

More accurate testing via parametric bootstrap

- P-values for MEME/FEL etc are derived from asymptotic approximations (large N)
- Not clear how well these hold for smaller and low-divergence datasets
- Can use a **much slower** simulation based method to derive more accurate p-values at each site
- Can result both in improved power and lower rates of false positives

FEL on internal branches of Spike finds most selected sites

Codon	Partition	alpha	beta	LRT	Selection detected?
1	1	10000.000	0.075	1.004	Neg. p = 0.0099
5	1	0.000	19.044	2.890	Pos. p = 0.0297
12	1	0.000	20.323	2.989	Pos. p = 0.0297
13	1	0.000	23.877	2.589	Pos. p = 0.0396
18	1	0.000	19.110	2.885	Pos. p = 0.0396
49	1	0.000	22.768	2.519	Pos. p = 0.0594
69	1	0.000	22.573	2.464	Pos. p = 0.0396
138	1	0.000	26.821	2.736	Pos. p = 0.0693
177	1	10000.000	0.000	1.187	Neg. p = 0.0099
245	1	0.000	22.499	2.501	Pos. p = 0.0792
367	1	0.000	44.535	9.044	Pos. p = 0.0099
439	1	0.000	34.481	4.986	Pos. p = 0.0099
452	1	0.000	30.321	5.518	Pos. p = 0.0099
477	1	0.000	23.678	4.325	Pos. p = 0.0297
501	1	0.000	38.373	3.317	Pos. p = 0.0297
570	1	0.000	21.039	3.047	Pos. p = 0.0594
614	1	0.000	22.215	3.099	Pos. p = 0.0594
681	1	0.000	18.344	2.819	Pos. p = 0.0297
701	1	0.000	20.245	1.998	Pos. p = 0.0792
716	1	0.000	25.450	2.349	Pos. p = 0.0792
769	1	0.000	28.107	2.470	Pos. p = 0.0693
854	1	0.000	25.908	2.506	Pos. p = 0.0495
941	1	0.000	25.251	2.344	Pos. p = 0.0891
1118	1	3.081	27.290	2.023	Pos. p = 0.0495
1176	1	0.000	22.141	3.039	Pos. p = 0.0297
1237	1	10000.000	0.151	0.876	Neg. p = 0.0099
1248	1	0.000	20.075	2.336	Pos. p = 0.0891

CAUTION: A VERY TIME CONSUMING ANALYSIS (SEVERAL HOURS)

```
hyphy fel --alignment data/spike.fas --tree data/spike.tree --branches Internal --
output Spike-pbs.FEL.json --resample 100
```