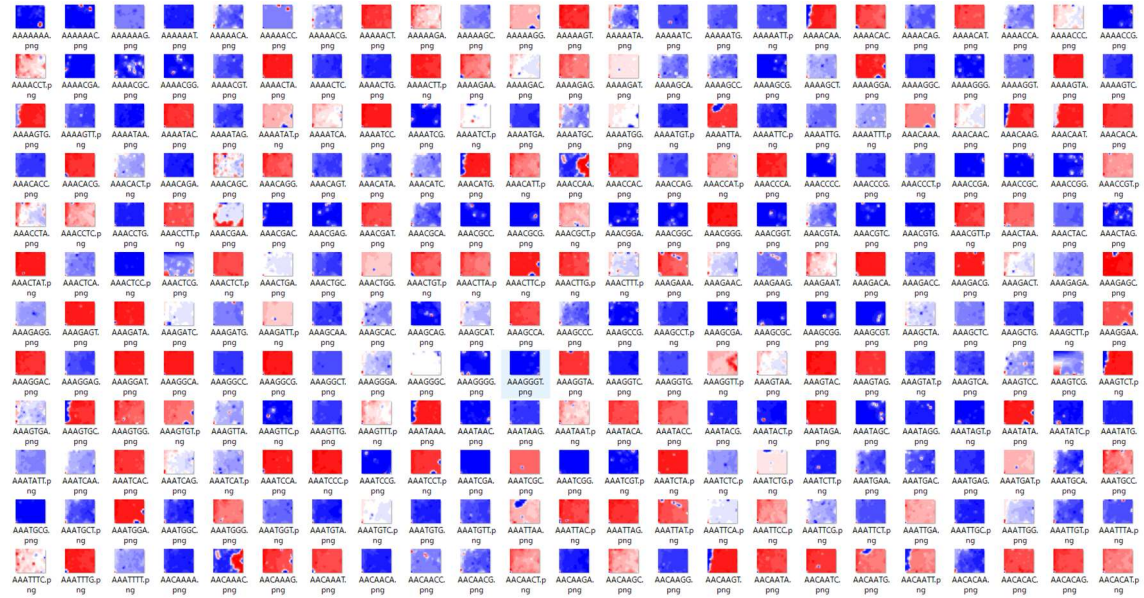


Supplementary Fig. S1. Heatmaps of a portion of 7-mers for the BLSOM of 16,384 types of 7-mers presented in Fig. 2a. The choice of the “T” base was made for compatibility reasons with the NCBI reference genome notation, while the actual RNA base in the SARS-CoV-2 genome is a “U”.



Supplementary Fig. S3. Pre- and postmutation sequences of 29-mers with the mutation listed in Table 1: 31-mers for M7. The mutated bases are shown in parentheses; N1 is not a mutation. The 15-mers used for time-series analyses were derived from these sequences.

M1 : GCACAUCUAGGUUUC (U) GUCCGGGUGUGACC
M2 : AUGUAUUGUUCUUUC (U) UACCCUCCAGAUGA
M3 : UACAGUGUCCCCACC (U) UACAAGUUUUGGAC
M4 : UGUUCUUUAUCAGGA (G) UGUUAACUGCACAG
M5 : CUUGCUGUUUUUCAG (U) AGCGCUUCCAAAAU
M6 : AAAGAAAUUUGACAC (U) CUUCA AUGGGGAAU
M7 : UCCAGGCAGCAGUAGGG (AAC) GAACUUCUCCUGCU
M8 : UCAUGACUAGGUGUC (U) UAGCUGUCCACGAG
M9 : UUGUUCUUUUUUUUG (U) UAUGAAA AUGCCUU
M10 : UACAGUUUCCUGUUU (C) ACCUUUUAACAAUUA
M11 : GACACAUGGUUUAGC (U) CAGCGUGGUGGUAG
M12 : AGCGAUUAUGACUAC (U) UAUCGUUAUAAUCU
M13 : UAGUCGCAACAGUUC (U) AAGAAAUCAACUC
N1 : ACAGUUUAUGAUCCU (NN) UUGCAACCUGAAU

Supplementary Fig. S4. BLSOM of 2087 types of 15-mers.



The number of nodes was set as described in Fig. 2a. Nodes are colored as described in Fig. 1.