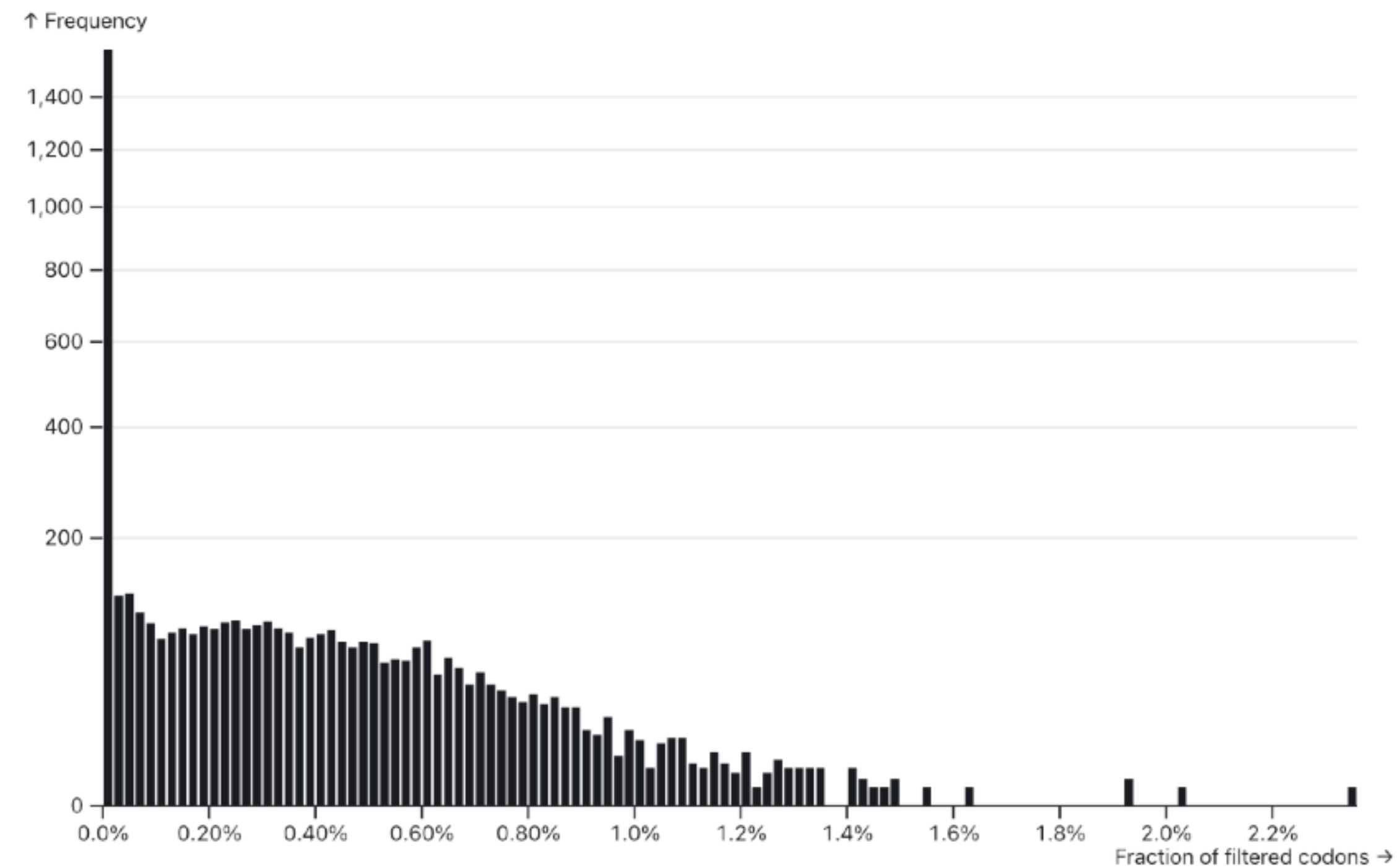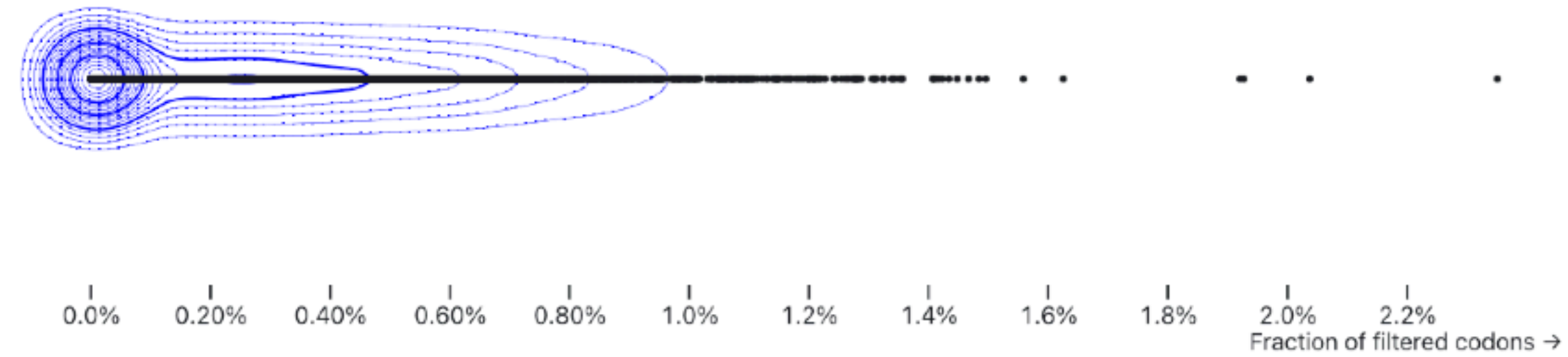- The distribution of per MSA filtered codon counts

Fractions of filtered sites (of the entire alignment character count, seqs x codons)

64-taxon Zoonomia tree
Codons filtered / 1000 codons

- dN/dS methods can be used to both handle some residual MSA error during selection screens and also find it in alignments

- The structure of the underlying substitution model can be adjusted to reflect more biological realism (e.g. MNM)

- Can identify specific genes and/or species which have higher putative errors

- Implemented in the HyPhy package

- Not computationally cheap, but delivers a useful result while screening for errors

  - The main cost growth dimension is the # of genomes

  - For 32 genomes, per MSA run time is on the order of 1-10 minutes per core