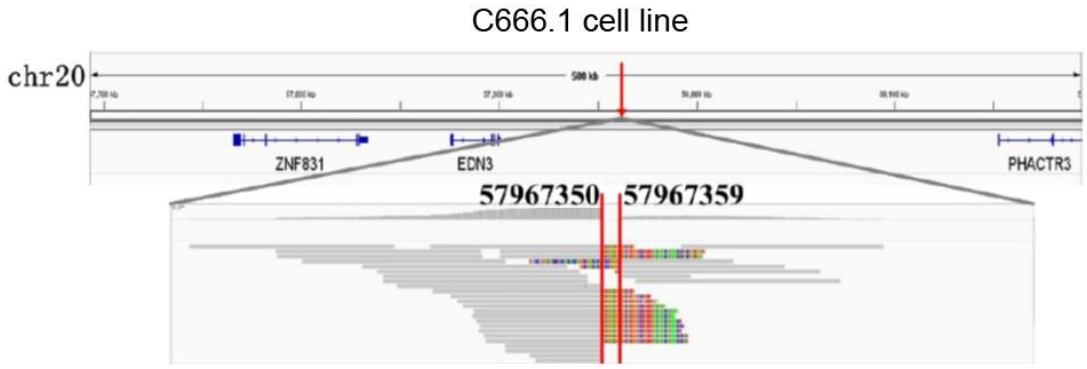


**Figure S1. Mapping of EBV integration breakpoint detected in C666-1 cell line**. EBV-human chimeric read pairs from bam file to detect EBV DNA integration in C666-1 cell line were mapped to chr20: 57967350-57967359 in the human hg19 reference genome.

