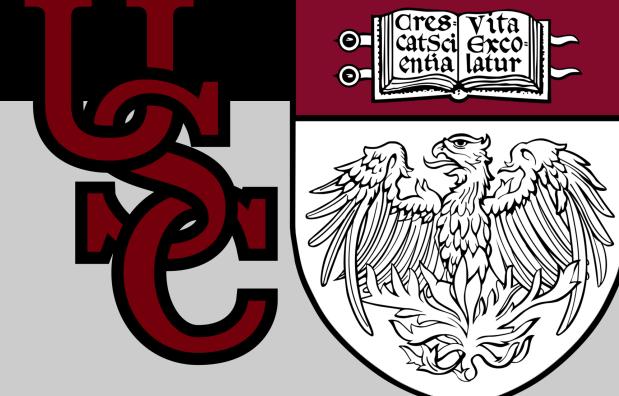


Future-proof Countermeasures

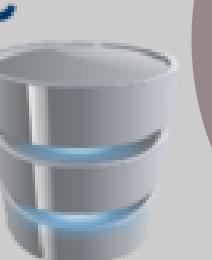
ishanu Chattopadhyay
Rick Stevens
Aaron Esser-Kahn
Paul Bogdan
S. Gnanakaran



Argonne
NATIONAL
LABORATORY



sequence database
(NCBI, GISAID)



