Supplemental Figures

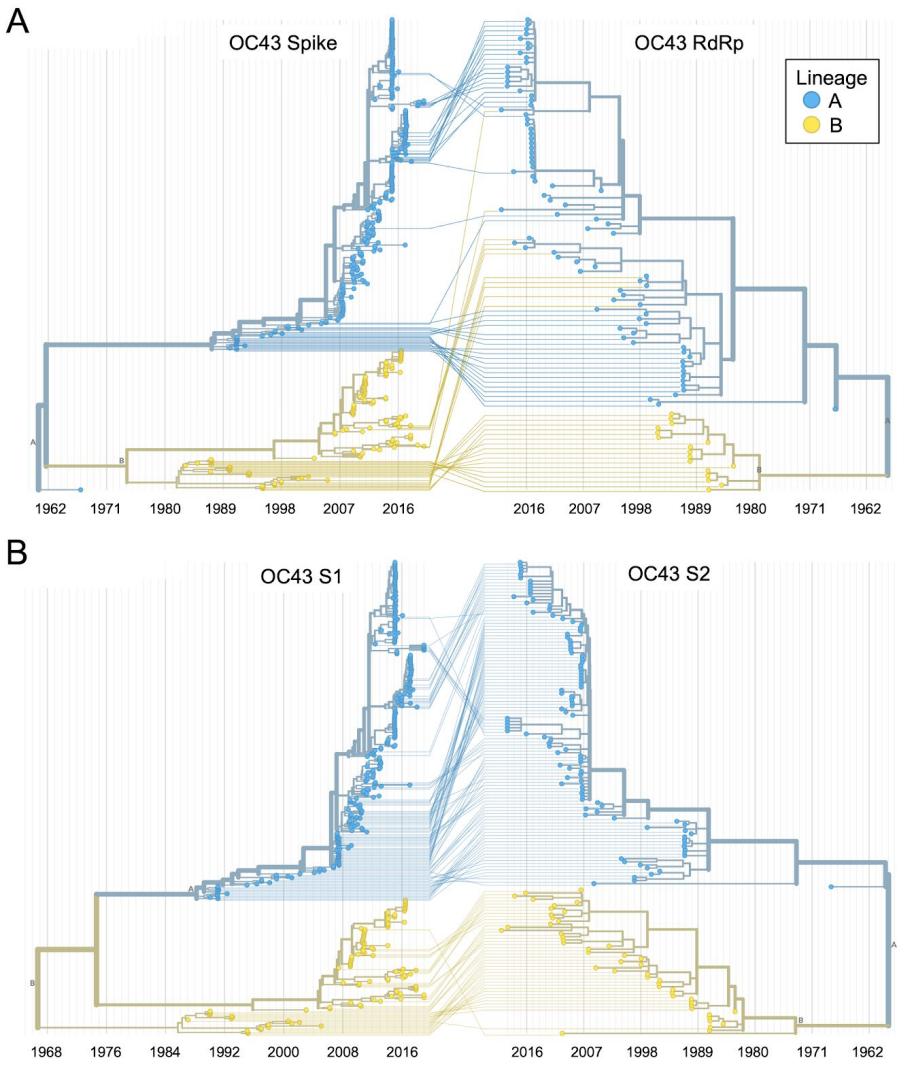

Figure 1 Supplement 1. Recombination occurs between HCoV isolates. A tanglegram draws lines between an isolate's position on two phylogenies built on different genes (or genomic regions). Dramatic differences in an isolate's position on one tree versus another is indicative of recombination. A) Phylogenetic relationships between OC43 isolates based on RdRp sequences versus relationships based on Spike sequences. Blue lines that connect isolates classified as lineage A based on their RdRp sequence to isolates classified as lineage B based on their Spike sequence, suggest that recombination occurred in these isolates or their ancestors. B) Phylogenetic reconstruction of OC43 isolates based on S1 sequences versus S2 sequences. Year is shown on the x-axis.