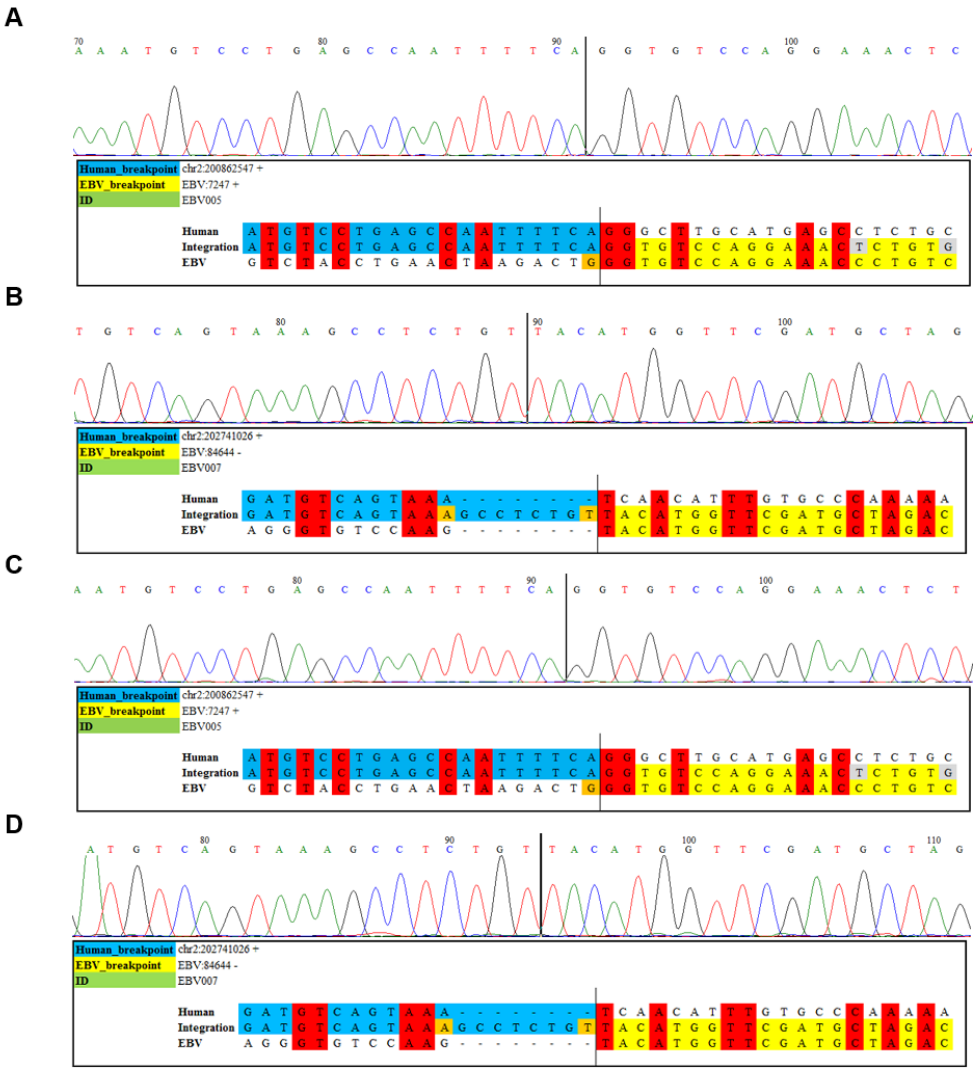


**Figure S2. Two EBV integrations were observed in both primary and metastatic NPC tumors from the same patient. (A-D)** Two EBV integrations were observed in both primary **(A-B)** and metastatic NPC tumors **(C-D)** from the same patient. Sequence alignment around the integration site between the human genome and EBV genome are presented. All EBV and human sequences are from the reference strand. Human sequences are in blue, and EBV sequences are in yellow. Nucleotides that align to both reference sequences are micro-homologies and are highlighted in red.