Immunological
and structural
information

Infer dependencies
across mutations
and genomic
variations

Mechanistic
Knowledge
+
Novel Machine
Learning

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

$$Pr(x \rightarrow y)$$



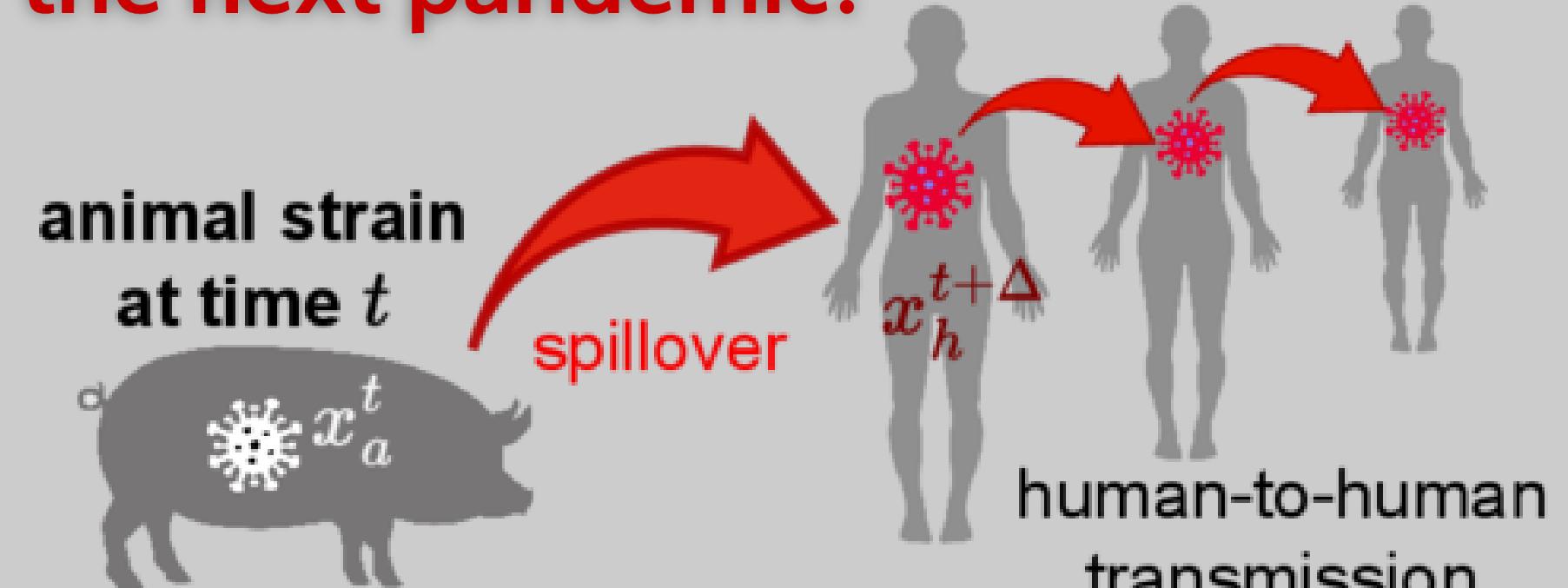
Emergenet

What are we trying to do?
