Supplemental Tables

Name	Original length	Length in MSA
NSP1	539	539
NSP2	1913	1913
NSP3	5834	5521
NSP4	1499	1499
NSP5	917	917
NSP6	869	869
NSP7	248	248
NSP8	593	593
NSP9	338	338
NSP10	416	416
NSP11	38	38
NSP12	2794	2794
NSP13	1802	1802
NSP14	1580	1580
NSP15	1037	1037
NSP16	893	893
S	3821	2341
ORF3a	827	830
E	227	230
M	668	668
ORF6	185	192
ORF7a	365	368
ORF7b	131	134
ORF8	365	21
N	1259	1268
ORF10	116	116

Table 1: Overview of the genes for which trees were built, including their name/symbol, original length in the SARS-CoV-2 genome and the length of the corresponding sequence in the multiple sequence alignment. Order corresponds to their order in the SARS-CoV-2 genome. Note that all non-structural proteins (NSP's) together form the ORF-1a and ORF-1ab proteins, but their sequences are used as separate 'genes' here.

Breakpoints	Selection	Edge contraction thresholds					
		None None	None s=70	None s=90	l=0.01 None	l=0.01 s=70	l=0.01 s=90
Blocks	A (n=12)	9	9	9	9	9	9
	A- (n=11)	8	7	7	7	6	7
	B (n=9)	9	7	7	8	7	7
	B- (n=8)	5	5	5	5	5	4
	C (n=7)	5	5	4	5	5	4
	C- (n=6)	1	1	1	1	1	1
Genes	A (n=12)	23	22	21	19	19	20
	A- (n=11)	21	17	17	16	15	12
	B (n=9)	20	22	19	18	19	18
	B- (n=8)	12	9	9	10	8	8
	C (n=7)	10	9	9	9	8	8
	C- (n=6)	1	1	1	1	1	1

Table 2: Unique trees for different breakpoint location sets, taxon selections (n = number of taxa) and thresholds for edge contraction based on branch length (l) and BS support values (s). "–" indicates that no solution was found within the runtime limit of 5 minutes.

Breakpoints	Selection	Edge contraction thresholds					
		None None	None s=70	None s=90	l=0.01 None	l=0.01 s=70	l=0.01 s=90
Blocks	A (n=12)	2 1 1 0 0	4 2 1 1 1	2 0 0 0 0	5 5 6 5 6	4 2 3 2 1	1 1 1 0 0
	A- (n=11)	0 0 0 0 1	1 0 0 0 1	1 1 0 0 1	1 1 0 5 4	1 1 1 1 0	0 1 0 0 0
	B (n=9)	8 8 8 8 7	9 9 9 5 7	3 2 2 2 2	4 4 3 4 5	2 2 1 2 1	9 9 2 9 9
	B- (n=8)	2 1 1 1 0	3 3 3 0 0	5 5 6 5 5	1 1 1 0 1	1 1 0 1 0	1 1 1 1 0
	C (n=7)	6 6 6 1 1	4 4 2 4 4	11 11 11 8 9	1 3 2 2 1	5 5 9 8 11	4 2 2 2 1
	C- (n=6)	1 2 0 1 1	2 2 1 0 1	1 1 0 1 1	2 2 2 2 2	2 1 1 0 0	2 2 2 2 1
Genes	A (n=12)	7 6 0 0 0	1 3 3 1 0	5 5 3 4 3	2 2 3 3 2	3 2 2 3 3	0 1 8 0 1
	A- (n=11)	1 1 2 2 1	5 5 5 4 1	2 1 1 1 0	0 2 3 3 3	6 6 6 6 8	0 0 0 1 1
	B (n=9)	9 9 9 9 2	2 2 2 2 2	1 2 6 1 3	3 3 3 3 3	7 8 3 3 3	5 4 4 4 4
	B- (n=8)	0 0 0 2 3	4 1 1 1 0	5 2 5 2 3	1 1 1 0 1	1 1 1 1 2	1 0 1 0 0
	C (n=7)	1 2 1 3 1	5 5 5 1 1	2 6 2 3 2	7 6 6 6 2	3 8 8 5 8	3 3 3 4 4
	C- (n=6)	1 1 1 0 2	2 2 2 0 1	3 2 1 3 0	5 8 7 7 7	1 1 1 1 0	6 1 1 0 1

Table 3: Reticulation numbers of each of the five networks constructed by the Maximum Pseudo-Likelihood algorithm for different breakpoint location sets, taxon selections (n = number of taxa) and thresholds for edge contraction based on branch length (l) and BS support values (s).