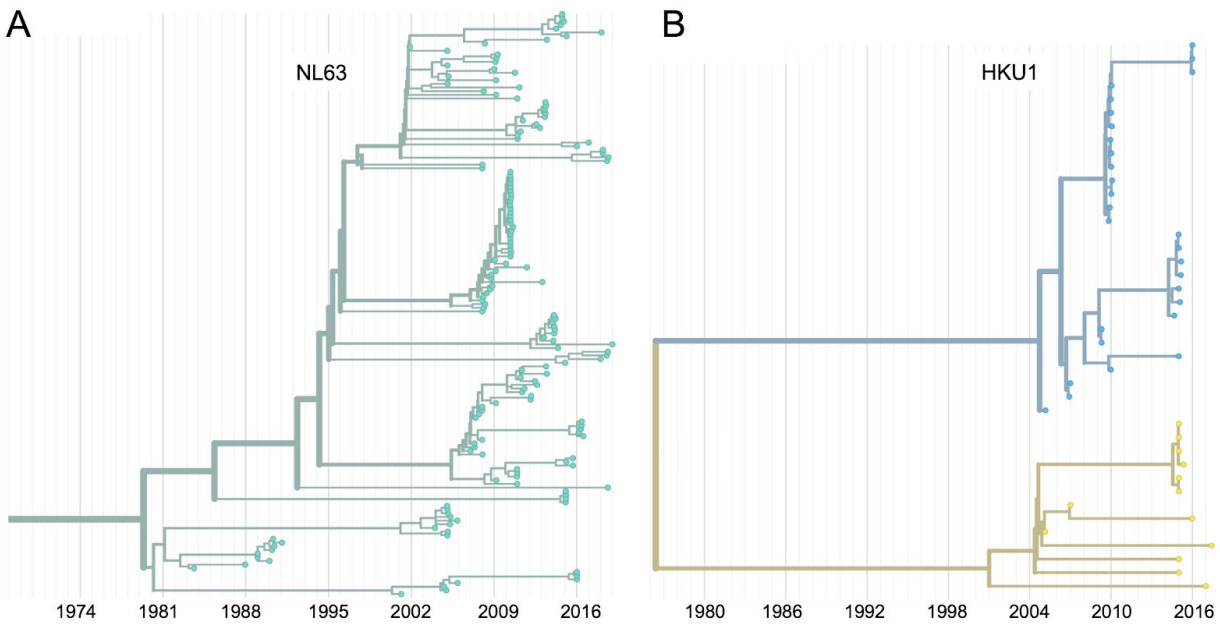

Figure 1 Supplement 2. Phylogenetic trees for seasonal HCoVs NL63 and HKU1. Phylogenies built from A: NL63 spike sequences from 159 isolates over 37 years, and B: HKU1 spike sequences from 41 isolates over 13 years. HCoVs that bifurcate immediately after the root are split into blue and yellow lineages. NL63 contains just one lineage (teal). Both HCoVs are rooted on an outgroup sequence. For the analyses in this paper, the evolution of each gene (or genomic region) is considered separately, so phylogenies are built for each viral gene and those phylogenies are used to split isolates into lineages for each gene. These are temporally resolved phylogenies with year shown on the x-axis. The clock rate of each HCoV is listed in the Methods “Phylogenetic inference” section.

