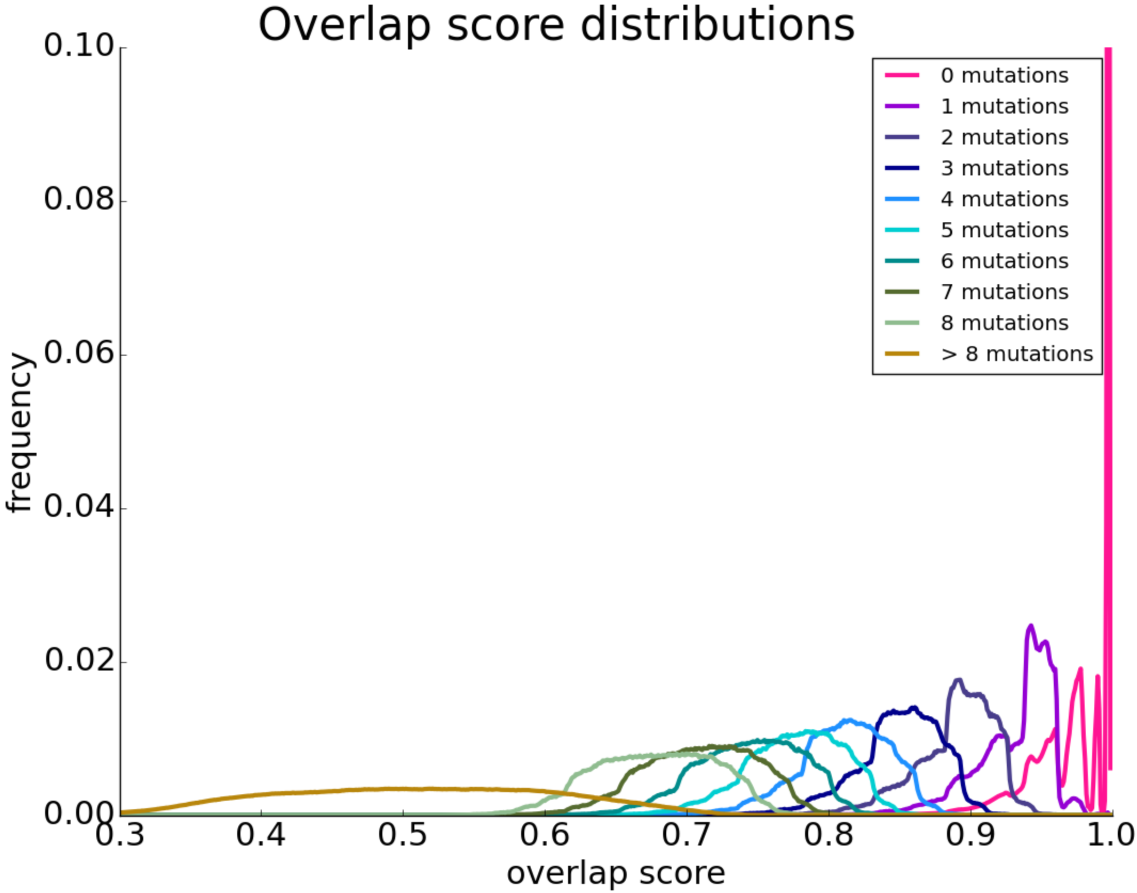

Supplemental Fig S2: Overlap score distributions. To analyze the behaviour of the overlap score function, we simulated 2x250 bp Illumina MiSeq reads from different genomes, diverging between 1% and 10%. We computed all overlaps among those reads and classified them by the number of true mutations in the overlap (not counting mismatches that are due to sequencing errors). This resulted in distributions P_i, $i \geq 0$, representing the overlap scores found in case of i true mutations.