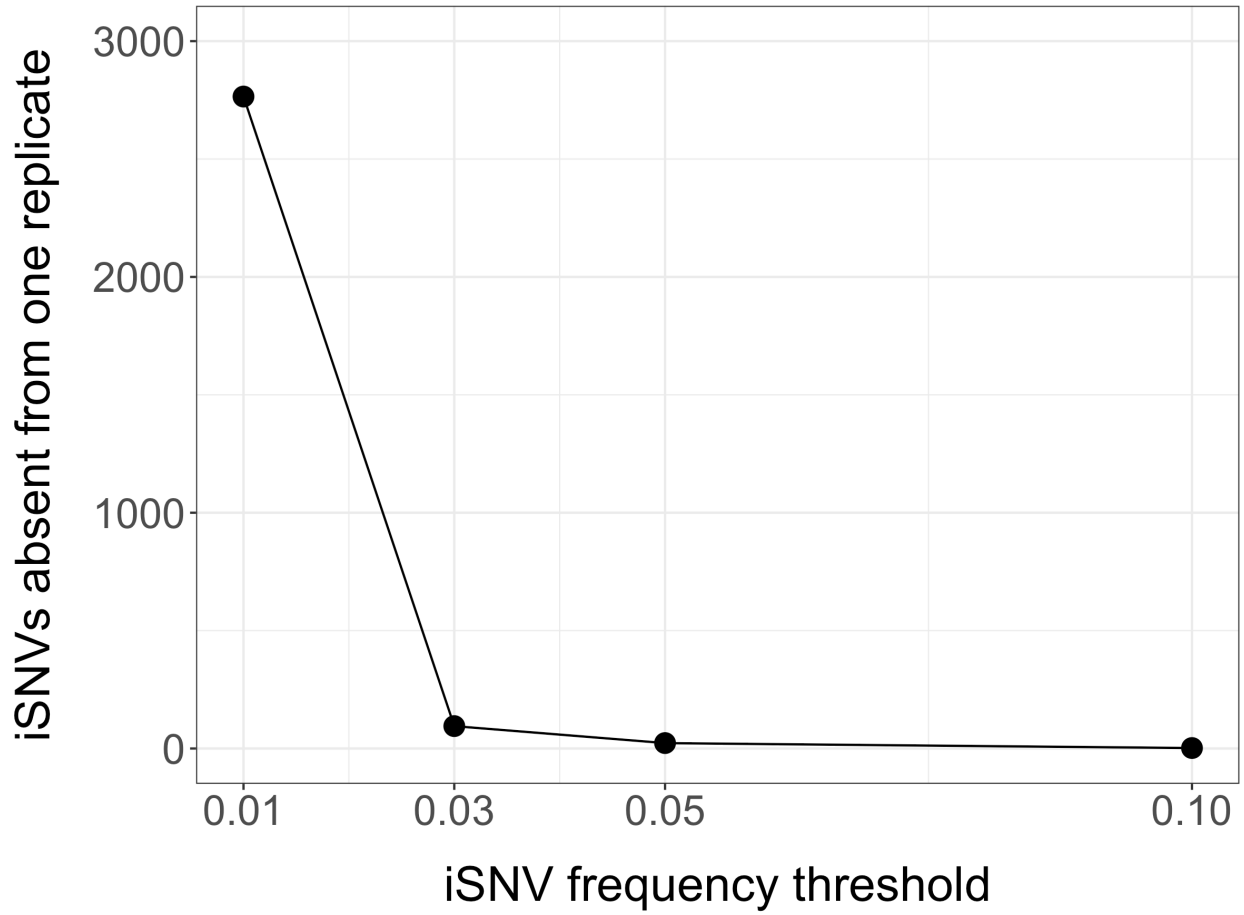
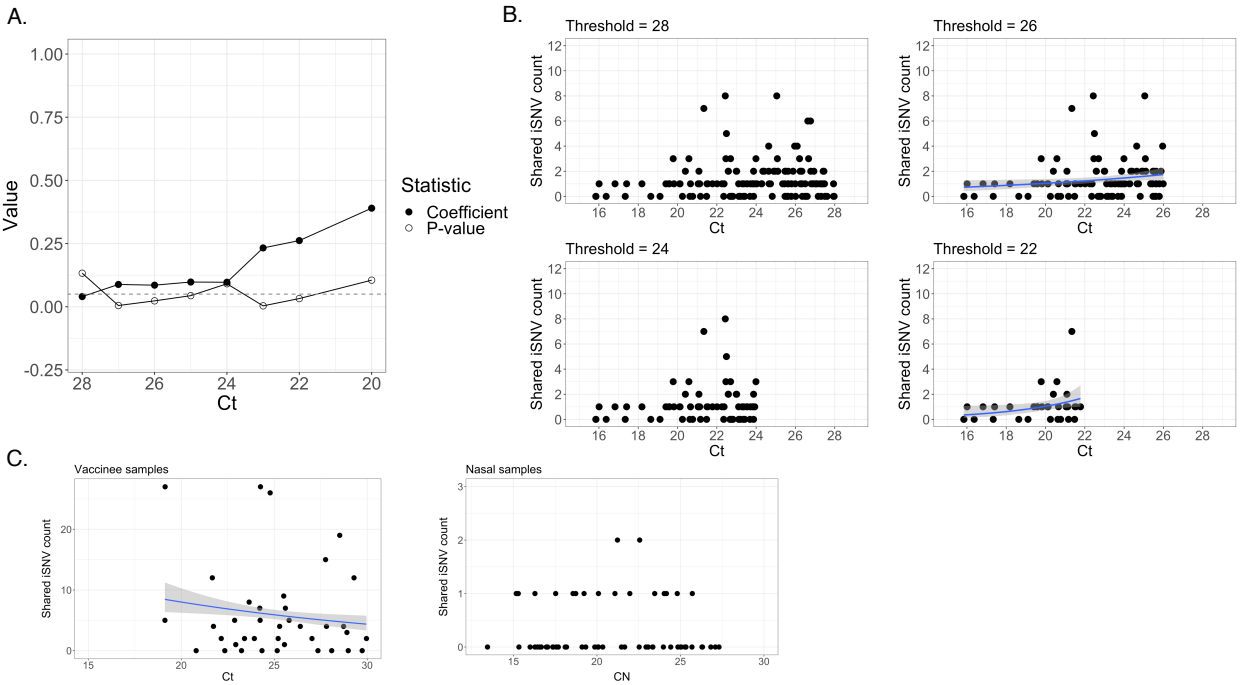
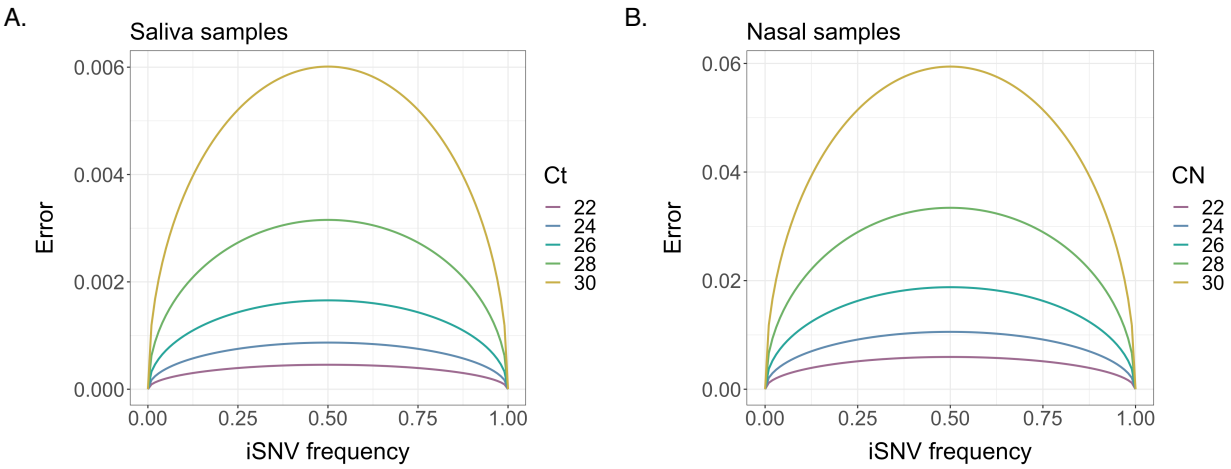


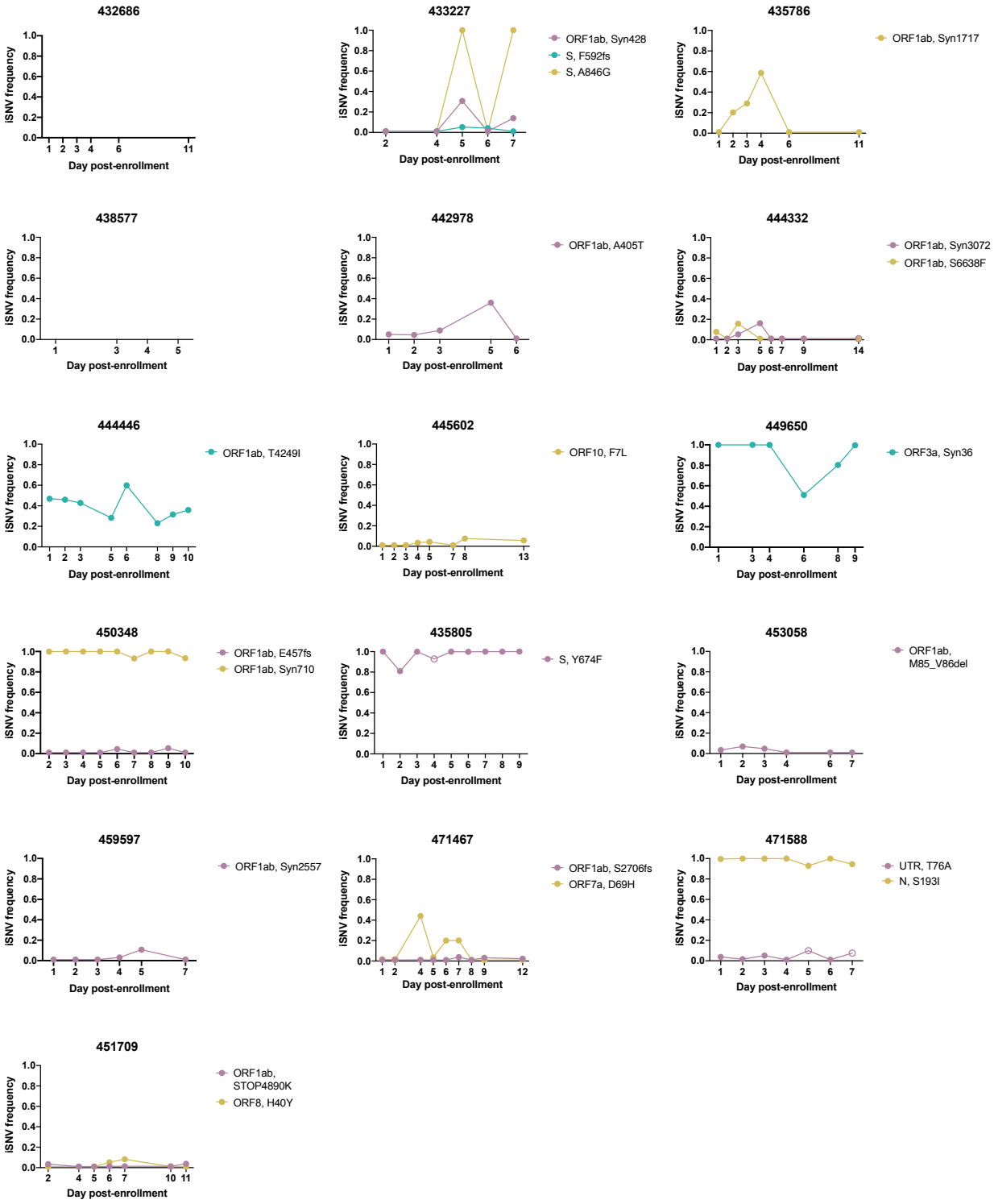
Supplemental Figure 1. Effect of iSNV frequency cutoff on number of iSNVs absent between technical replicates. Number of iSNVs present in one technical replicate but absent in the other, with iSNVs called using a range of frequency cutoffs from 0.01 – 0.10. Missing iSNVs were counted using a pooled dataset combining all spike-in dilutions at all Ct values.