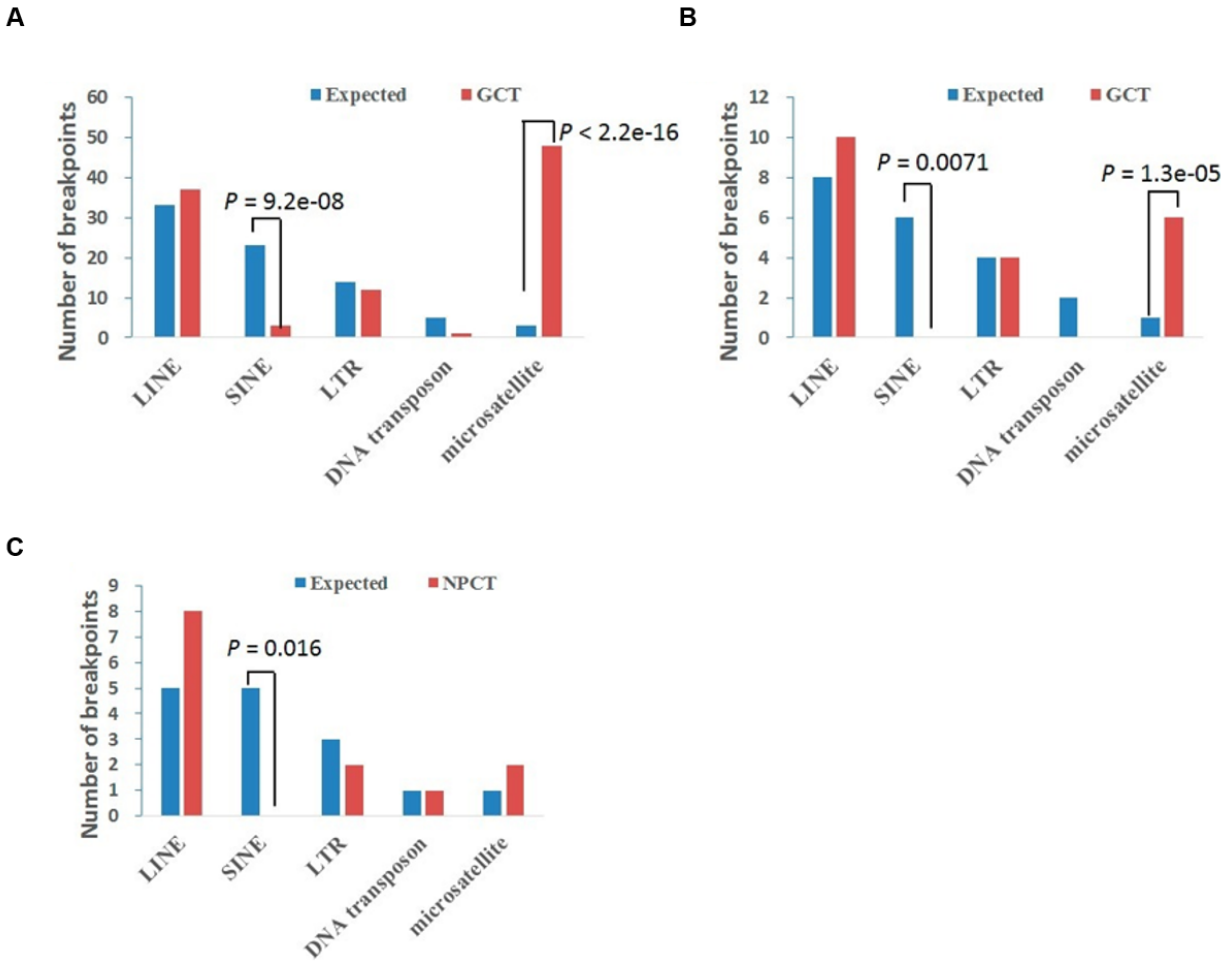


**Figure S3. Significant enrichment of integration breakpoints with microsatellite repeats in gastric carcinoma.** The expected and observed frequency of breakpoints from all gastric carcinomas **(A)** all gastric carcinomas except GCT015 **(B)** and NPC tumors **(C)** co-localized with repeat elements LINE, SINE, LTR, DNA transposon and microsatellite are shown. *P*-values were calculated using the binomial exact test. LINE, long interspersed nuclear element; SINE, long interspersed nuclear element; LTR, long terminal repeat.

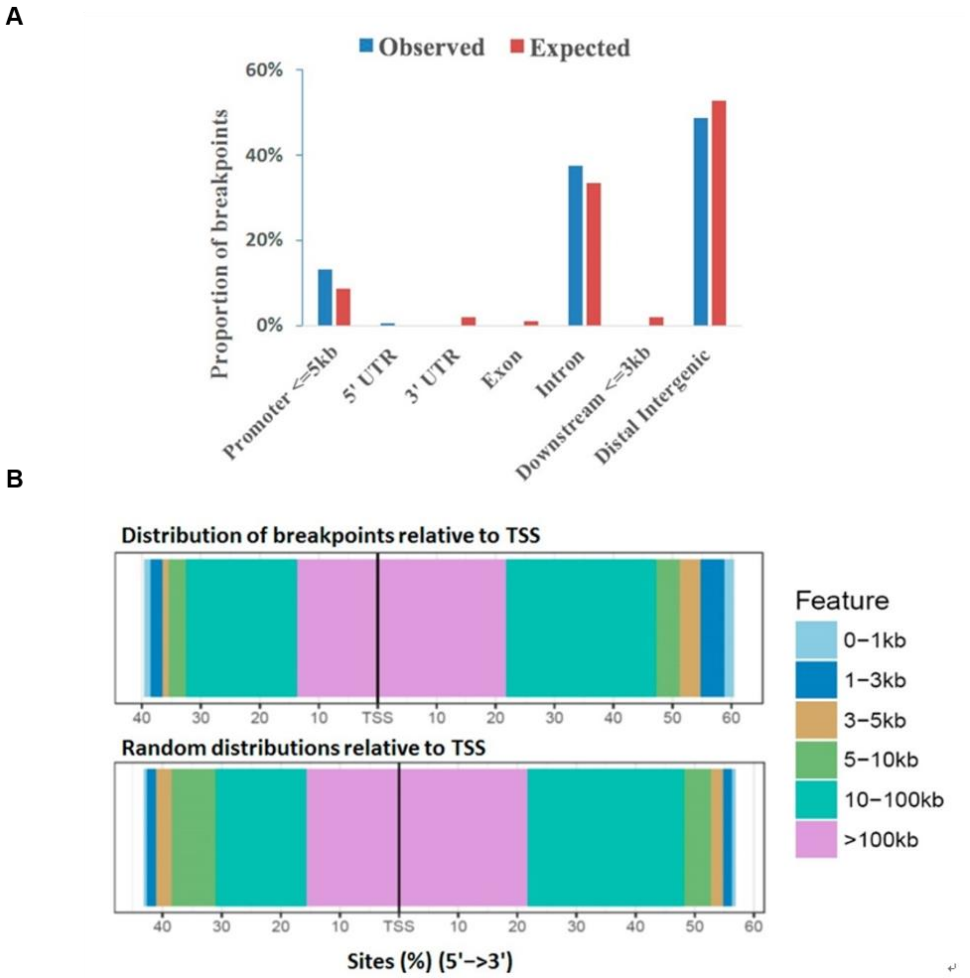


**Figure S4. Distribution of EBV integration breakpoints relative to genes.** **(A)** Proportions of EBV integration breakpoints located near genes and transcriptional start sites. **(B)** The observed (upper panel) and expected (lower panel) random distributions of breakpoints relative to transcriptional start sites (TSS) are shown. Different colored blocks indicate the distance of breakpoints to TSS and are ordered according to the 5' to 3' transcriptional direction. The X-axis and width of colored blocks show the proportion of breakpoints falling into each category indicated. The breakpoints tended to occur adjacent (smaller proportion of breakpoints with distance to TSS > 100 kb, purple block) to TSS more frequently than expected.