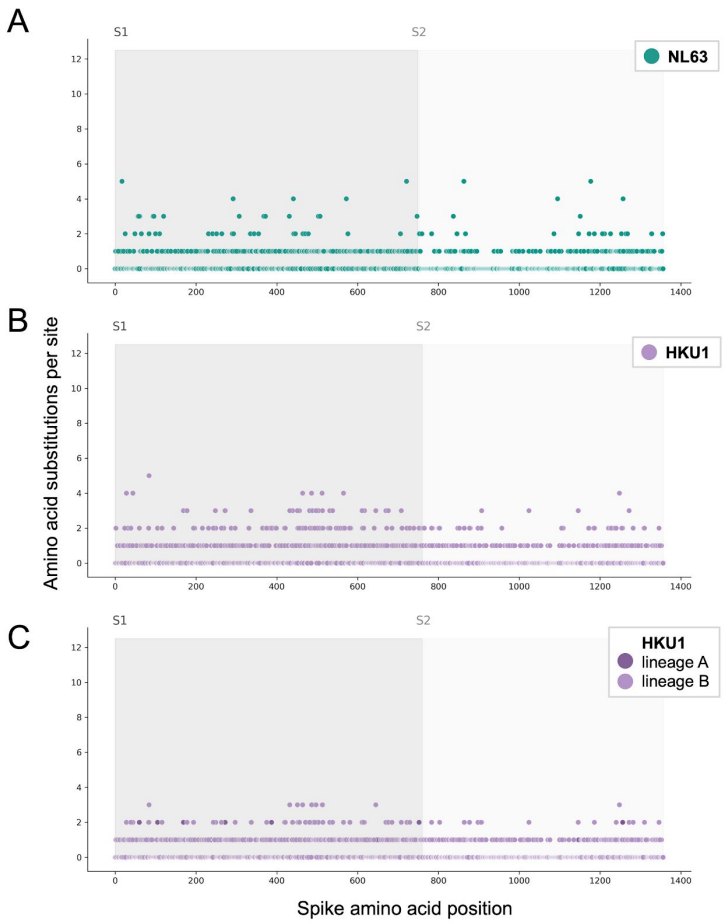

Figure 2 Supplement 1. Mutations per at each position within Spike for NL63 and HKU1. Number of mutations observed at each position in the Spike gene. S1 (darker gray) and S2 (light gray) are indicated by shading and the average number of mutations per site is indicated by a dot and color-coded by HCoV lineage. A: NL63, B: HKU1 (assuming all HKU1 isolates are a single lineage), C: HKU1 (assuming there are 2 co-circulating HKU1 lineages).