Develop future-proof vaccines and countermeasures, escape-resistant to emergence of variants and zoonosis

What is the new approach?
Learn deep generative models of emergent dependencies across non-colocated residue variations, which induce a biology aware metric

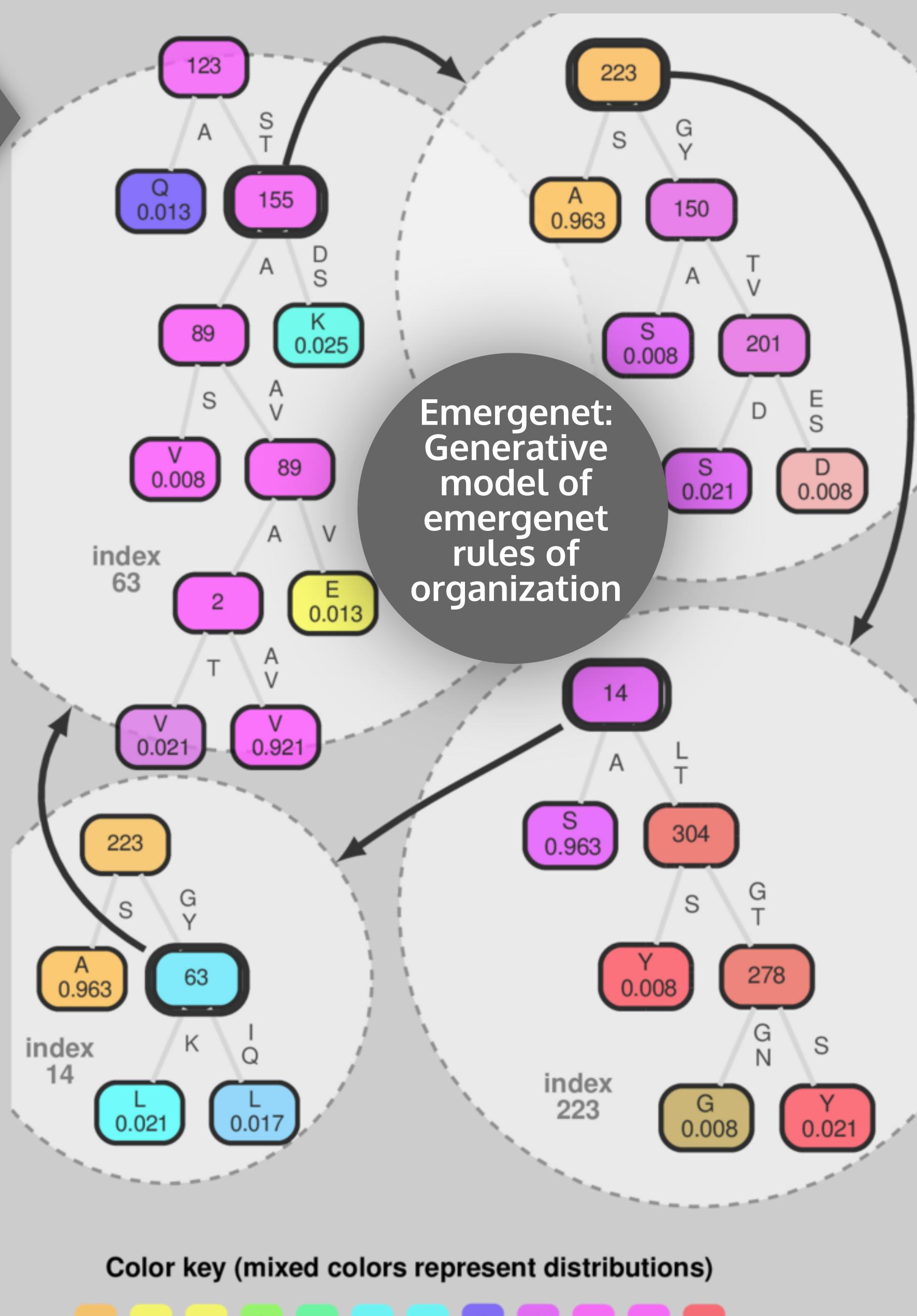
Can we predict future mutations?
Can we define the "edge of emergence"?