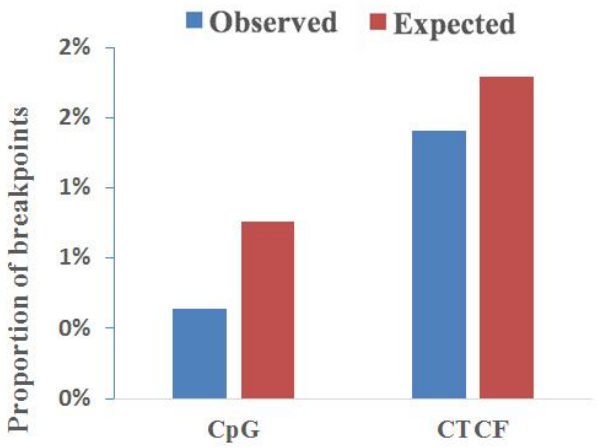


**Figure S5. Distribution of EBV integration breakpoints relative to CpG islands and CTCF binding sites.** The observed and expected proportions of EBV integration breakpoints co-localized with CpG islands and CTCF binding sites are shown.

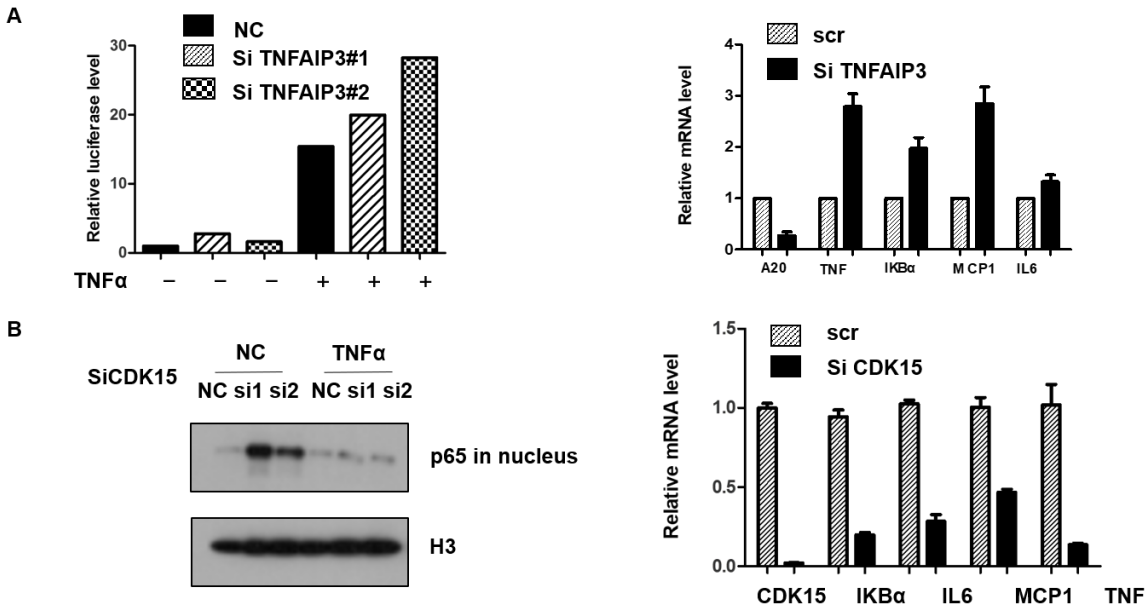


**Figure S6. Regulation of NF-κB pathway activity in NPC cells. (A)** *TNFAIP3* knockdown in NPC cells up-regulates NF-κB pathway activity. NF-κB–specific dual luciferase promoter assay to detect NF-κB activity in NPC cells with *TNFAIP3* knockdown. The relative promoter activity value was compared with the control. Quantitative PCR performed on RNA prepared from NPC cells with *TNFAIP3* knockdown to detect expression levels of NF-κB-pathway-targeted genes. **(B)** *CDK15* knockdown in NPC cells was down-regulated by NF-κB pathway activity. Western blot performed with nuclear lysates prepared from NPC cells with *CDK15* knockdown treated with *TNF-α* (10 mg/mL) and incubated with antibody against p65. Histone H3 was used as a control for loading. Quantitative PCR to detect levels of NF-κB-pathway-targeted genes in NPC cells with *CDK15* knockdown.

