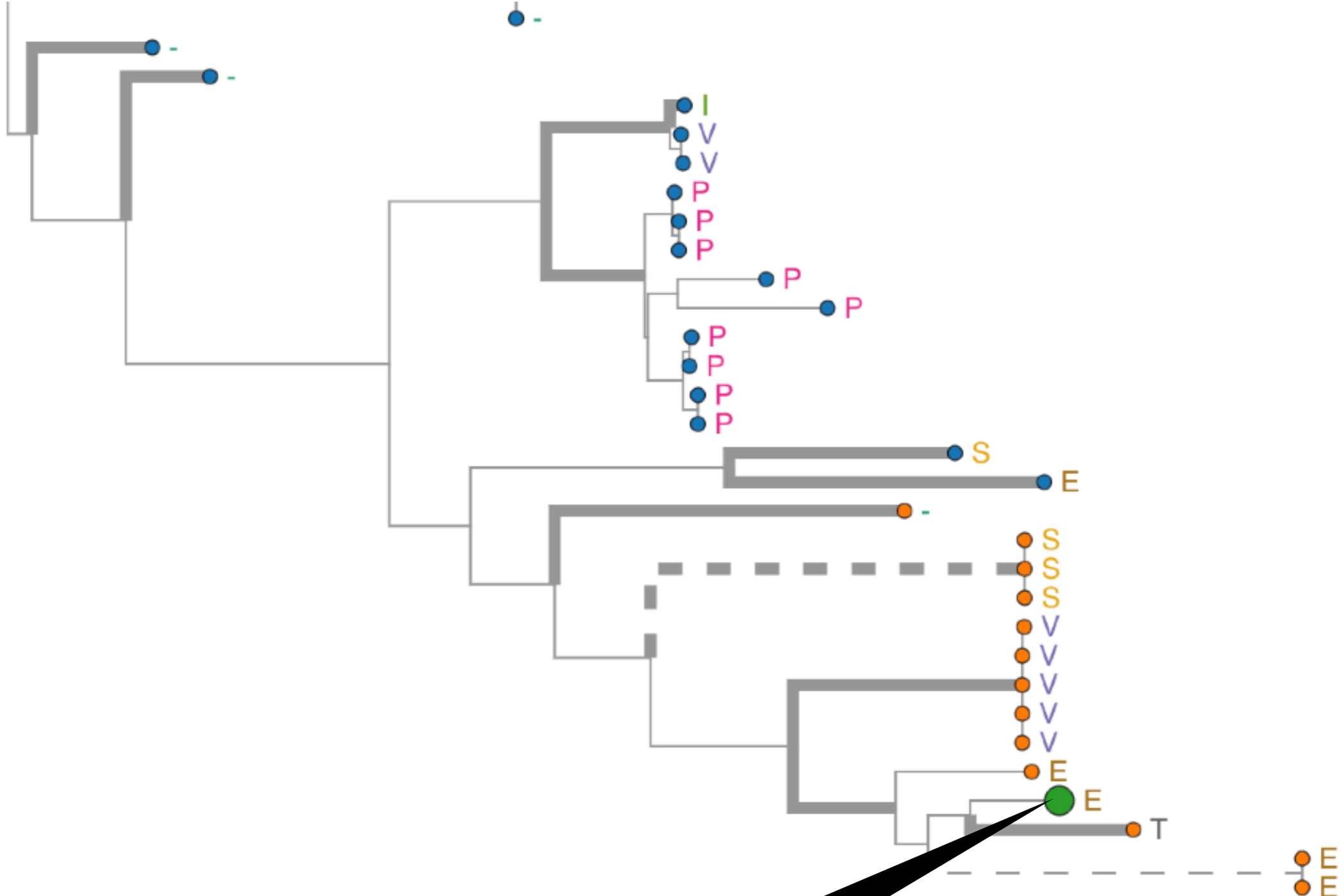
December 2019
November 2020
May 2021

- Use the evolutionary history in related *Sarbecoviruses* to predict which codons and amino-acids are “expected” in homologous SARS-CoV-2 positions.

- Our evolutionary model uses inferred site-level biochemical property importance to impute evolutionary credibility.

Continued evolution, complex selection dynamics, transition to endemic?

https://observablehq.com/@spond/visualizing-selection-analysis-results-for-evolution-of-n



Predicting S/484 possible states based on nCOV evolution

Evolutionary credibility report:

Codon	AA	Predicted probability in SARS-CoV-2
GAA	E	0.307
GAG	E	0.109
GTA	V	0.0818
GTT	V	0.0676
AAA	K	0.0477
GAT	D	0.0405
GCA	A	0.0315
GTG	V	0.0296
GGA	G	0.0272
GTC	V	0.0261
GAC	D	0.0195
AAG	K	0.0169
CAA	Q	0.0162
GCT	A	0.0154

Variable position => large admissible set of codons

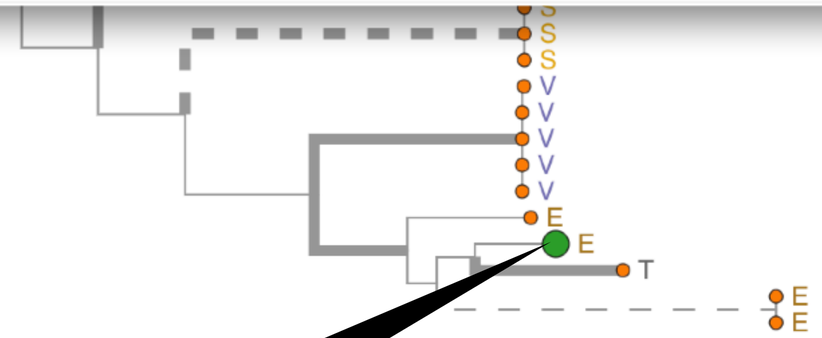
December 2019 | November 2020 | May 2021

Continued evolution, complex selection dynamics, transition to endemic?