Problem: Can we predict the next pandemic?



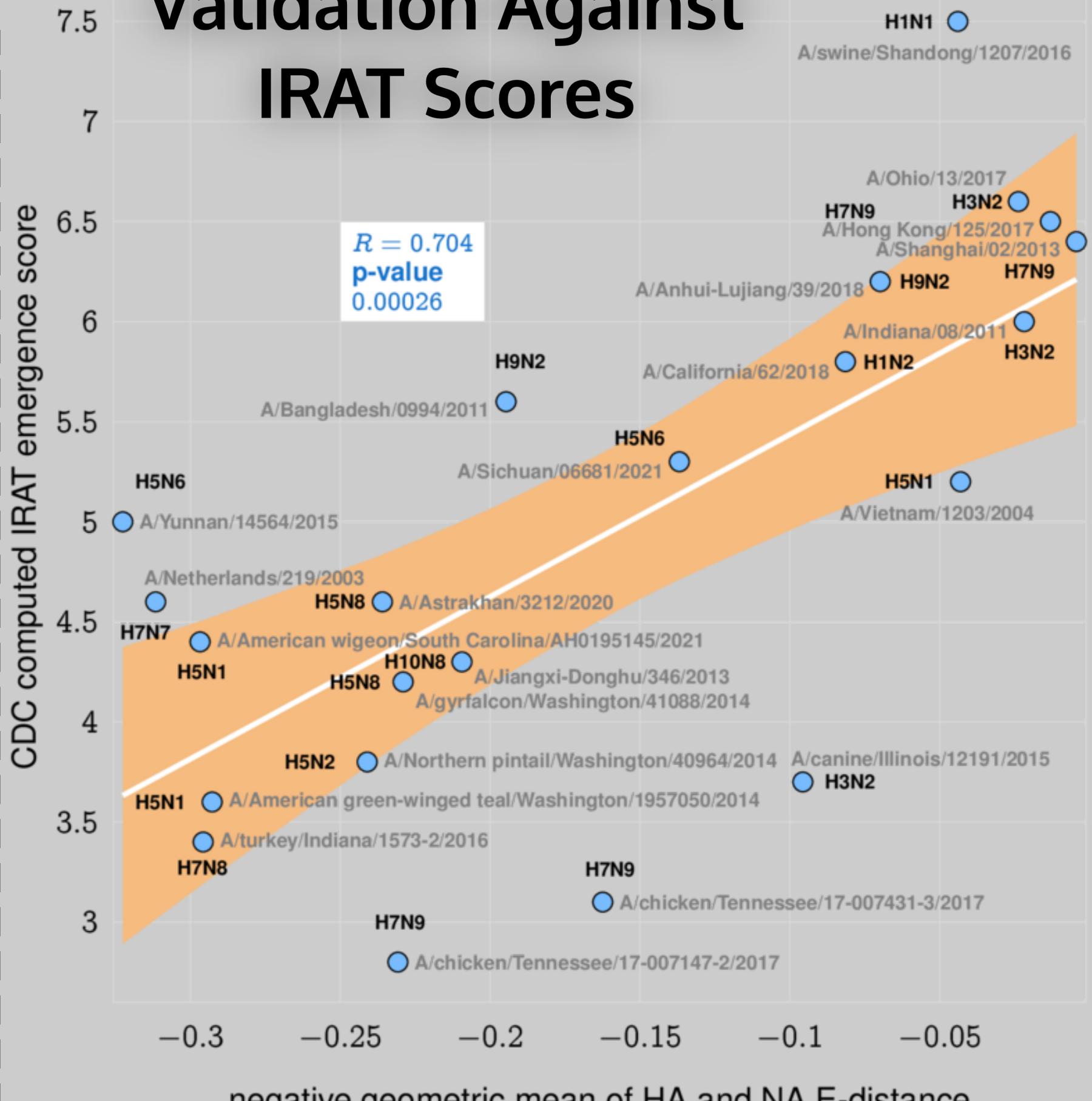
Who Cares?

