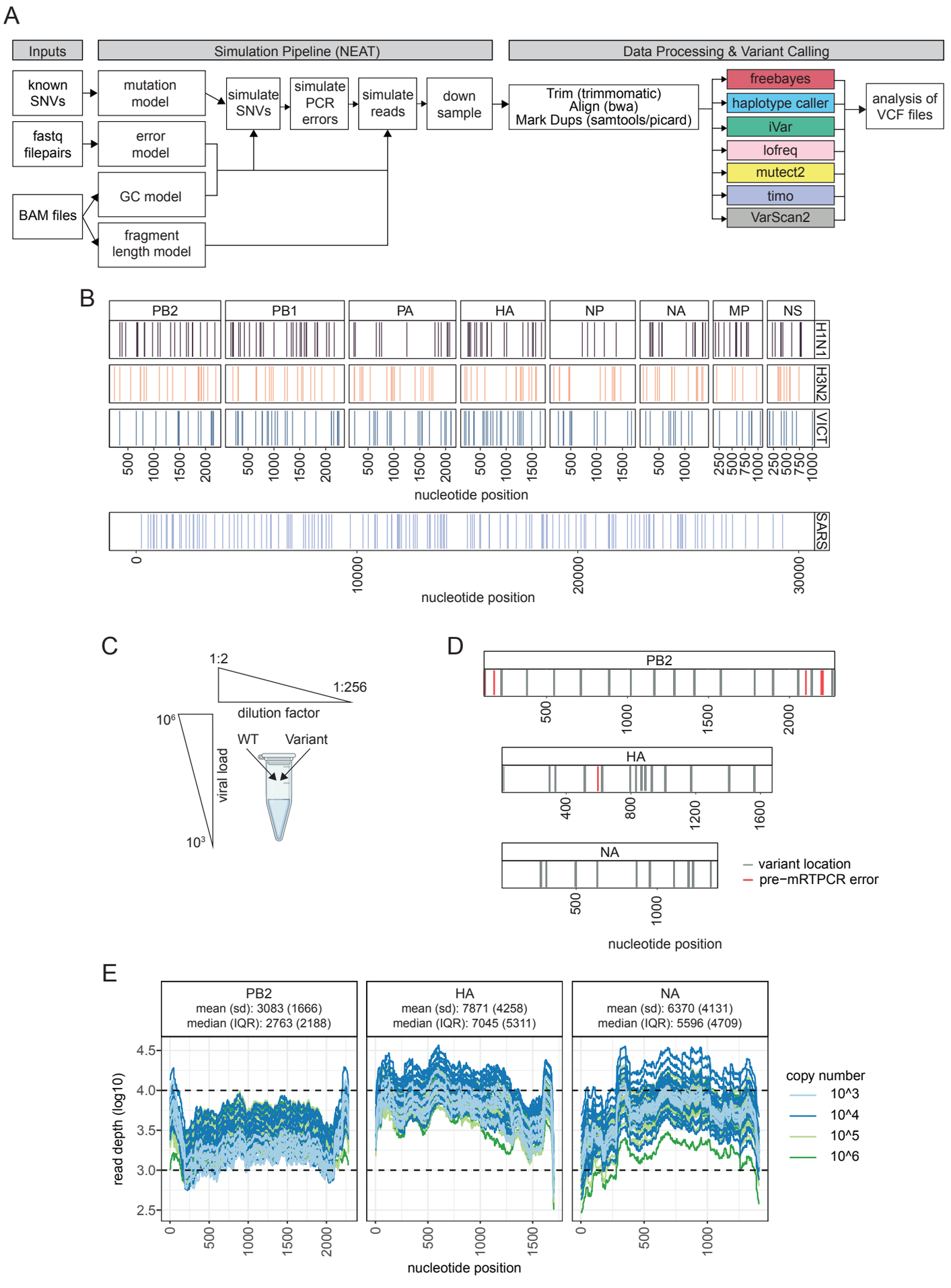

Figure S1. Experimental setup of simulated and synthetic data generation. (A) Schematic of the SNV simulation pipeline. **(B)** Nucleotide positions of SNVs in simulated data across the genomes of A/H1N1 (n=121), A/H3N2 (n=110), B/Victoria (n=118), and SARS-CoV-2 (n=144). Gene segments for influenza A and B strains are ordered largest (PB2) to smallest (NS), left to right. **(C)** Schematic dilution factor and viral loads used for mixing WT and variant RNA for library preparation and sequencing. **(D)** Location of synthetic SNVs across the three influenza gene segments. Gray lines represent the designed SNVs (PB2 (n=18), HA (n=14), NA (n=14)) and red lines represent the pre-mRT-PCR errors, likely generated during template preparation or in vitro-transcription. **(E)** Coverage plots showing log10 read depth for each synthetic influenza control across the PB2, HA, and NA gene segments. The mean, standard deviation (sd), median, and interquartile range (IQR), are represented for each segment and were calculated across the different synthetic mixtures. Color represents copy number (10^3-10^6) of all sequenced synthetic samples.