

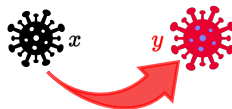
sequence database
(NCBI,GISAID)

Infer dependencies
across mutations
and genomic
variations



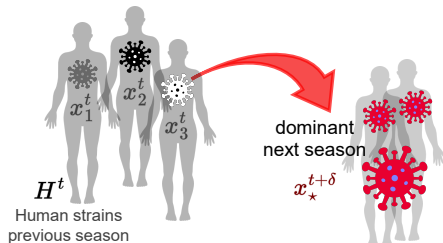
Emergenet

$Pr(x \rightarrow y)$



Application 1:

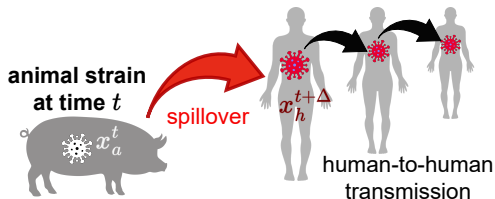
Forecast dominant strain(s) in upcoming season



$$x_*^{t+\delta} = \underset{x}{\operatorname{argmax}} \prod_{x^t \in H^t} Pr(x^t \rightarrow x)$$

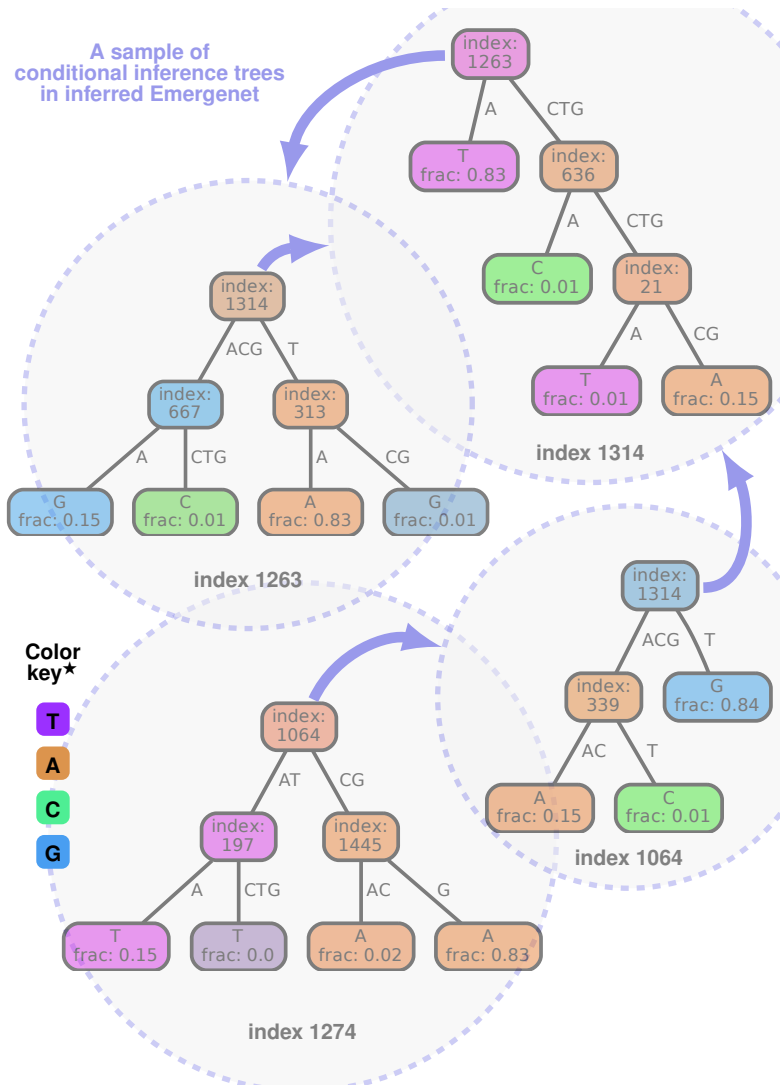
Application 2:

Estimate pandemic potential of animal strains



$$\text{risk}(x_a^t) = \log Pr(x_a^t \rightarrow x_h^{t+\Delta})$$

A sample of
conditional inference trees
in inferred Emergenet



H1N1 2018

Haemagglutinin Sequences

GGAAACAAAAGCAACAAAA... TGA^AAAAAAGA...
AGCAAAAGCAGGGGAAAAACA... GTT^CAACCAC...
AGCAAAAGCAGGGGAAAAACA... GTT^TAACCAC...
ATGAAGACTATCATTGCTTT... ACC^TTGAGAA...

residue
at index 1274

A/Italy/7366/2018

A/Baltimore/P0264/2018

A/Baltimore/P0278/2018

A/Florida/61/2018