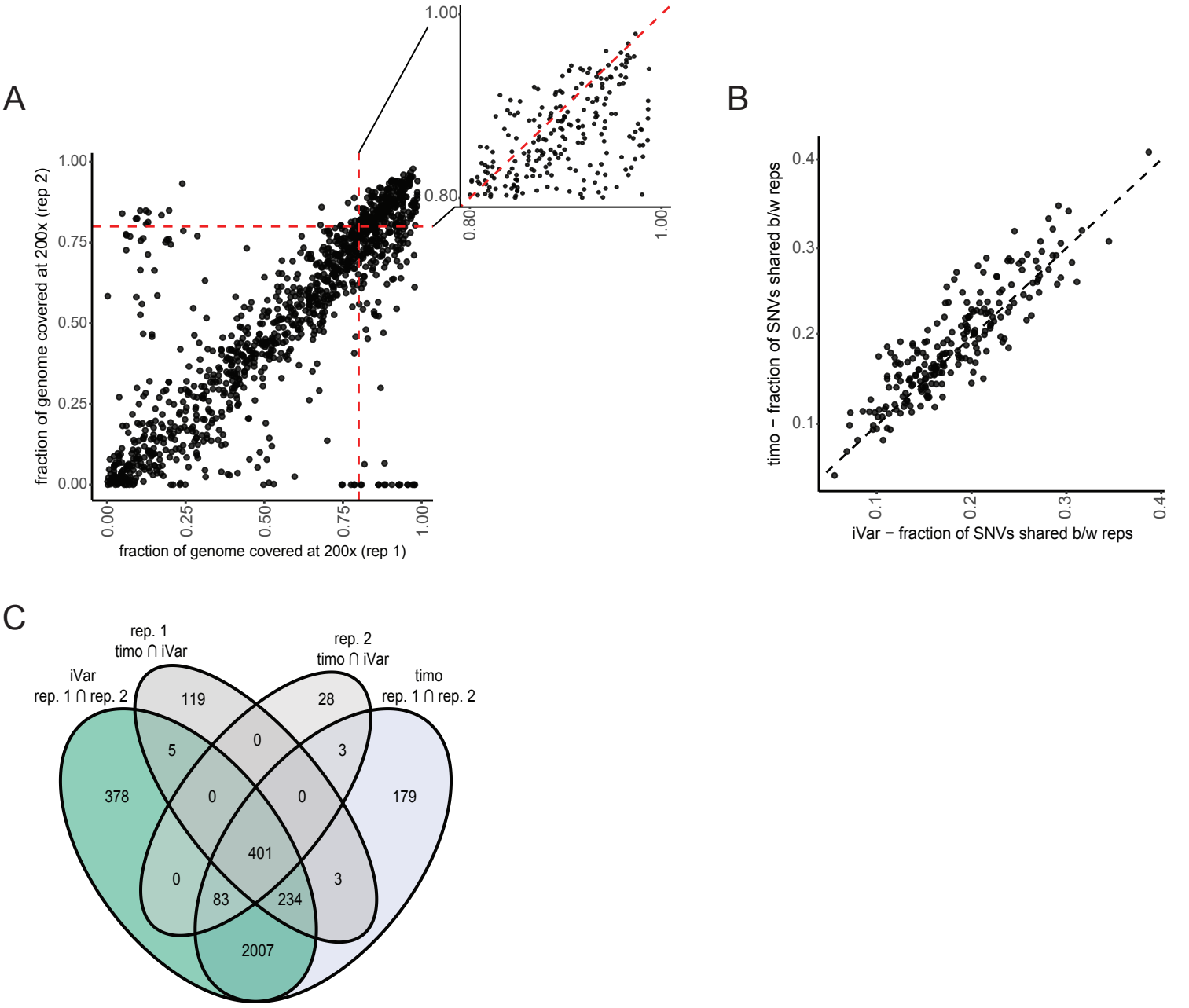

Figure 7. Reproducibility of minority variants across sequencing replicates from the Tonkin-Hill et al., 2021 sequencing dataset.

(A) The fraction of the SARS-CoV-2 genome covered at 200x read depth for replicate 1 (x-axis) and replicate 2 (y-axis) sequencing data outlined in Tonkin-Hill et al., 2021 (40). Dashed lines mark the 80% coverage cutoff. The figure inset highlights the 227 samples used for SNV analyses as they had 80% of the genome covered at 200x in both replicates.
(B) A dot plot representing the fraction of SNVs shared between the sequencing replicates (# shared/ total called) for each of the 227 samples when using timo (y-axis) or iVar (x-axis) to identify SNVs. The diagonal line represents the x=y axis.
(C) A Venn diagram representing the number of single nucleotide variants that overlap when keeping the intersection of SNVs from sequencing replicates using iVar outputs (green circle, iVar rep. 1 ∩ rep. 2), timo outputs (blue circle, timo rep. 1∩rep. 2), or timo and iVar outputs of single replicate data from either the first (rep. 1 timo ∩ iVar) or second (rep. 2 timo ∩iVar) replicate. SNV counts are calculated across the 227 samples used for analyses (outlined in **(A)**).