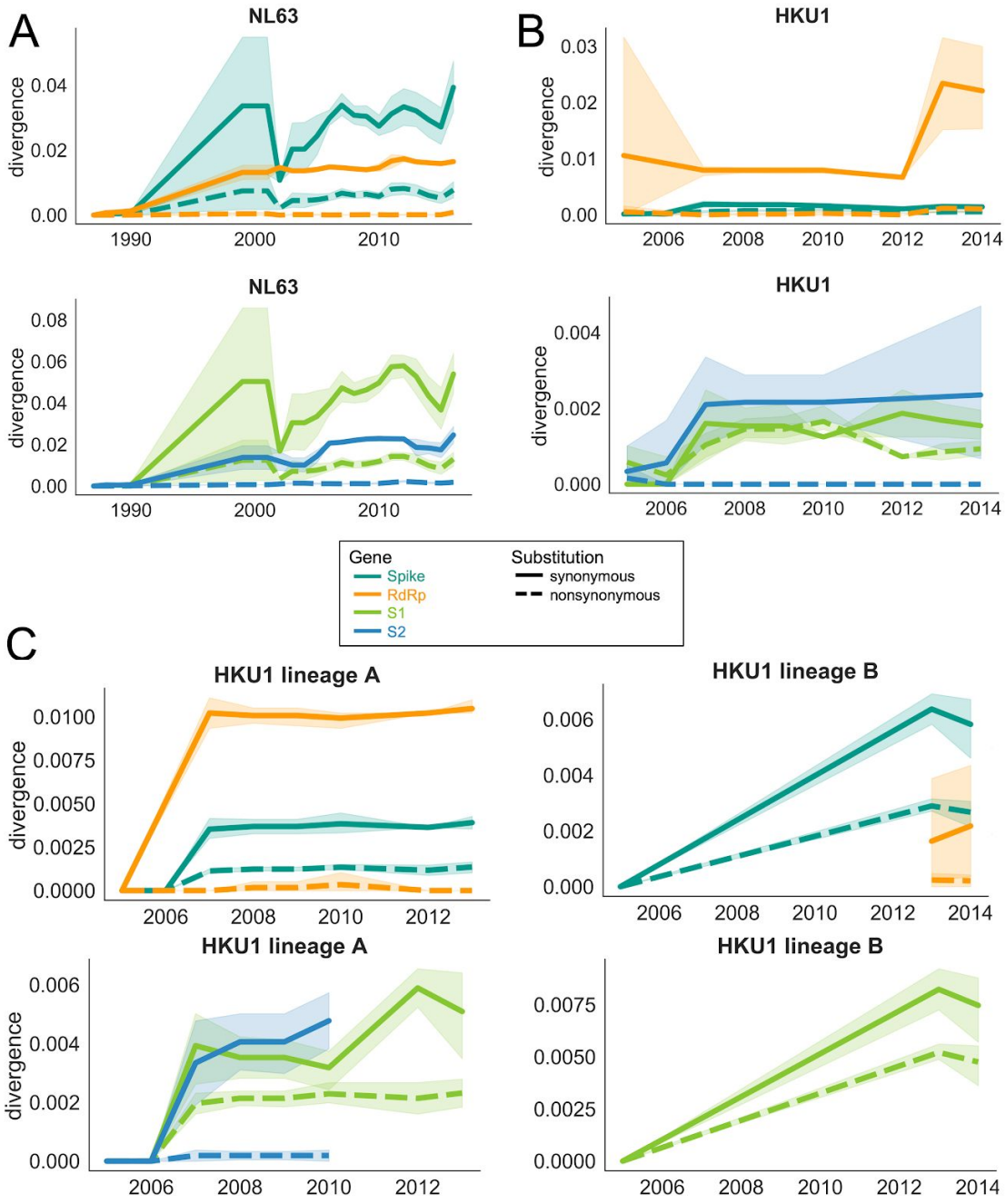

Figure 3 Supplement 1. Nonsynonymous divergence in NL63 and HKU1. Nonsynonymous (dashed lines) and synonymous divergence (solid lines) within the Spike (teal) and RdRp (orange) genes and within S1 (light green) and S2 (blue) over time. Divergence is the average Hamming distance from the ancestral sequence, computed in sliding 3-year windows which contain at least 2 sequenced isolates. Shaded region shows 95% confidence intervals. A: NL63, B: HKU1 (assuming all HKU1 isolates belong to a single lineage), and C: HKU1 (divided into 2 co-circulating lineages). Year is shown on the x-axis. Note that x- and y-axis scales are not shared between plots.

