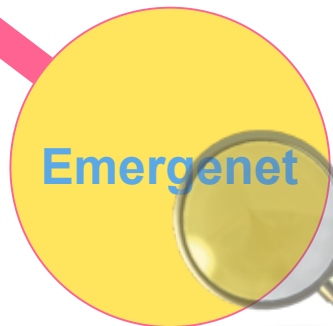
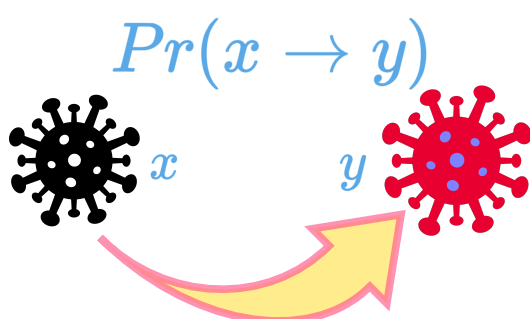
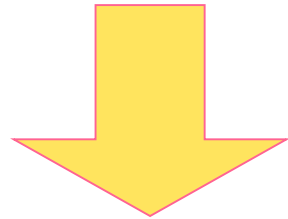


sequence database  
(NCBI, GISAID)



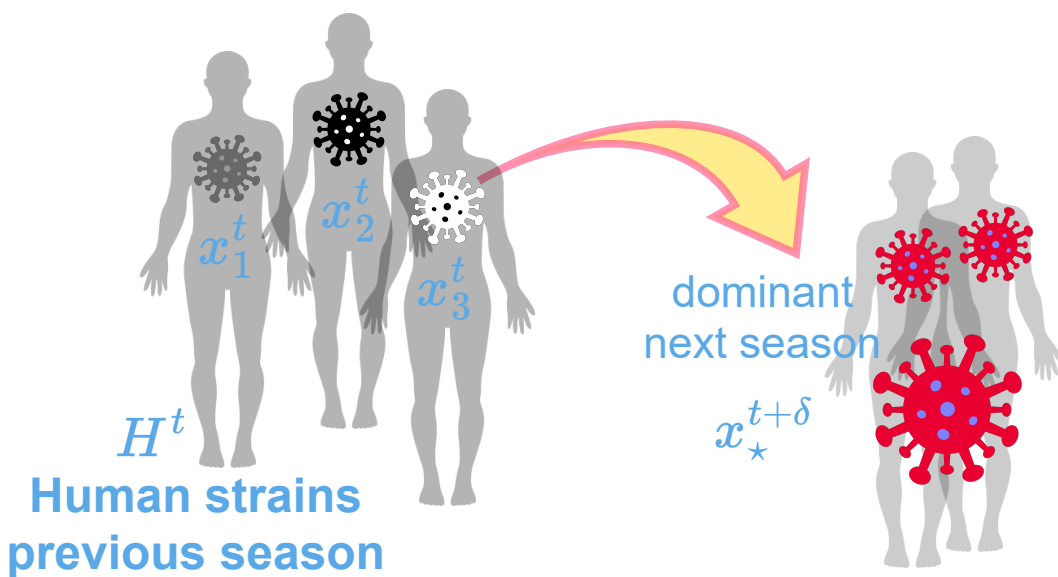
- Quantify probability of specific mutations in the wild
- Implicitly factor in functional constraints and selection pressures

Infer dependencies  
across mutations  
and genomic  
variations



## Application 1:

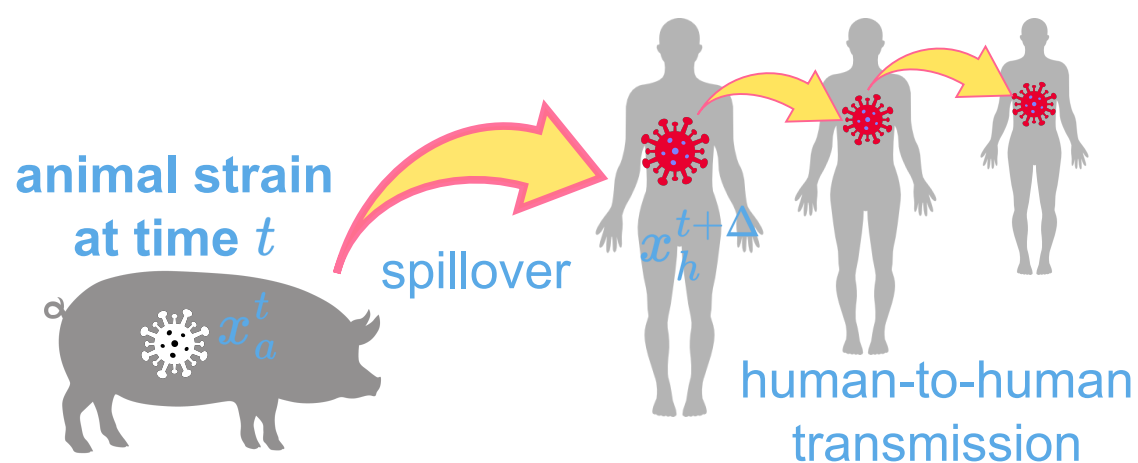
Forecast dominant strain(s) in upcoming season



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

## Application 2:

Estimate pandemic potential of animal strains



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