- Use the evolutionary history in related *Sarbecoviruses* to predict which codons and amino-acids are “expected” in homologous SARS-CoV-2 positions.
- Our evolutionary model uses inferred site-level biochemical property importanc
evolutionary credibi

Evolutionary credibility report:

Codon	AA	Predicted probability in SARS-CoV-2
GAA	E	0.307
GAG	E	0.109
GTA	V	0.0818
GTT	V	0.0676
AAA	K	0.0477
GAT	D	0.0405
GCA	A	0.0315
GTG	V	0.0296
GGA	G	0.0272
GTC	V	0.0261
GAC	D	0.0195
AAG	K	0.0169
CAA	Q	0.0162
GCT	A	0.0154



Variable position => large admissible set of codons

Predicting S/484 possible states based on nCOV evolution

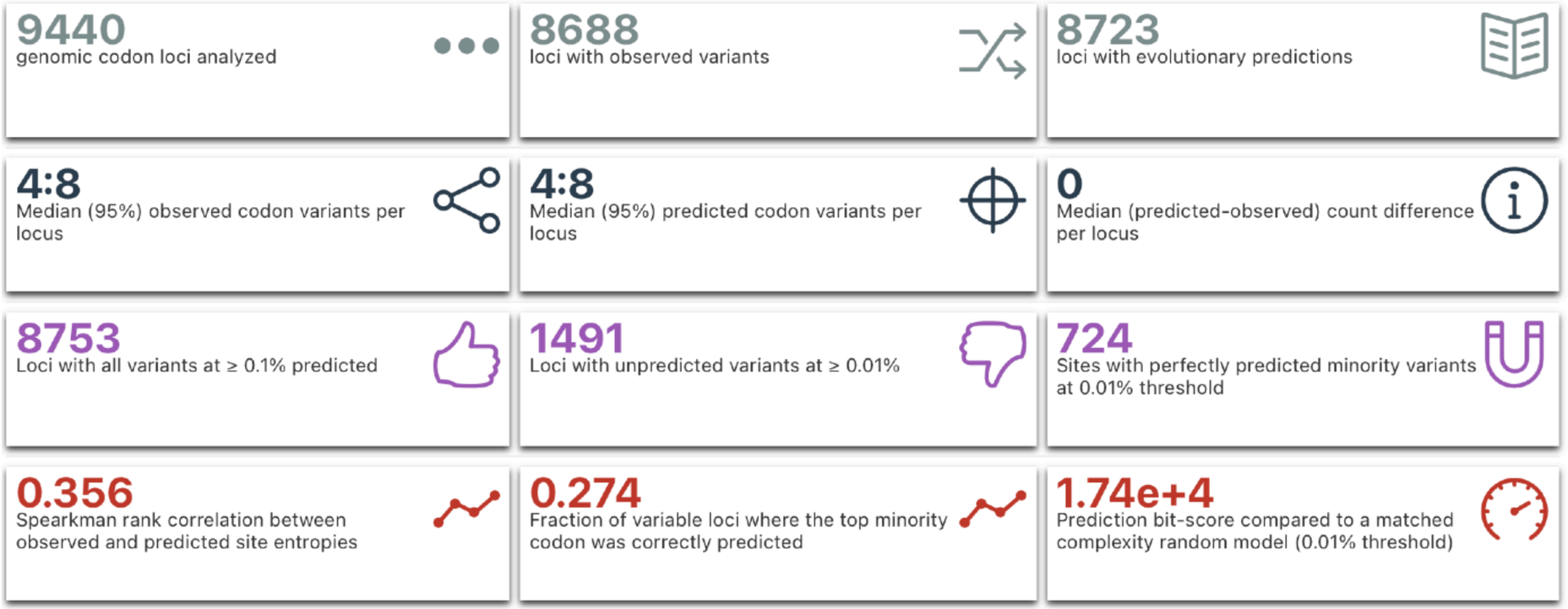
https://observablehq.com/@spond/visualizing-selection-analysis-results-for-evolution-of-n

December 2019 — November 2020 — May 2021

Continued evolution, complex selection dynamics, transition to endemic?

For a given set of SARS-CoV-2 genomic sites compare predicted probabilities of finding specific codons at given genomic sites (based on the evolutionary analysis of closely related animal sarbecoviruses) vs observed variation with a median of **3078961.5** consensus genomes per codon of SARS-CoV-2 from GISAID.

Download .JSON data



https://observablehq.com/@spond/sars-cov-2-pvo