Gene-wide selection analysis using a branch-site method (BUSTED), SARS-CoV-2 spike
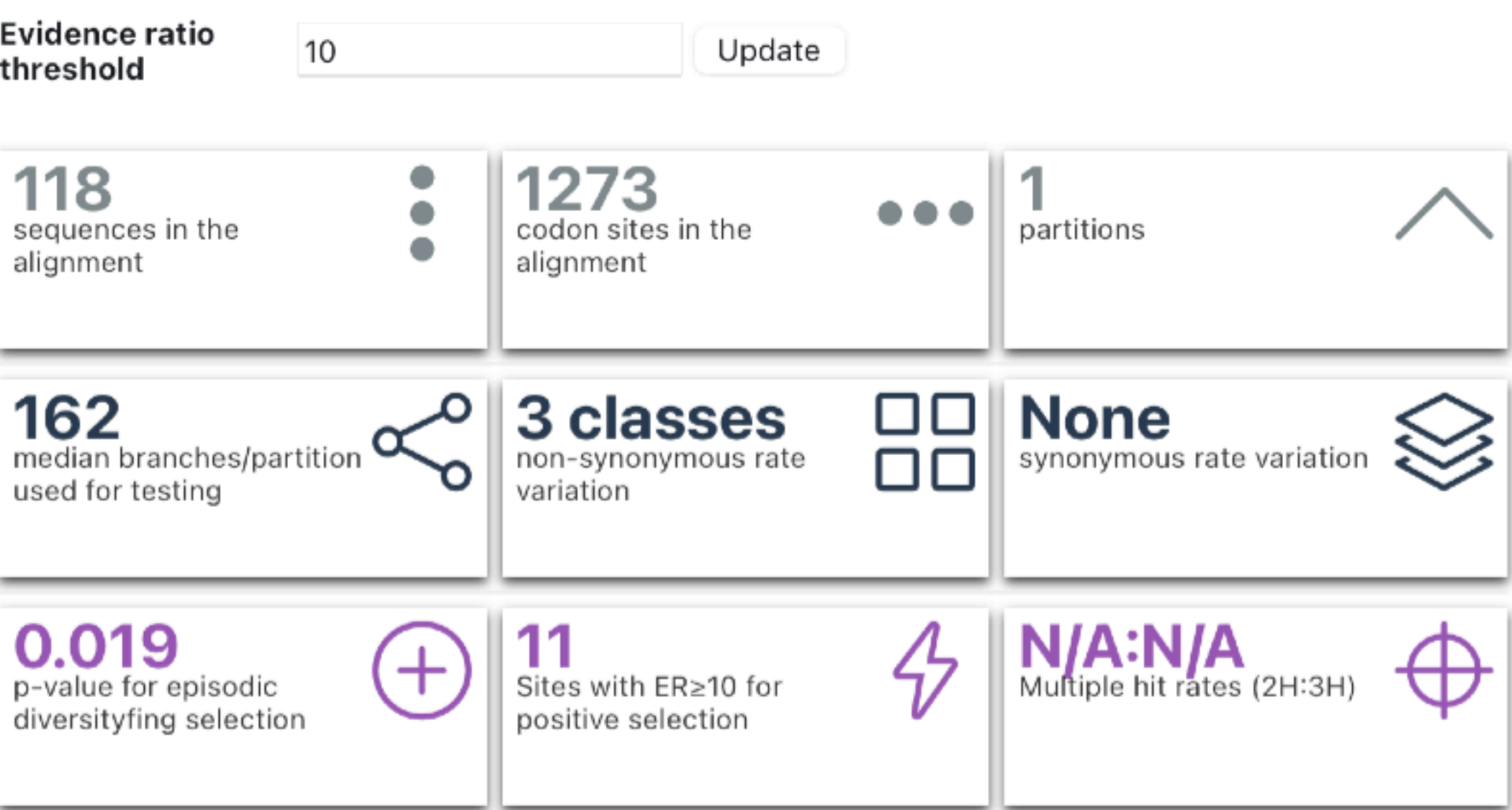
```
hyphy busted --srv No --alignment
data/spike.fas --tree data/
spike.tree --starting-points 5
```

Produces *spike.fas.BUSTED.json* file
View in http://vision.hyphy.org/BUSTED

Based on the likelihood ratio test, there **is** evidence of *episodic diversifying selection* in this dataset (p=0.01901).
BUSTED analysis (v4.0) was performed on the alignment from /Users/sergei/Dropbox/Talks/VEME-current/data/spike.fas using HyPhy v2.5.40.
This analysis **did not include** site-to-site synonymous rate variation.

Suggested citation: *Gene-wide identification of episodic selection*, Mol Biol Evol. 32(5):1365–71, *Synonymous Site-to-Site Substitution Rate Variation Dramatically Inflates False Positive Rates of Selection Analyses: Ignore at Your Own Peril*, Mol Biol Evol. 37(8):2430–2439

Evidence ratio threshold [10] [Update]

118 sequences in the alignment	1273 codon sites in the alignment	1 partitions
162 median branches/partition used for testing	3 classes non-synonymous rate variation	None synonymous rate variation
0.019 p-value for episodic diversityfing selection	11 Sites with ER≥10 for positive selection	N/A:N/A Multiple hit rates (2H:3H)

Alignment-wide results

Model	Log (L)	AIC-c	Params.	Rate distribution	Rate plot
Unconstrained model	−9287.35	18937.1	181	Tested ω 0.000 (0.45262%) 0.2561 (97.532%) 14.88 (2.0156%) Mean = 0.5497, CoV = 3.740	
Constrained model	−9290.62	18941.7	180	Tested ω 0.000 (46.919%) 1.000 (2.2974%) 1.000 (50.784%) Mean = 0.5308, CoV = 0.9402	

78

BUSTED analysis - summary

- **West Nile Virus NS3 protein**

  - No statistical support for selection; ML point estimate allocates a small proportion of sites (`~1%`) to the selected group (`dN/dS ~ 2`)

  - The rest of the gene is very strongly conserved (`dN/dS = 0.004`)

- **HIV-1 transmission pair**

  - Very strong evidence of strong episodic diversification (`dN/dS ~ 100`) on a small proportion of sites (`2%`)

  - The rest of the gene evolves with weak purifying selection (`dN/dS = 0.6-0.7`)

- **SARS-CoV-2 spike**

  - Evidence of episodic diversification (`dN/dS ~ 15`) on a small proportion of sites (`~2%`)

  - Most of the rest of the gene evolves with purifying selection (`dN/dS = 0.2`)