SARS CoV-2 Spike

Based on the likelihood ratio test, *episodic diversifying selection* has acted on **6** sites in this dataset (p≤0.1).

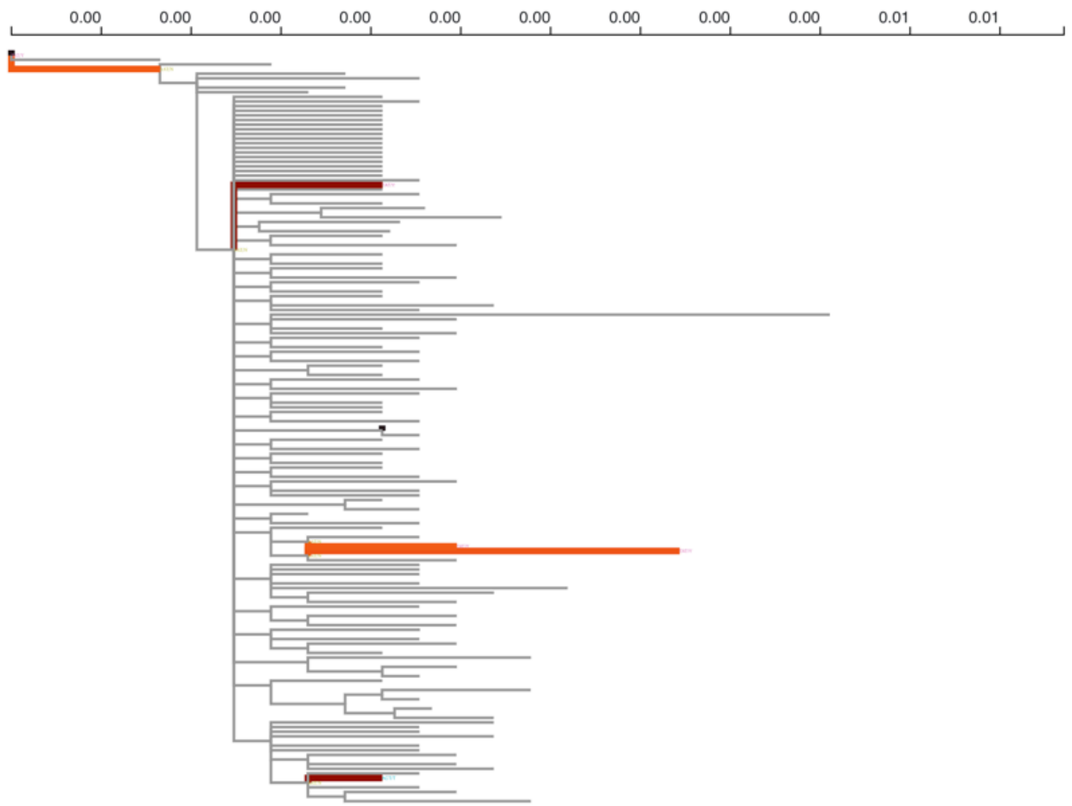
MEME analysis (v3.0) was performed on the alignment from /Users/sergei/Dropbox/Talks/VEME-current/data/spike.fas using HyPhy v2.5.40.

p-value threshold 0.1 Update

Suggested citation: Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLoS Genet* 8(7): e1002764.

118 sequences in the alignment	**1273** codon sites in the alignment
1 partitions	**162** median branches/partition used for testing
N/A bootstrap replicates	**6** sites subject to episodic diversifying selection
3.00 median branches with support for selection/selected site	**3** sites with variable ω across branches

Codon	α	β⁻	p⁻	β⁺	p⁺	LRT	p-value	# branches under selection	MEME LogL	FEL LogL
1	1.213	0.543	0.994	2,732.12	0.006	9.76	0.003	1	-19.143	-14.262
1,243	0	0	0.983	552.73	0.017	8.846	0.005	2	-23.278	-19.431
452	0	0	0.01	14.48	0.99	8.771	0.005	5	-36.485	-36.486
470	3.168	0.727	0.994	10,000	0.006	8.012	0.008	1	-24.317	-19.832
501	0.004	0.002	0.925	343.752	0.075	5.017	0.037	5	-37.067	-36.202
157	0	0	0.01	7.313	0.99	3.708	0.074	4	-29.546	-29.547




```
hyphy meme --alignment data/spike.fas --tree data/spike.tree
```

Interpreting dN/dS for intra-host and intra-species pathogen

- **dN/dS** can be estimated for all sorts of sequence data (e.g., it has been done for cancer SNP data)
- Traditional interpretation of dN/dS is based on the assumption that **substitution ~ fixation**
- Not the same for intra-species / intra-host pathogens
  - Much of variation is due to polymorphism, or even dead-end mutations
- This is because selection has not had a chance to “filter” mutations (except for patently deleterious ones)
- This often manifests as differences in selective “regimes” between tips and internal branches