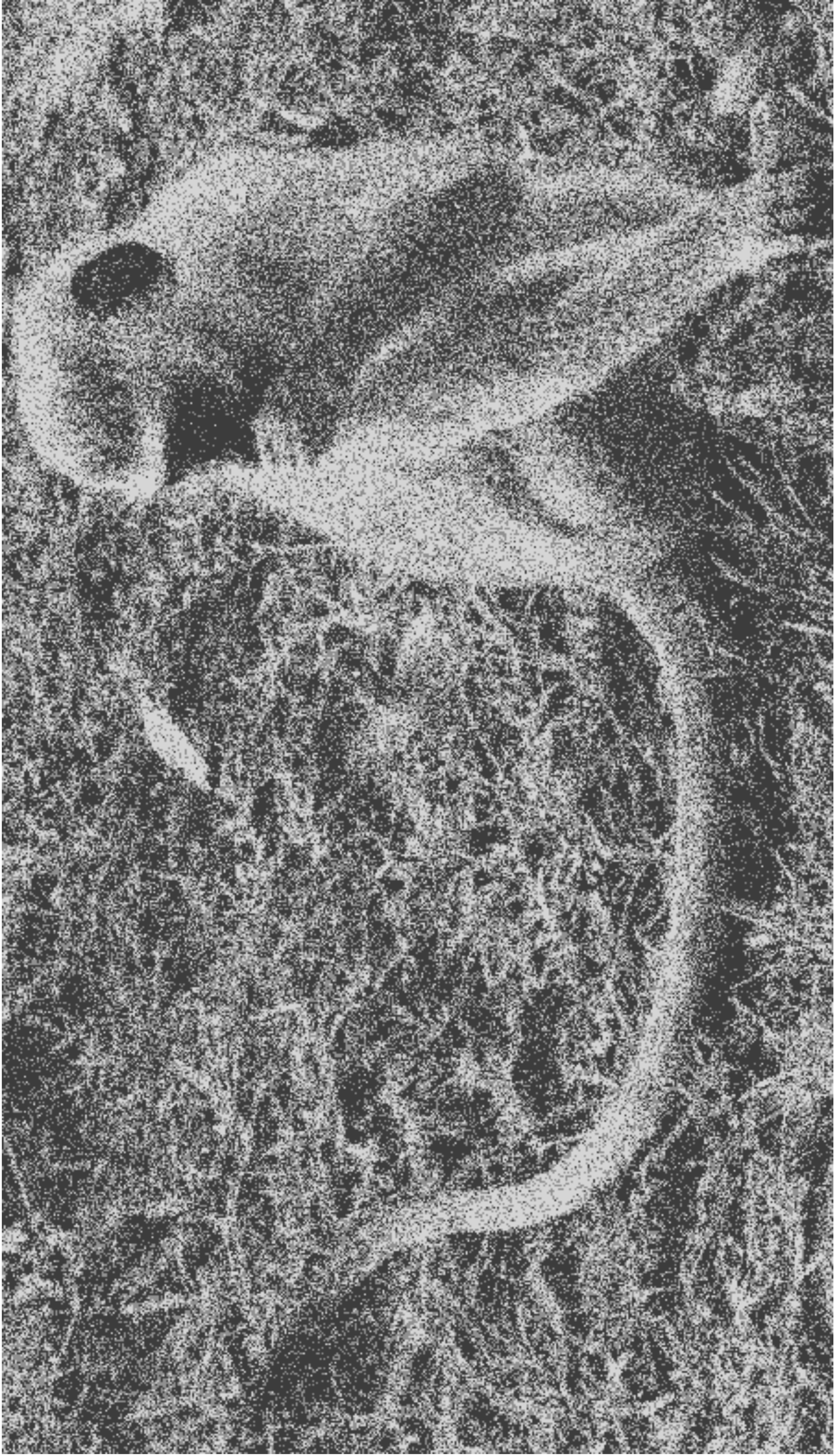
Gene-wide selection - random effects over sites and branches [BUSTED]