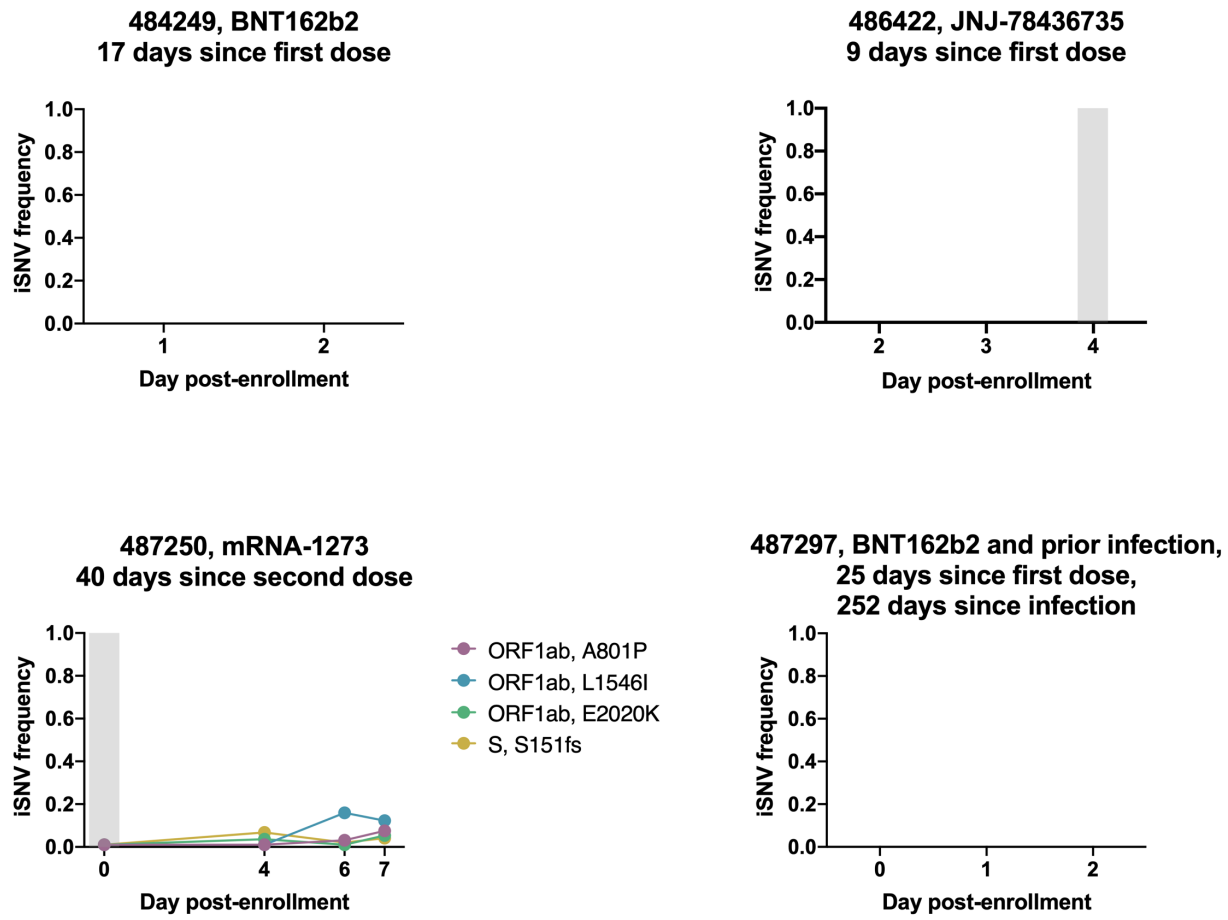
Supplemental Figure 2. Effect of Ct cutoff thresholds on correlation between Ct and shared iSNVs. (A) *Regression coefficients and P-values for Poisson regressions between Ct and shared iSNV count on datasets with increasingly strict Ct cutoffs. Dashed line indicates p = 0.05.* ***(B)*** *Plotted comparison of Ct vs. shared iSNV count for Ct<28 (Poisson regression, reg. coef. = 0.0402, p = 0.133), Ct<26 (reg. coef. = 0.0854, p = 0.0233), Ct<24 (reg. coef. = 0.0970, p = 0.909), and Ct<22 (reg. coef. = 0.262, p = 0.0324) datasets.* ***(C)*** *Plotted comparison of Ct vs. shared iSNV count for vaccinee samples (Poisson regression, reg. coef. = -0.0608, p = 0.00919).* ***(D)*** *Plotted comparison of CN vs. shared iSNV count for nasal samples (Poisson regression, reg. coef. = -0.0102, p = 0.869).*