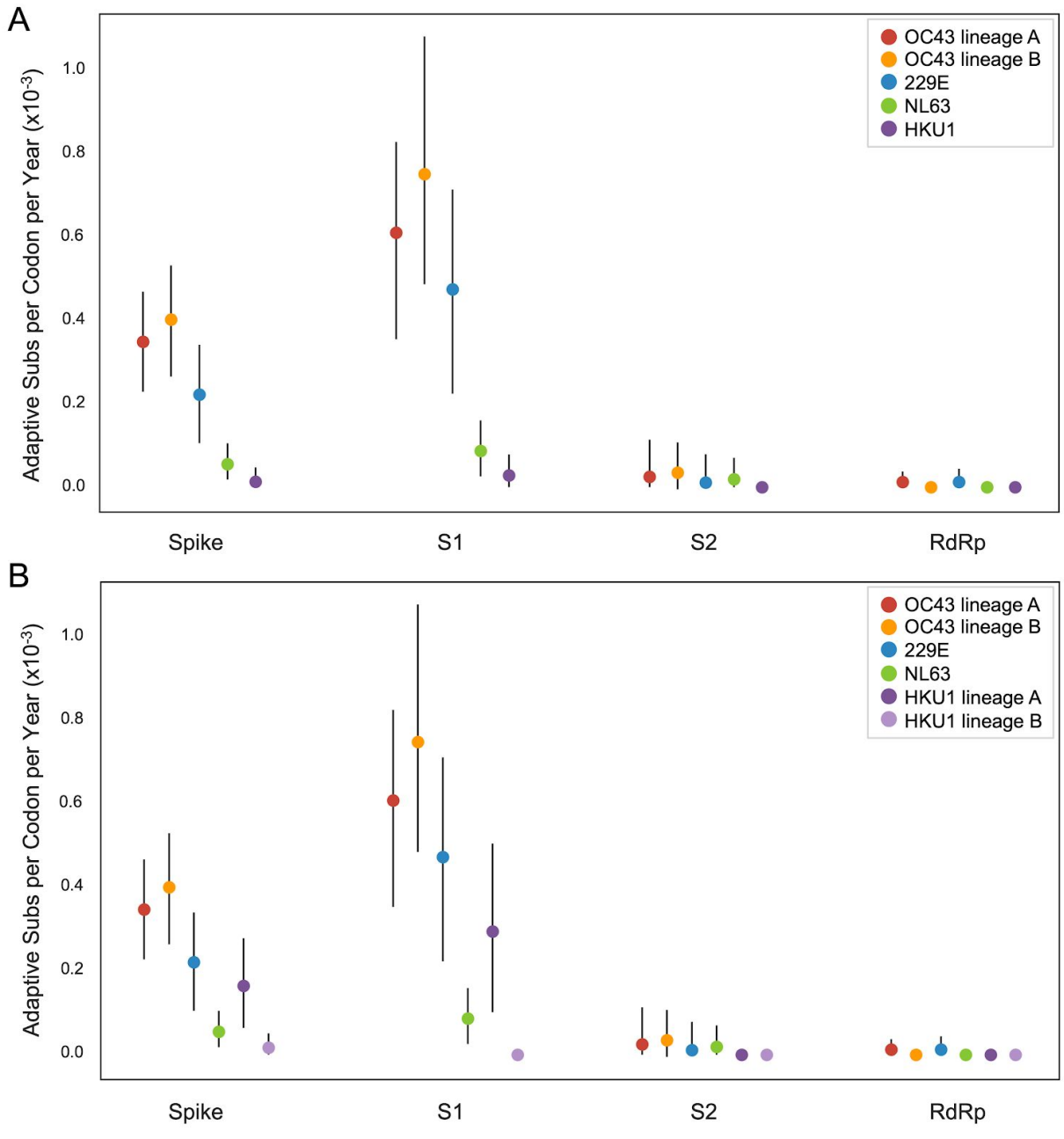

Figure 5 Supplement 1. NL63 and HKU1 have low rates of adaptation in Spike. As in Figure 4, adaptive substitutions per codon per year are calculated by our implementation of the Bhatt method. A: NL63 (teal) and HKU1 (purple) are both considered to consist of a single lineage. B: HKU1 is divided into 2 co-circulating lineages (dark and light purple). The calculated rates of adaptive substitution within Spike, S1, S2 and RdRp are plotted alongside 229E and OC43 for comparison. Error bars show 95% bootstrap percentiles from 100 bootstrapped datasets

