



Figure 1 – figure supplement 2: The relative indel frequency defined as the number of indels divided by the total bases of non-homopolymer regions in MMR-deficient tumors stratified into intergenic, exonic, 5'UTR, 3'UTR and intronic regions is shown. Indel frequencies in homopolymers are shown in the left panel, whereas indel frequencies in non-homopolymer regions are shown in the right panel. The algorithm we used to correct for homopolymer content, composition and length can be found in the supplementary file. In homopolymer regions, there was a 16% decrease in indel frequency in exonic regions. In non-homopolymer regions, a clear decrease was also observed for exonic regions, confirming that the decrease in exonic indels is not only due to differences in homopolymer characteristics between exonic regions and the rest of the genome. This reveals apparent negative selection in exonic regions, independent of homopolmer content, composition or length of the homopolymers.