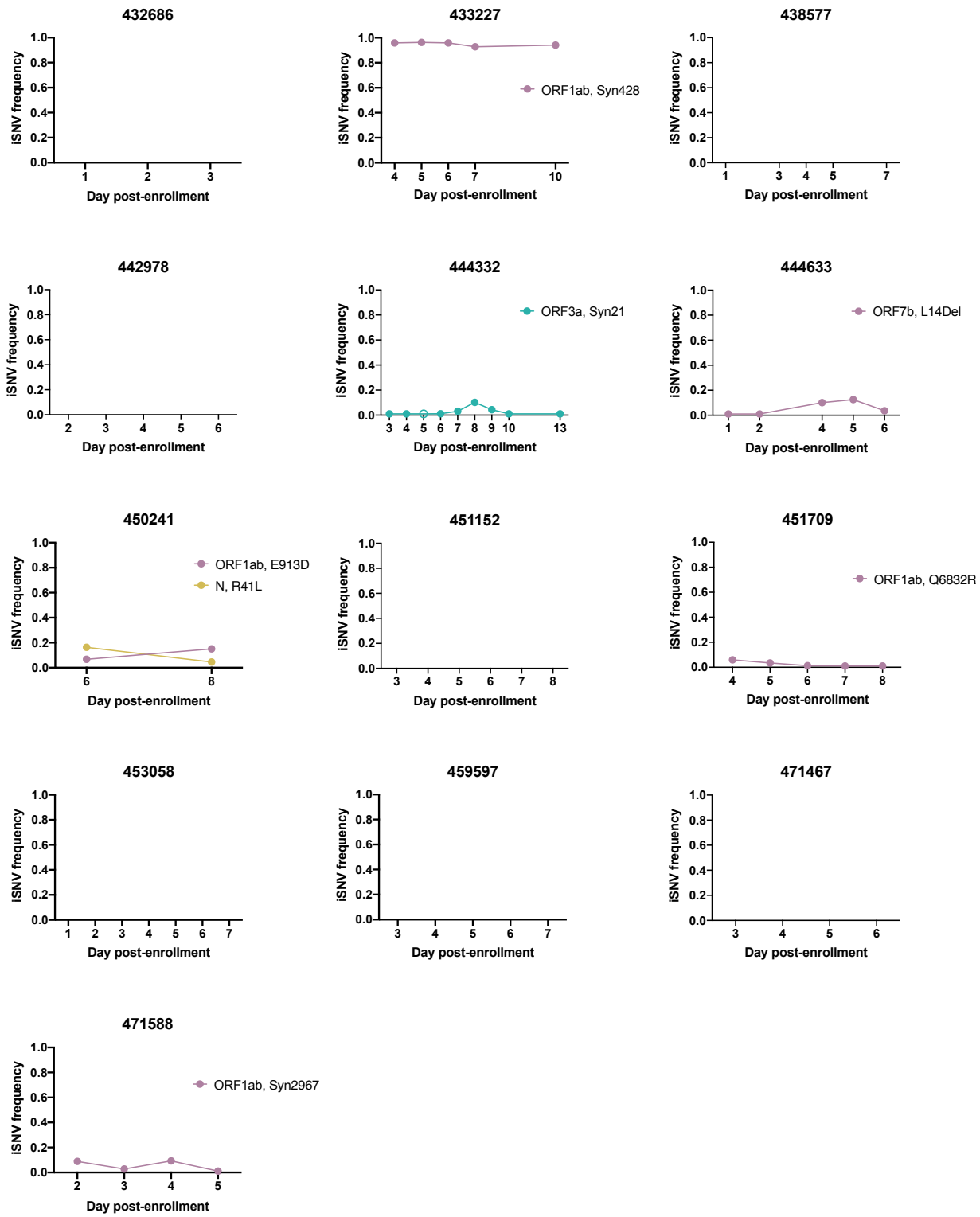
Supplemental Figure 3. Standard error estimates for saliva and nasal samples. (A) *Error estimates for saliva samples with Ct values ranging from 22 – 30, at frequencies of 0.0 – 1.0. Genome load was estimated based on Ct value, as in Ke et al. 2022* [15]*.* ***(B)*** *Error estimates for nasal samples with CN values ranging from 22 – 30, at frequencies of 0.0 – 1.0. Genome load was estimated based on CN value, as in Ke et al. 2022*[15]*.*



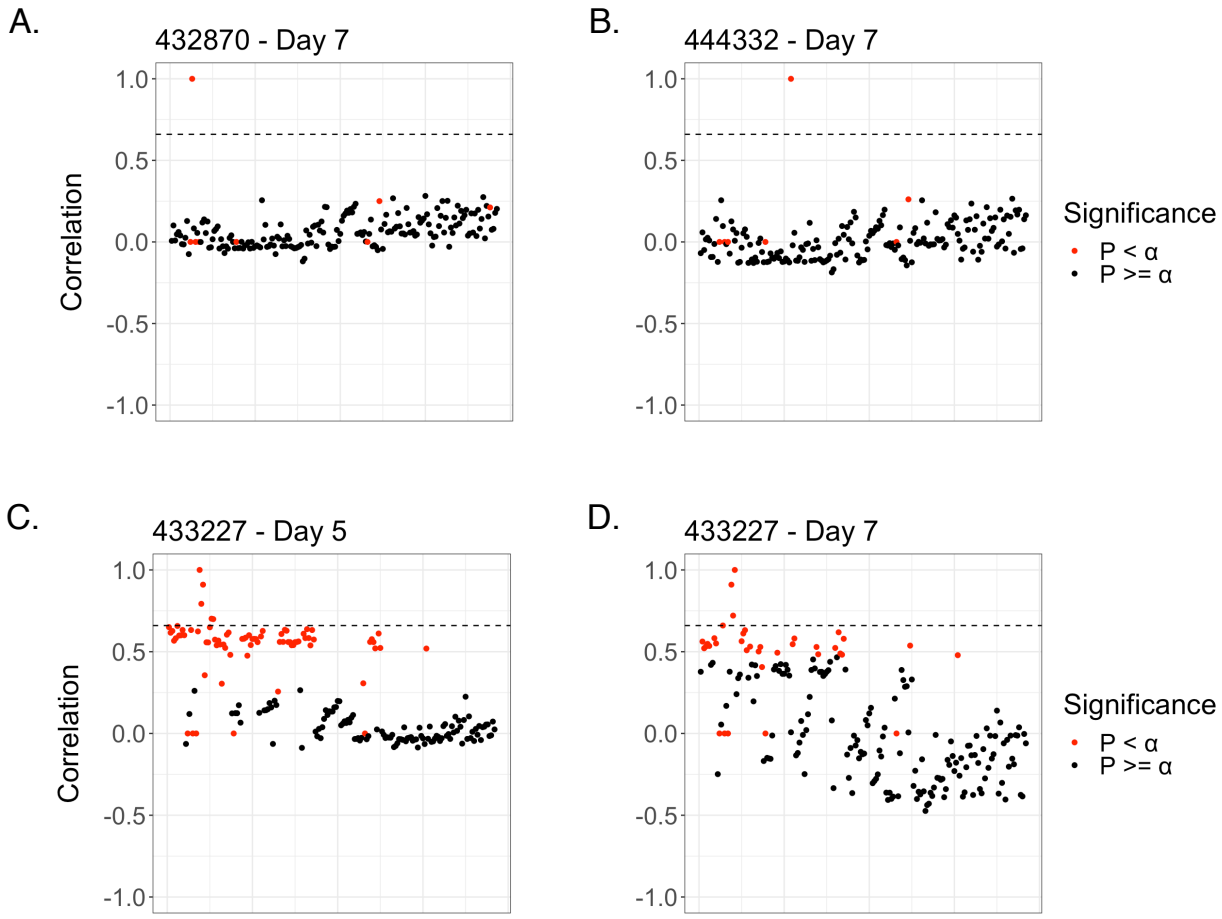
Supplemental Figure 4. iSNV dynamics over time from additional unvaccinated participants. *Frequency tracking of selected iSNVs in unvaccinated participants. Unfilled points mark iSNVs with read depths below the threshold of 1000 reads.*