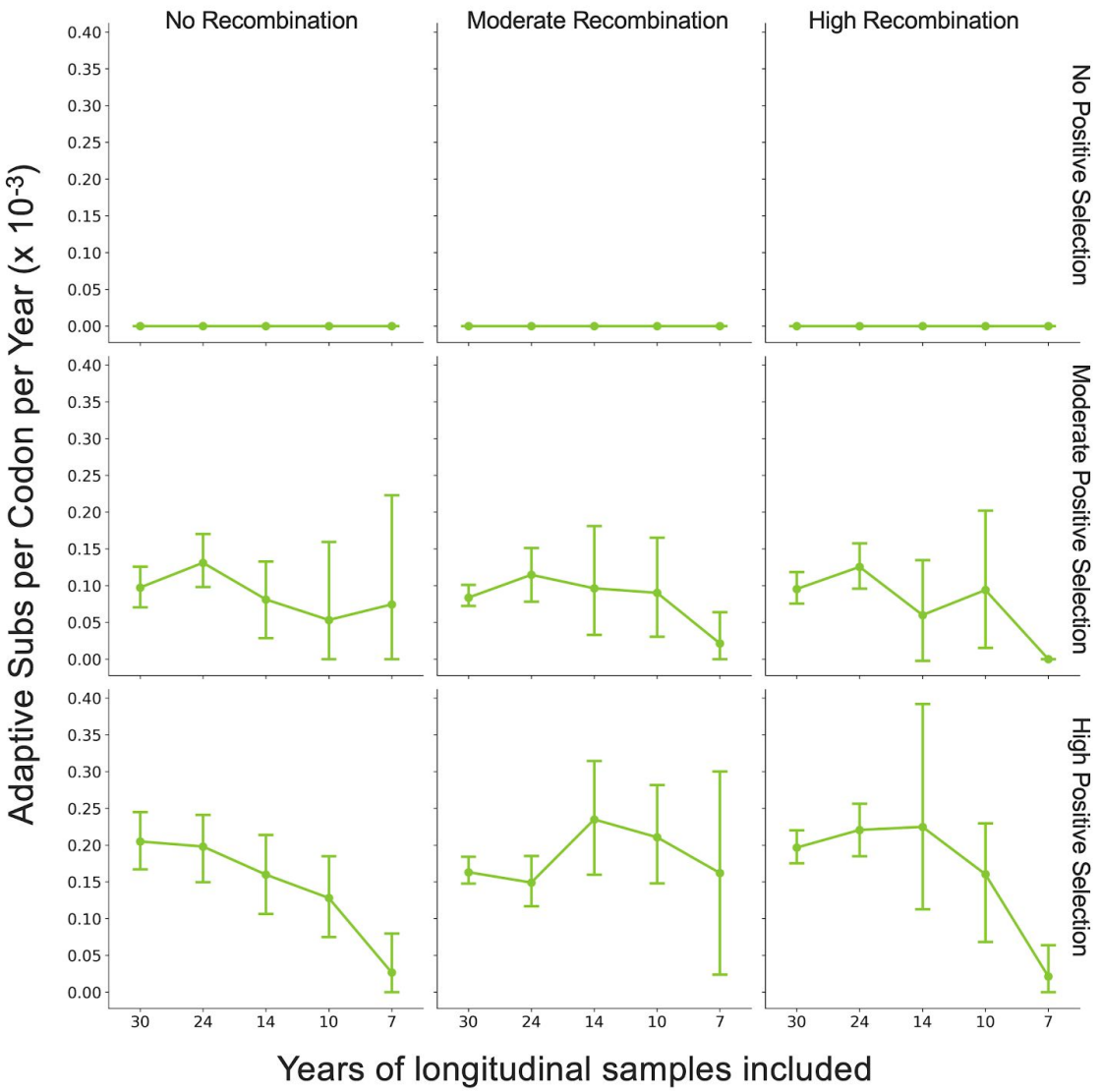

Figure 7 Supplement 1. Fewer years of longitudinally-sampled isolates reduces ability to detect rate of adaptation. OC43 lineage A S1 sequences were simulated under conditions of no, moderate and high rates of recombination in combination with no, moderate or high strength of positive selection. The Bhatt method was used to calculate the rate of adaptive evolution under each of these scenarios using all available sequence data (30 years), or only the most recent 24, 14, 10 or 7 years of simulated sequences.