementary figure 4.tif