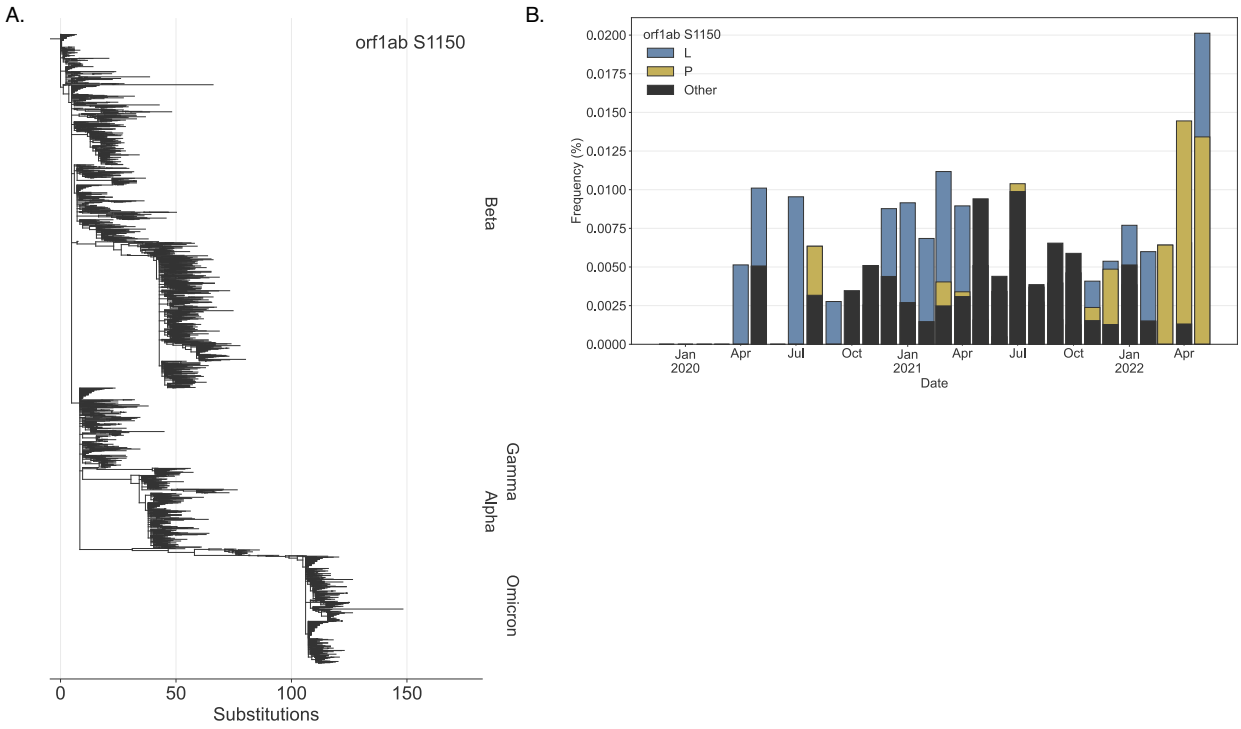
Supplemental Figure 5. iSNV dynamics over time from additional vaccinated participants. *Frequency tracking of selected iSNVs in vaccinated participants. Unfilled points mark iSNVs with read depths below the threshold of 1000 reads. Gray boxes mark samples with mean per-nucleotide coverages below 1000 reads. Panel headings indicate vaccine received and time between enrollment and last vaccine dose.*



Supplemental Figure 6. iSNV dynamics over time from all nasal samples. *Frequency tracking of selected iSNVs from nasal swab samples. Unfilled points mark iSNVs with read depths below the threshold of 1000 reads. Gray boxes mark samples with mean per-nucleotide coverages below 1000 reads.*



Supplemental Figure 7. Correlations between samples of interest and all other samples in dataset. *Pearson correlations between samples suspected of contamination and each other sample in the dataset. Dashed line marks threshold of minimum correlation between identical samples, as determined using sequencing replicates.* α *= 0.05/192 samples = 0.000260.*



***Supplemental Figure 8. Global occurrence of selected iSNVs**. **(A)** ORF1ab:S1150-associated amino acid changes plotted on a downsampled (100 sequences per month) global phylogeny of SARS-CoV-2 sequences (no deviations from WT are observed). **(B)** Frequency of ORF1ab:S1150-associated amino acid changes from January 2020 to April 2022 in all quality filtered SARS-CoV-2 sequences.*