

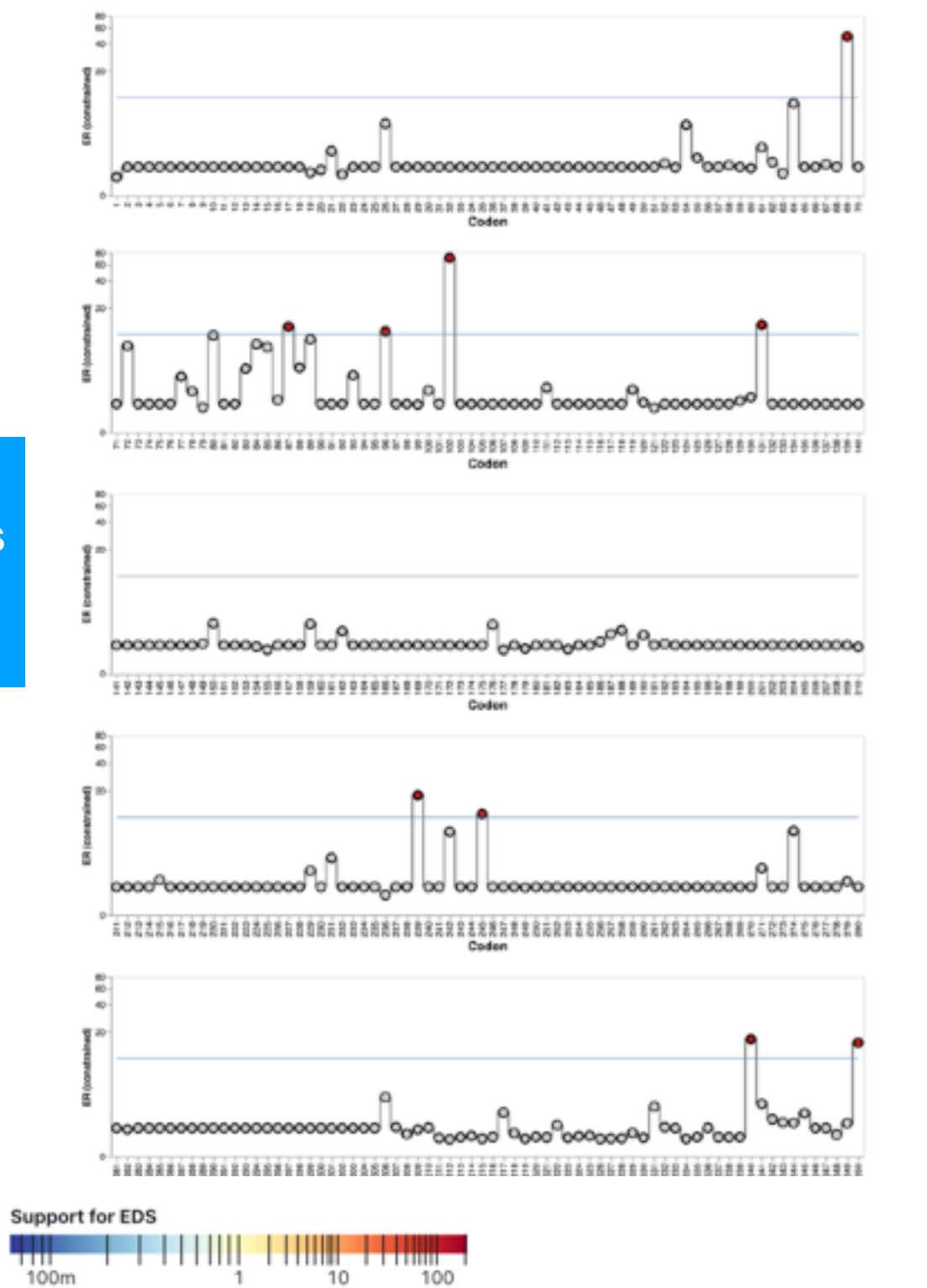
REAL

IQCF1

$\omega_1=0.4085$ (79.597%)
 $\omega_2=0.4078$ (15.784%)
 $\omega_3=11.16$ (4.6197%)

