SUPPLEMENTARY FIGURES

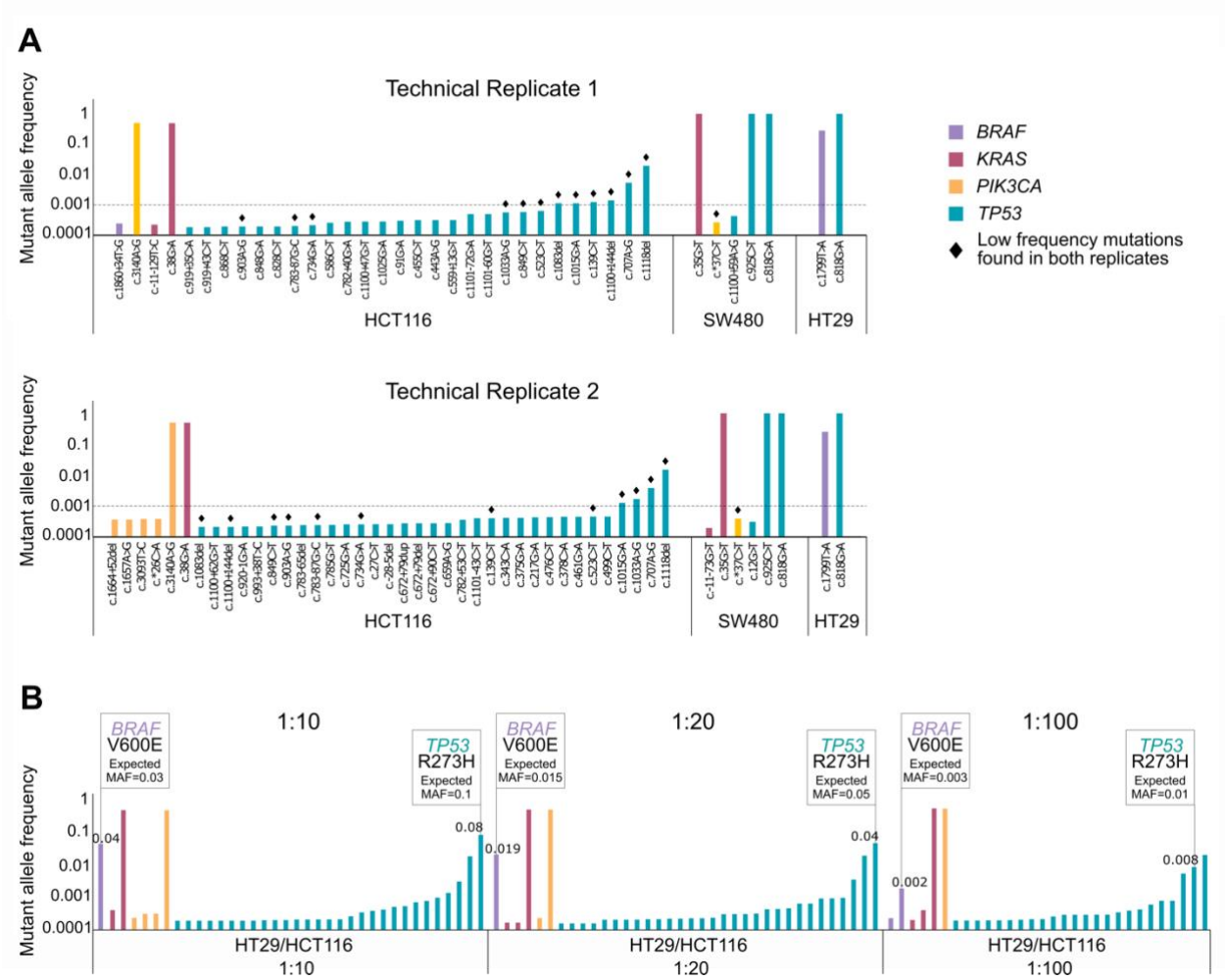

Figure S1. Ultra-deep sequencing of colorectal cancer cell lines with CRISPR-DS. A. Distribution of mutations in two technical replicates of HCT116, SW480 and HT29 cell lines. Each bar represents a mutation, color coded by gene. The height of the bars indicates Mutant Allele Frequency (MAF), calculated as the number of mutant alleles divided by the sequencing depth at a given position. Mutations are sorted by ascending MAF within each gene for each sample. Mutations found in both replicates are highlighted with a black diamond. **B.** Spike in of HT29 in HCT116 at different concentrations: 1/10, 1/20 and 1/100. Expected HT29 mutations and their MAF for each dilution are indicated with boxes and the observed frequency is indicated above each of the corresponding bars.

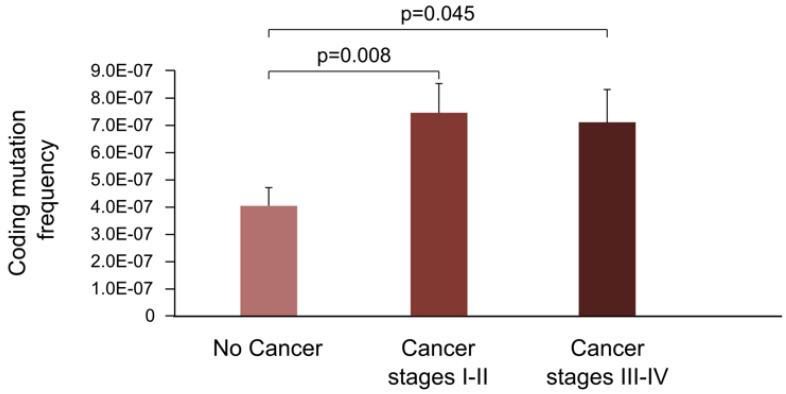

Figure S2. Normal colon from individuals with CRC carries higher coding mutation frequency than individuals without cancer regardless of the cancer being early or late stage. Coding mutation frequency is calculated as the number of mutations divided by the total number of duplex nucleotides sequenced in the coding target regions. P-values correspond to Mann-Whitney U test. Error bars represent standard error of the mean.

