



**Figure S2. COVGAP identifies all SNPs in the mock genomes. A.** Levenshtein distance between the mutations in the input mock genomes (y axis) and in the genomes recovered by COVGAP (x axis); the marginal plots show the frequency of presence / detection across samples. Only two deletions (5845, 16281) and one insertion (16145) were not detected. **B.** Phylogeny of input and COVGAP-derived consensus output genomes, showing that all SNPs were identified.