Reasonable ω value
with broad support

Nicely
dispersed sites
with evidence
of selection



VS_HLNYCCOU1	GCA	GTT	ACT	CTG	CAG	TCC	TGG	GCC	CGC	ATG	TGG	TTC	ATC	CGT	CGG	CGC	TAC
VS_HLTRAJAV1	GCA	GTC	AGG	CTG	CAG	TCT	TGG	GTC	CGC	ATG	TGG	CGC	ATC	CGT	CGG	CGC	TAC
VS_HLNOCLEP1	GCG	GTC	CGG	CTG	CAG	TCC	TGG	GTC	CGC	ATG	TGG	CGC	ATC	CGC	CTG	CGC	TAC
VS_HLCRATHO1	GCG	GCG	CGG	CTG	CAG	AGC	TGG	CTG	CGC	ATG	TGG	CGC	GCG	CGG	CGG	CGC	TAC
VS_HLUROGRA1	GTG	GTC	AGG	CTT	CAG	TCC	TGG	GTC	CGC	ATG	TGG	TGC	ATC	---	---	---	---
VS_HLSOLPAR1	GCG	GTC	AGG	TTG	CAG	TCC	TGG	GTC	CGC	ATG	TGG	CGC	ATC	CGT	AGG	CGC	TAC
VS_ERIEUR2	GCG	GTC	AGC	CTG	CAG	TCC	TGG	GTC	CGC	ACG	TGG	CGC	ATT	CGG	AGG	CGC	TAC
VS_SORARA2	GCA	GTC	AAG	CTG	CAG	TCC	TGG	GTC	CGA	ATG	TGG	CGC	ATC	CGT	AAG	CGC	TAT
VS_OCHPRI3	GCT	GTG	AGG	CTG	CAG	TCC	TGG	GCT	CGC	ATG	TGG	CGC	ATC	CGC	CGG	CGC	TAC
VS_HLPSAOBE1	GTG	GTC	AGG	TTG	CAG	TCC	TGG	GTC	CGC	ATG	TGG	CGA	ATC	CGC	AGG	CGC	TAC
VS_HLSIGHIS1	GTG	GTC	AGG	GTG	CAG	TCC	TGG	ATC	CGC	ATG	TGG	CTT	ATT	CGT	AGA	CAT	TAC
VS_HLHYSCRI1	GCA	GTC	AGA	CTG	CAG	TCC	---	GTC	CGC	ATG	TGG	TGT	GCT	CGC	TGG	CAC	TAC
VS_CHILAN1	GCG	GTC	AGA	CTG	CAG	TCC	TGT	GTC	CGC	ATG	TGG	CGT	GCT	CGC	CAG	CAC	TAC
VS_HLCTESOC1	GCG	GTC	AGA	CTG	CAG	TCC	TGG	GTC	CGC	ATG	CGG	CGT	GCT	CGC	CAG	CAC	TAC