Sites

Branches

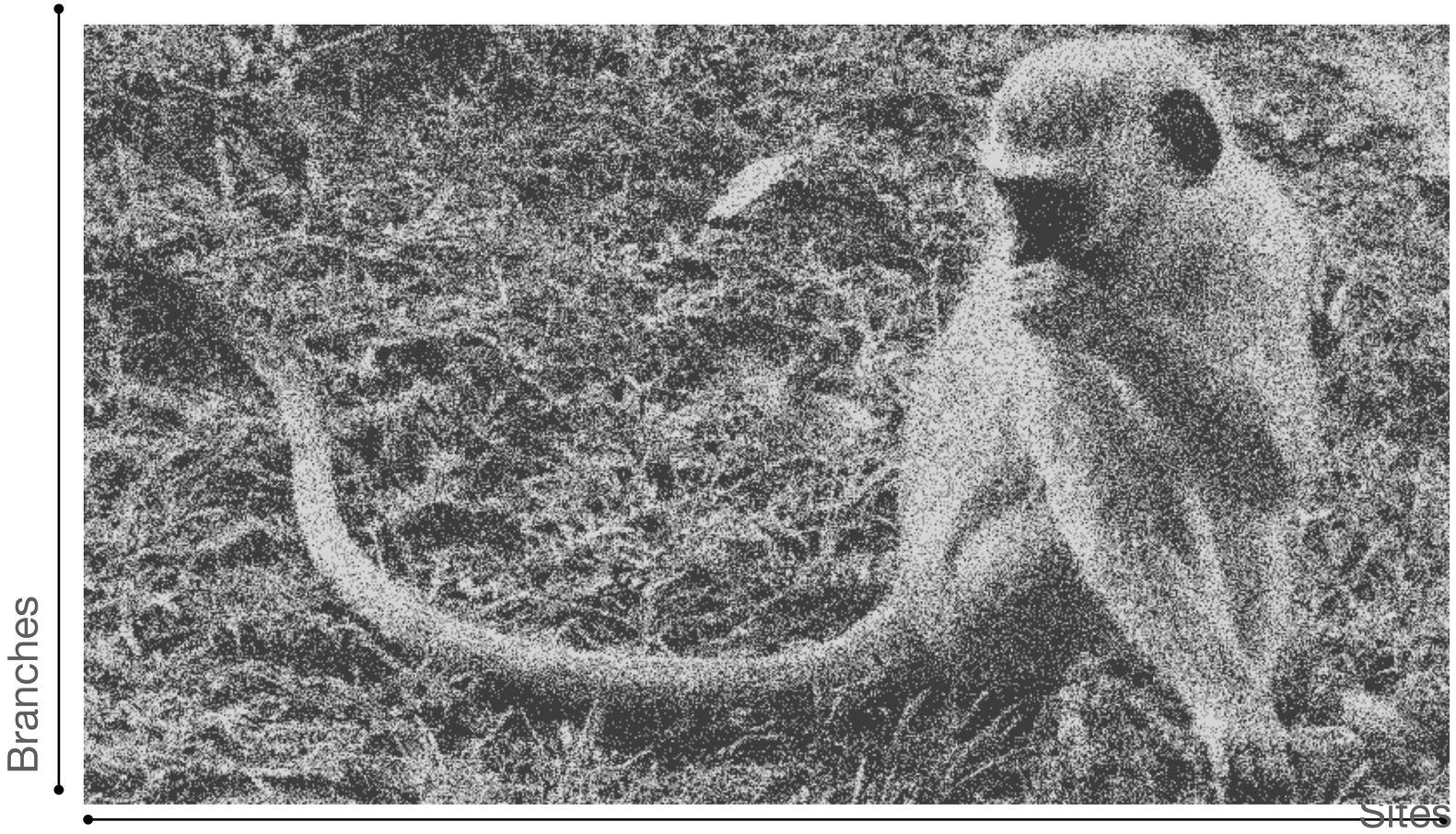
Is there enough **image area** that is sufficiently bright; allow each pixel to be one of K (=3) colors, chosen adaptively, e.g. to minimize perceptual differences

[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



71

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



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

BUSTED inference

- **Because BUSTED is a random-effects method, it pools information across multiple sites and branches to gain power**
- The cost to this pooling is lack of site-level **resolution**, i.e., it is not immediately obvious which sites and/or branches drive the signal
- Standard ways to extract individual site contributions to the overall signal is to perform a post-hoc analysis, such as empirical Bayes, or “category loading”
- For BUSTED, “category loading” is faster and experimentally better
- Can also compute exploratory evidence for selection support along individual branches at specific sites

Murrell et al | Mol. Biol. Evol | 32(5) | 1365–1371