Gene-wide selection
random effects over sites and branches [BUSTED]

Sites

Branches

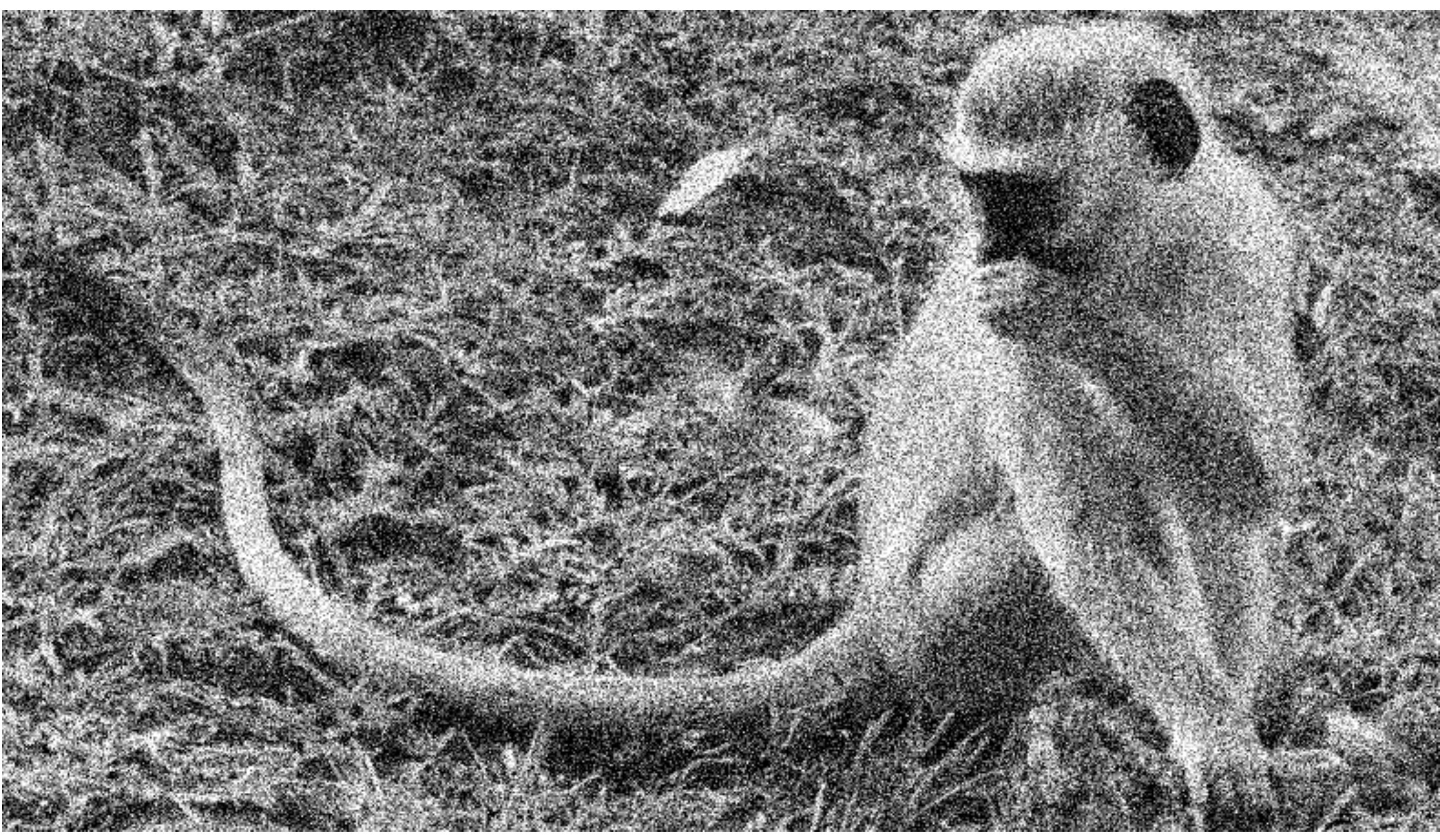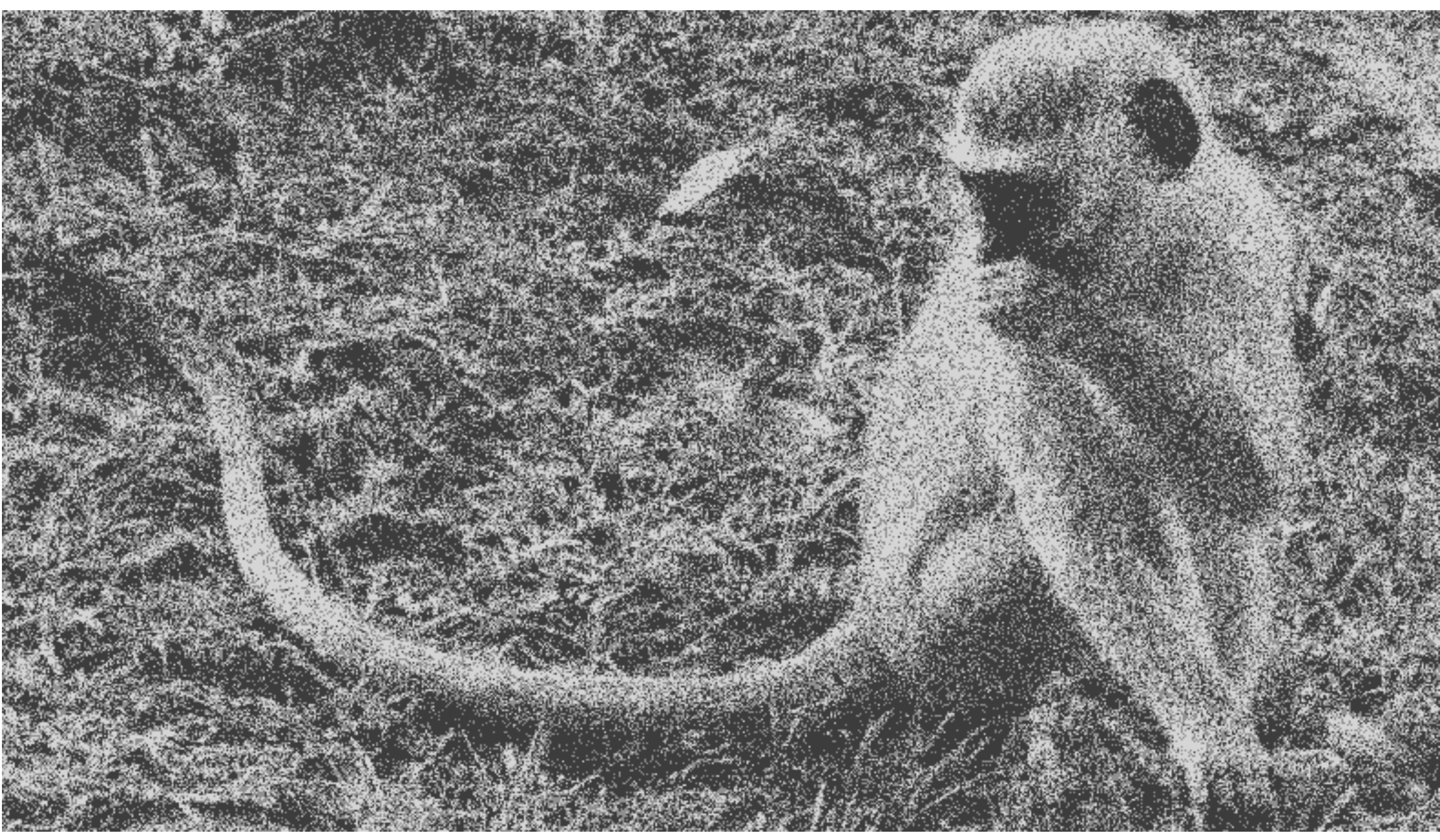Is there enough **image area** that is sufficiently bright; allow each pixel to be one of K (=3) colors, chosen adaptively, e.g. to minimize perceptual differences