Standard BUSTED method results, $p \leq 0.001$ for both

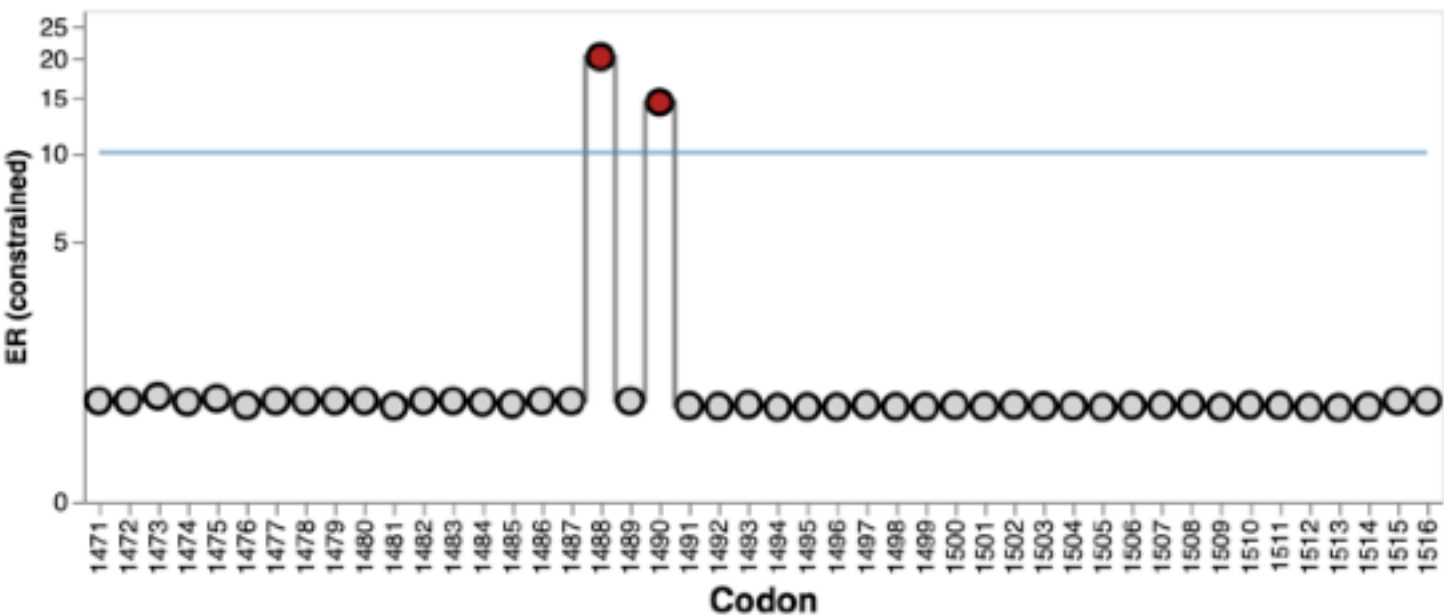
NOT REAL

KRT8

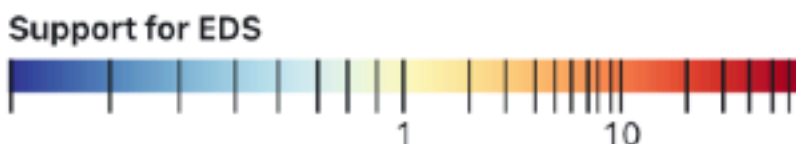
$\omega_1=0.02107$ (80.763%)
 $\omega_2=0.6961$ (18.533%)
 $\omega_3=136.3$ (0.70395%)

A large ω value with
narrow support

Evidence of
selection is
a localized
clump



Obvious alignment/
homology issues,
here in one sequence



VS_TUPBEL1	---	/	---	/	---	/	---	/	---	/	---	/	---	/	---	/	---	/	---	/		
VS_OCHPRI3	---	/	---	/	AGC/S	TCC/S	TCC/S	AGC/S	---	/	---	/	TCC/S	CGA/R	---	/	---	/	GGC/G	AGC/S	AGC/S	
VS_HLPERLONPAC1	---	/	---	/	AGC/S	TCC/S	TCA/S	AGC/S	---	/	---	/	TCT/S	AGG/R	GTG/V	GGC/G	AGT/S	GGC/G	AGC/S	AGC/S		
VS_HLHYSCRI1	---	/	---	/	AGC/S	TCC/S	TCC/S	AGT/S	---	/	---	/	TCC/S	CGA/R	---	/	---	/	GGC/G	AGC/S	AGC/S	
VS_HLCTESOC1	---	/	---	/	AGC/S	TCT/S	TCC/S	AGC/S	---	/	---	/	TCC/S	AGA/R	---	/	---	/	GGC/G	AGC/S	AGC/S	
VS_HLTRAJAV1	---	/	---	/	AGC/S	TCC/S	TCG/S	ACC/T	---	/	---	/	ACC/T	CGA/R	---	/	---	/	GGC/G	AGC/S	GGC/G	
VS_HLCRATHO1	---	/	---	/	AGC/S	TCC/S	TCG/S	TCC/S	---	/	---	/	TCC/S	CGG/R	---	/	---	/	GGC/G	AGC/S	AGC/S	
VS_HLUROGRA1	---	/	AAC/N	CGC/R	ATC/I	AGC/S	TCC/S	---	/	---	/	TCG/S	TCC/S	TTC/F	---	/	---	/	GGC/G	CGA/R	GTG/V	GGA/G
VS_HLSOLPAR1	---	/	---	/	AGC/S	TCC/S	TCC/S	TCC/S	---	/	---	/	TCC/S	CGG/R	---	/	---	/	GGC/G	AGC/S	AGC/S	
VS_ERIEUR2	---	/	---	/	AGC/S	CCC/P	TCC/S	GCC/A	---	/	---	/	TCC/S	CGG/R	---	/	---	/	GGC/G	AGC/S	AGC/S	
VS_SORARA2	---	/	---	/	AGC/S	TCC/S	TCC/S	ACC/T	---	/	---	/	ACC/T	CGG/R	---	/	---	/	GGC/G	AGC/S	ACC/T	
	10		10		10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	

- The simple fix here is to include an **explicit error component** in the model
- We simply allow a small fraction of the alignment (e.g. $\leq 1\%$) to evolve with abiologically high rates (e.g. $dN/dS \geq 100$)
- This is in addition to the standard model which allows negative,

neutral, and positive selection regimes.

- The primary goal of this analysis is to classify the selective regime on a gene (in the presence of specific types of errors)
- So error detection (and filtering) is a byproduct of an already useful analysis