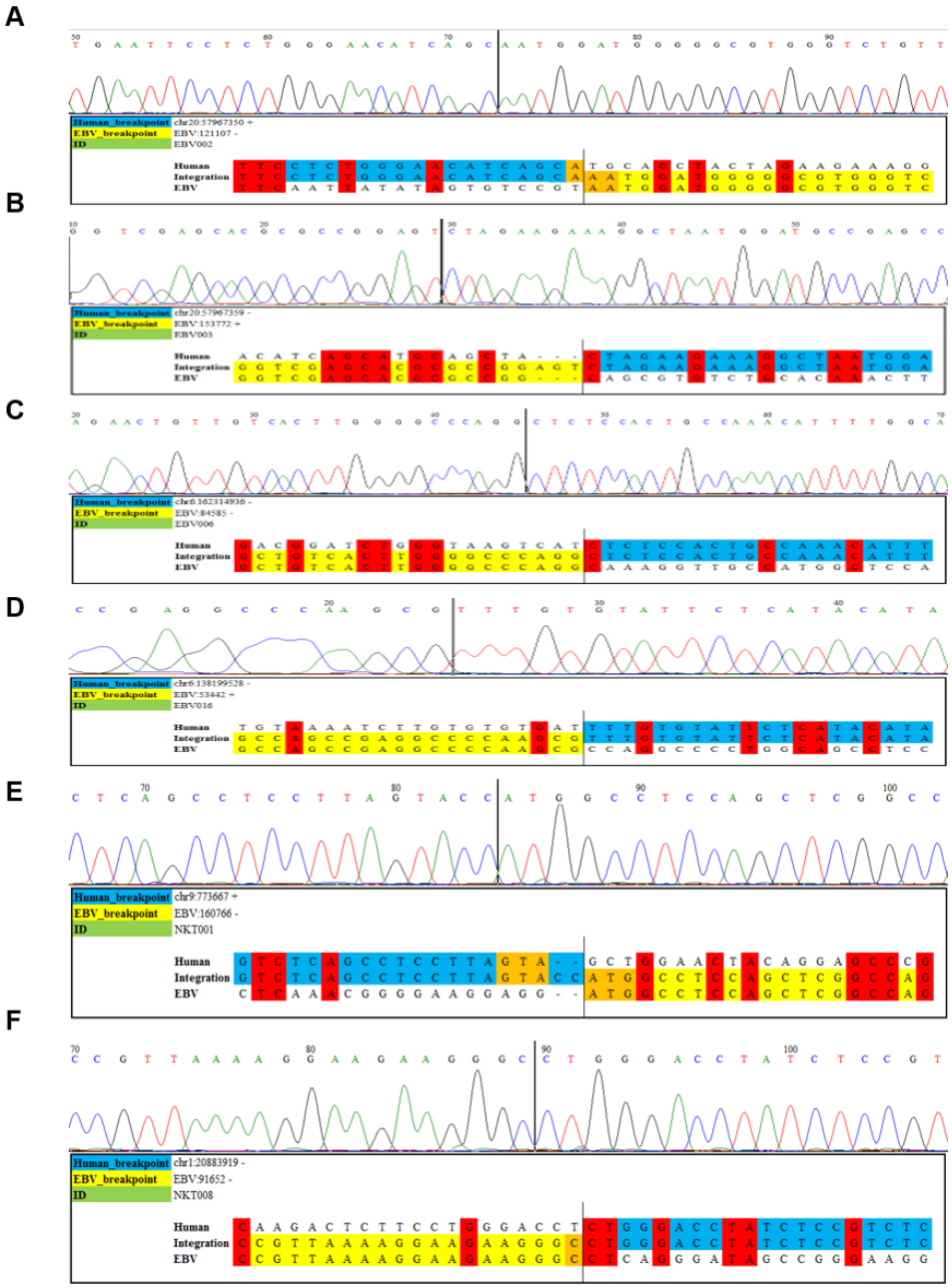


**Figure S7. Microhomology (MH) sequences in the regions flanking integration sites and Sanger sequencing validation.**

(A-F) Microhomology (MH) sequences in the regions flanking integration sites observed in NPC (A-D) and NK/T cell lymphoma (E-F) samples. All EBV and human sequences are from the reference strand. Human partner is in blue,

and EBV partner is in yellow. Nucleotides highlighted in red are micro-homologies aligning to both reference sequences.