Allows rapid reliable prediction of future mutations, variants and rank orders strains by their pandemic risk. Can target viruses BEFORE the first human infection

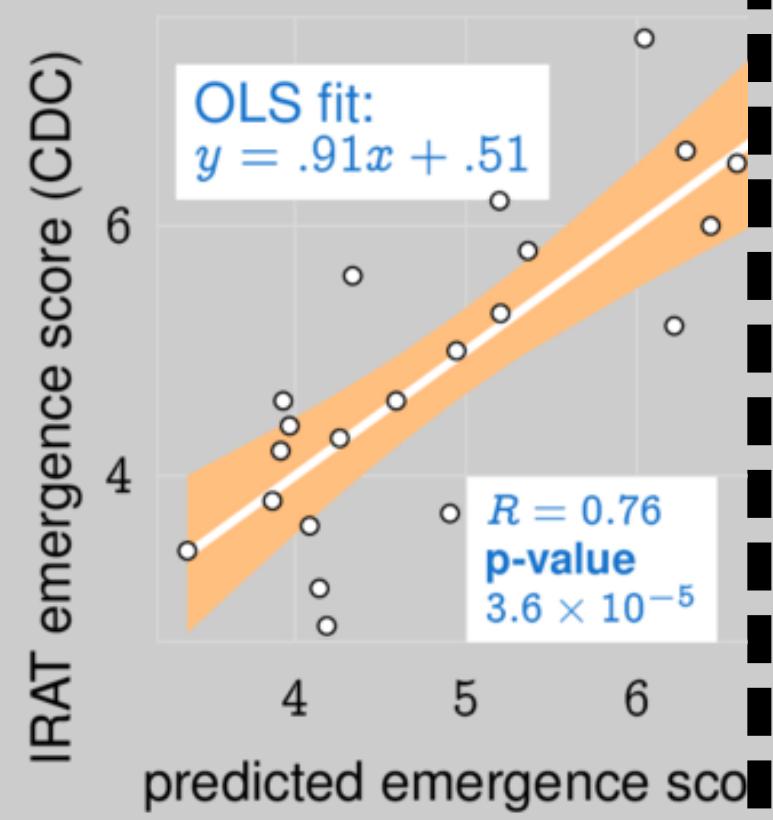


Predicted emergence risk vs published IRAT scores

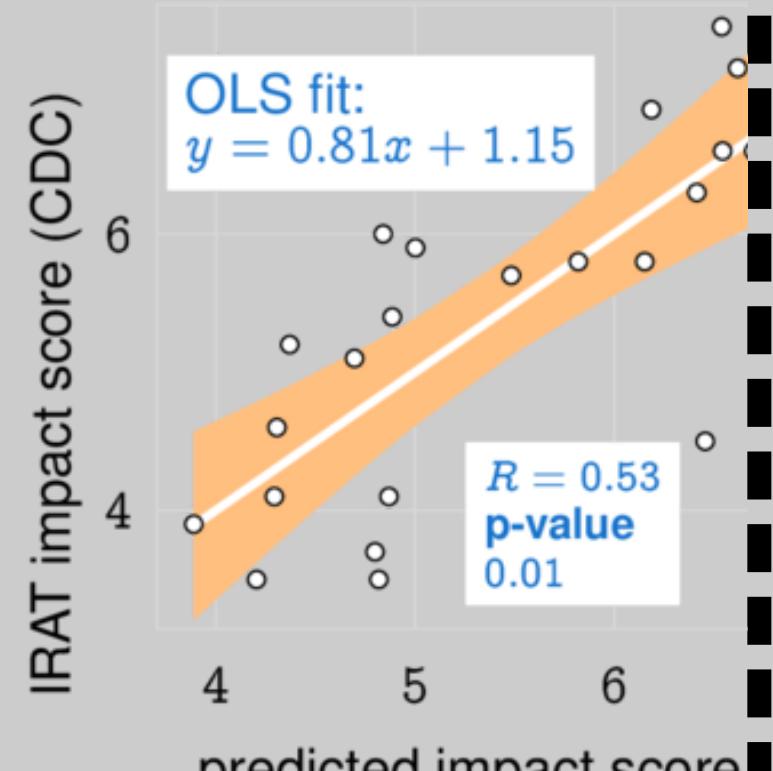
Validation Against IRAT Scores



b. Estimating emergence



c. Estimating impact



H1N1 2020-2021
Haemagglutinin Sequences
...GTSRYSKFKPEIATRPKVRDQEGR...
...GTSKYGKKFMPEIARRPKVRNQEGR...
...GSSKYKRFTEPIVARPKVREQAGR...
...GSSKYKRFTEPIVARPKVREQAGR...

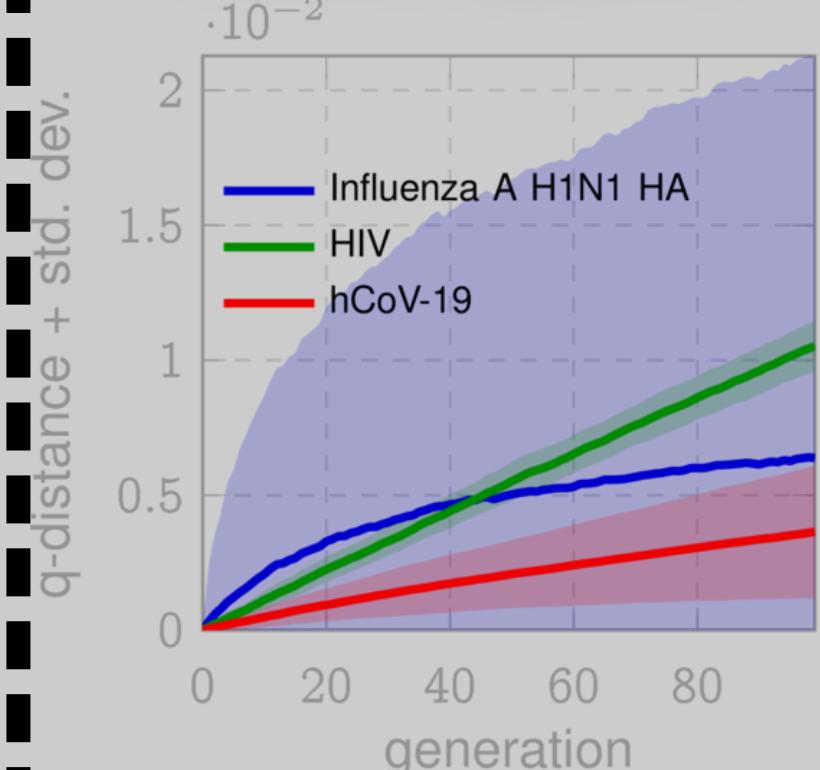
index 223

A/Niger/8327/2020
A/Parana/10835/2021
A/Gansu-Xifeng/1143/2021
A/Sichuan/01208/2021

$$\theta(x, y) \triangleq$$

$$E_i \left(\mathbb{J}^{\frac{1}{2}} \left(\Phi_i^P(x_{-i}), \Phi_i^Q(y_{-i}) \right) \right)$$

Predicting diffusion rate in strain space



Bio_NORAD

Global prediction of IRAT scores for all Influenza A sequences collected since 2020



highly vulnerable
vulnerable
affected