[BUSTED]: each branch-site combination is a drawn from a K-bin (dS,dN) distribution. The distribution is estimated from the entire alignment. Tests if dN/dS>1 for some branch/site pairs in the alignment

Gene-wide selection
random effects over sites and branches [BUSTED]

Branches (y-axis)

Sites (x-axis)

Is there enough **image area** that is sufficiently bright; allow each pixel to be one of K (=3) colors, chosen adaptively, e.g. to minimize perceptual differences

[BUSTED]: each branch-site combination is a drawn from a K-bin (dS,dN) distribution. The distribution is estimated from the entire alignment. Tests if dN/dS>1 for some branch/site pairs in the alignment
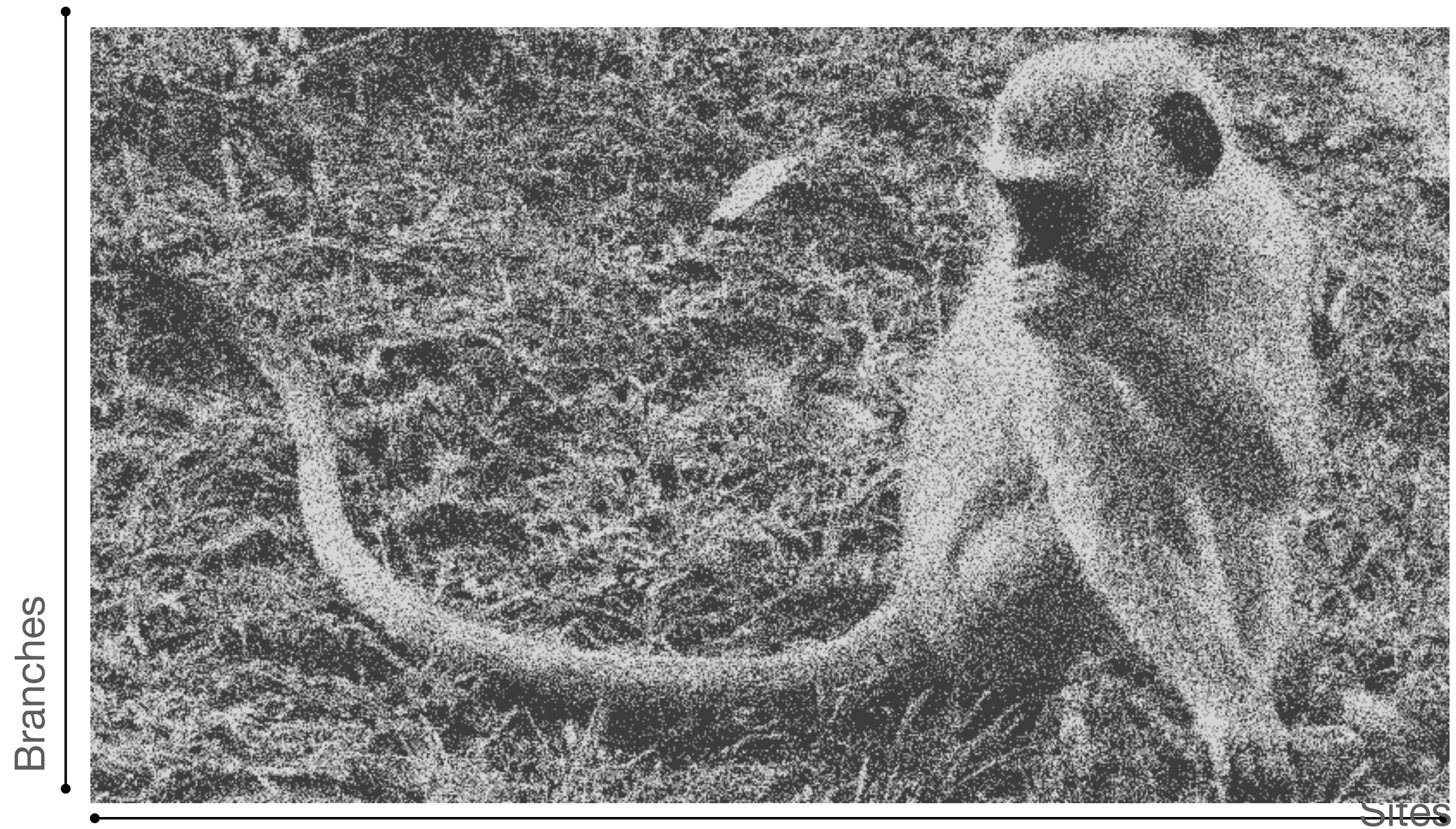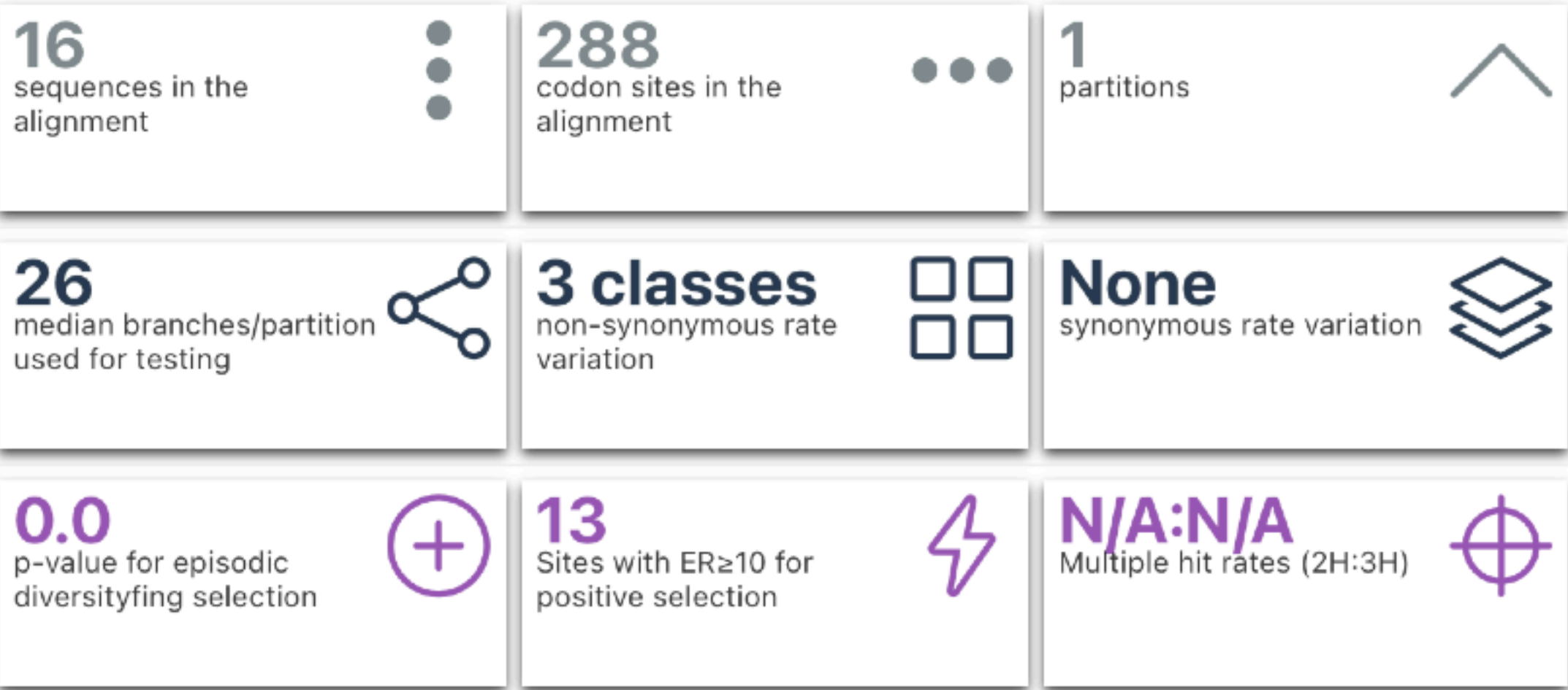
59

Based on the likelihood ratio test, there **is** evidence of *episodic diversifying selection* in this dataset (p=0.000).

BUSTED analysis (v4.0) was performed on the alignment from /Users/sergei/Dropbox/Talks/VEME-current/data/HIV-sets.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

Gene-wide selection analysis using a branch-site method (BUSTED), HIV-1 *env*

```
hyphy busted --srv No --alignment
data/HIV-sets.nex --starting-
points 5
```

Produces *HIV-sets.nex.BUSTED.json* file
View in http://vision.hyphy.org/BUSTED or
https://observablehq.com/@spond/busted

Evidence ratio threshold [10] [Update]

16 sequences in the alignment

288 codon sites in the alignment

1 partitions

26 median branches/partition used for testing

3 classes non-synonymous rate variation

None synonymous rate variation

0.0 p-value for episodic diversityfing selection

13 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	−2039.96	4170.83	45	Tested ω 0.5596 (86.941%) 0.9885 (10.960%) 96.09 (2.0981%) Mean = 2.611, CoV = 5.242	
Constrained model	−2078.31	4245.48	44	Tested ω 1.000 (14.819%) 1.000 (20.229%) 1.000 (64.952%) Mean = 1.000, CoV = NaN	

60