Tutorial-2

By 2018101075

1. Multiple Sequence Alignment
   1.)
   Multiple Sequence Alignment of Protein Sequences
   Identity Matrix

SARS CoV	100.00	77.30	77.68	77.22	32.91	32.19
PCoV	77.30	100.00	92.98	92.43	32.62	31.57
BatCoV RaTG13	77.68	92.98	100.00	97.71	32.95	32.15
SARS CoV2	77.22	92.43	97.71	100.00	33.20	32.05
BatCoV	32.91	32.62	32.95	33.20	100.00	65.54
MERS CoV	32.19	31.57	32.15	32.05	65.54	100.00

Multiple Sequence Alignment of Gene Sequences
Identity Matrix

SARS CoV	100.00	79.05	79.41	79.60	53.47	53.44
PCoV	79.05	100.00	85.26	85.44	53.57	53.76

BatCoV RaTG13	79.41	85.26	100.00	96.18	53.79	53.63
SARS CoV2	79.60	85.44	96.18	100.00	54.06	53.71
BatCoV	53.47	53.57	53.79	54.06	100.00	83.57
MERS CoV	53.44	53.76	53.63	53.71	83.57	100.00

Closest Relative according to Identity Matrix

	Closest Relative
SARS-CoV	Bat-Cov RaTG13, SARS-CoV2
P-Cov	Bat-Cov RaTG13, SARS-CoV2
Bat-CoV RaTG13	SARS-CoV2
SARS-CoV2	Bat-Cov RaTG13
Bat-CoV	MERS-CoV
MERS-CoV	Bat-CoV

2.) Possible source of origin for SARS-CoV2 seems to be Bat-Cov RaTG13 as protein identity matrix gives 97.71 and gene identity matrix gives 96.18 identity percentage.

Possible source of origin for MERS-CoV seems to be Bat-CoV as protein identity matrix gives 65.54 and gene identity matrix gives 83.57 identity percentage, but due to less percentage, source of origin is not very clear.

2.

1.) For the proteins, parsimony tree has 4 levels, while other have 3, but for dna, they have same topology-wise

2.) They both are giving more or less the same tree. Bootstrapping gives us extra details like percentage of tree with the same link, and better graphs and more confidence.

3.) Yes, the inferences are in agreement with those in Q1b). All the trees show that there is a close link between (SARS-CoV2 and BatCoVRaTG14), (BatCoV and MERSCoV).