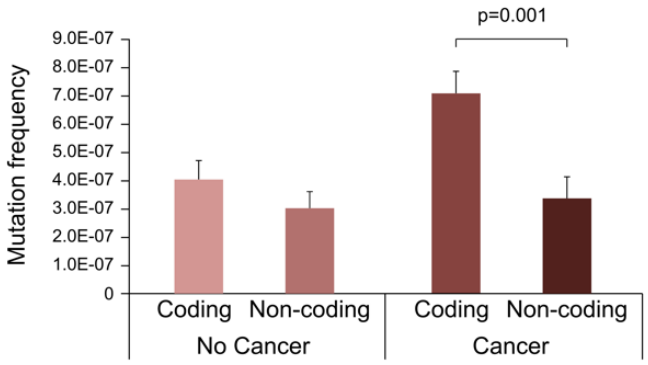

Figure S3. Coding mutation frequency is higher than non-coding mutation frequency in normal colon of individuals with CRC. Coding and non-coding mutation frequencies are calculated as the number of mutations divided by the total number of duplex nucleotides sequenced in the coding and non-coding target regions, respectively. P-value corresponds to Mann-Whitney U test. Error bars represent standard error of the mean.

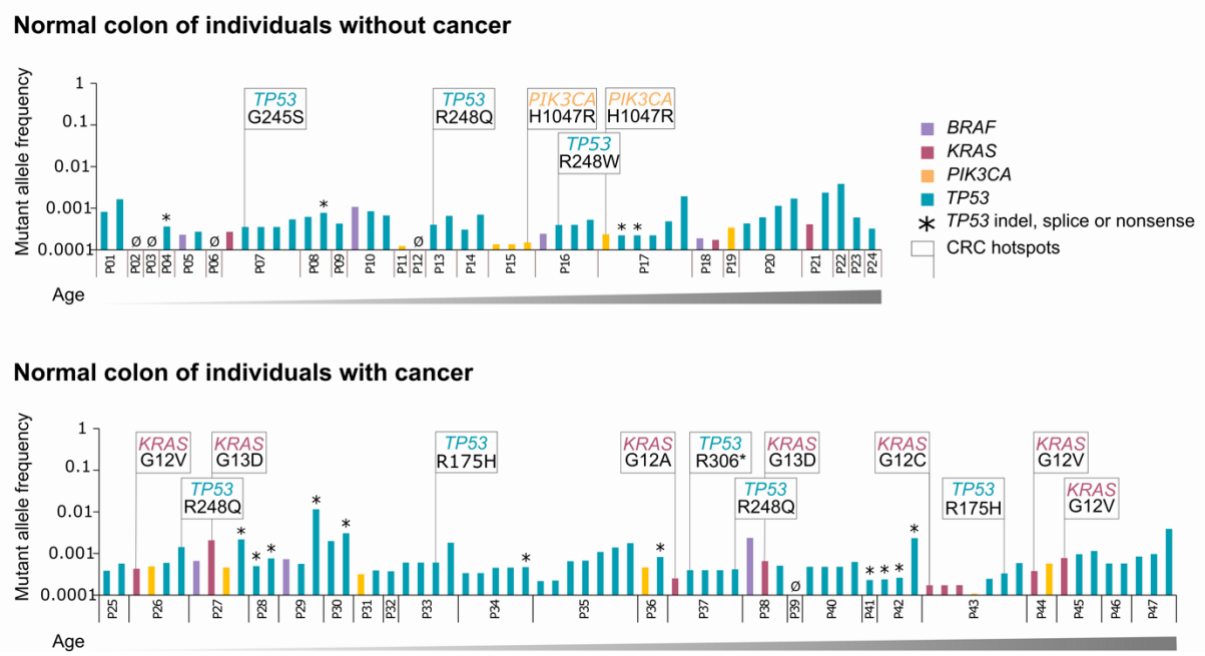

Figure S4. Normal colon from individuals with and without CRC carries mutations in common cancer genes. Distribution of mutations in normal colon from individuals without CRC (top) and individuals with CRC (bottom)**.** Patients are sorted by ascending age and patient IDs are indicated in the x-axis. Each bar represents a mutation, color coded by gene. The height of the bars indicates Mutant Allele Frequency (MAF), calculated as the number of mutant alleles divided by the sequencing depth at a given position. Mutations are sorted by ascending MAF within each gene for each patient. Hotspot mutations are indicated with boxes and *TP53* indels, nonsense, and splice mutations are indicated with asterisks.

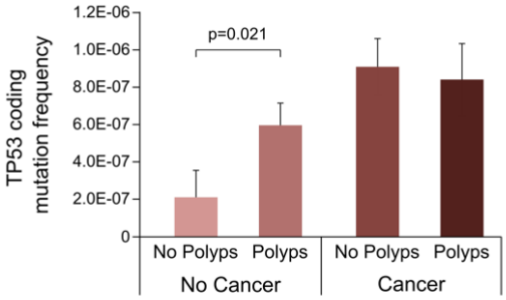

Figure S5. *TP53* coding mutations are more frequent in normal colon from individuals without CRC that have a history of polyp formation. Comparison of *TP53* mutation frequency in the normal colon of polyp and non-polyp formers with and without CRC. P-value corresponds to Mann-Whitney U test. Error bars represent standard error of the mean.