

Deep Learning Methods for Learning Phenotype from Genotype

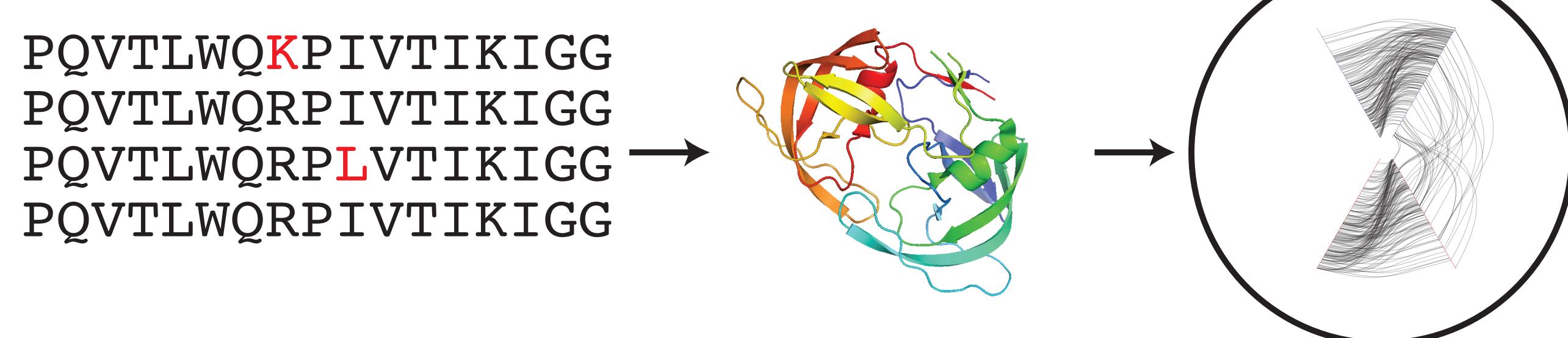
Eric J. Ma¹, David K. Duvenaud², Jonathan A. Rungstadler¹

¹MIT Biological Engineering, ²Univ. Toronto Computer Science & Statistics

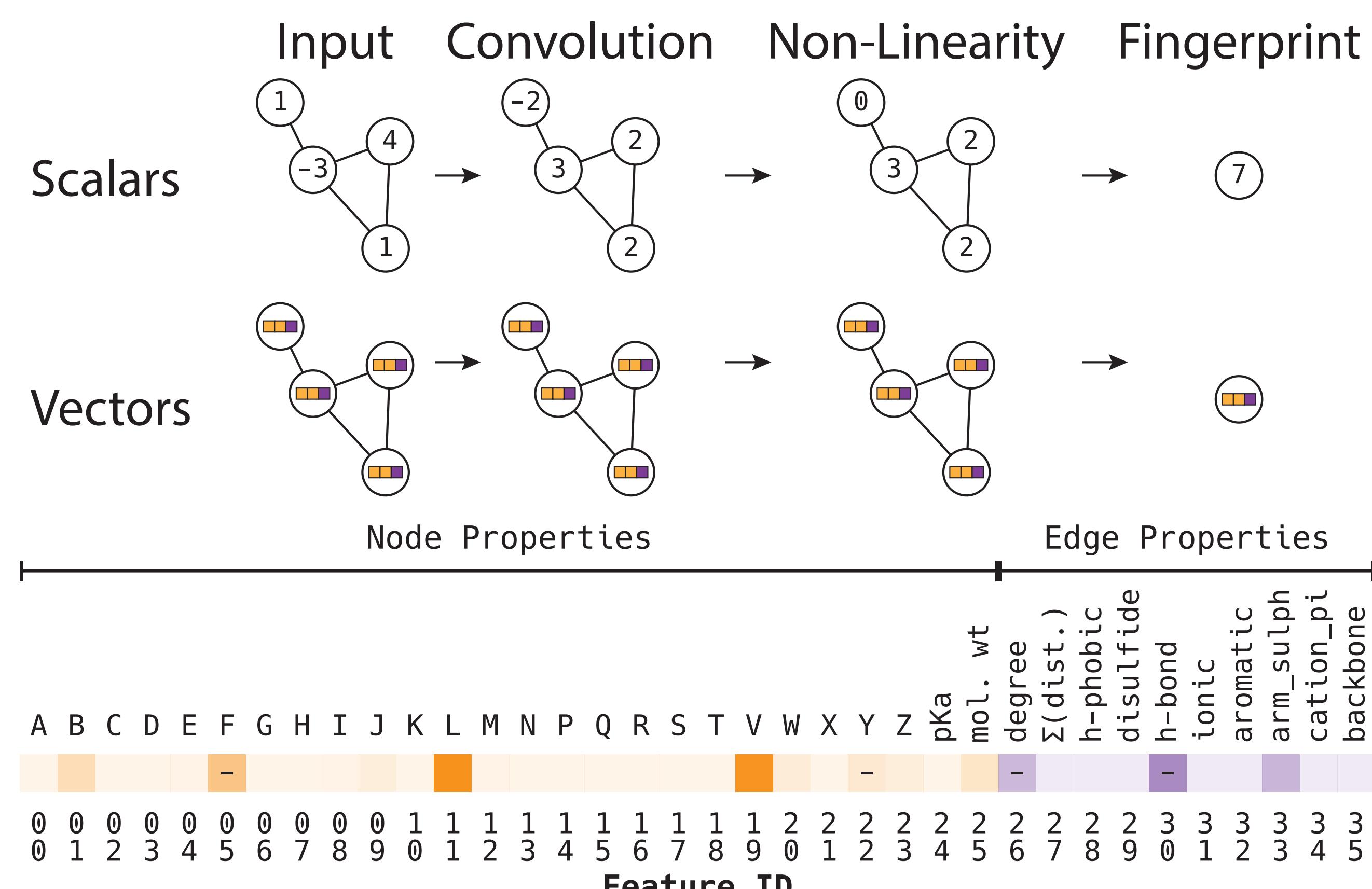
Automatic discovery of structural features predictive of phenotype

- Can we learn interactions between amino acids that predict phenotype?

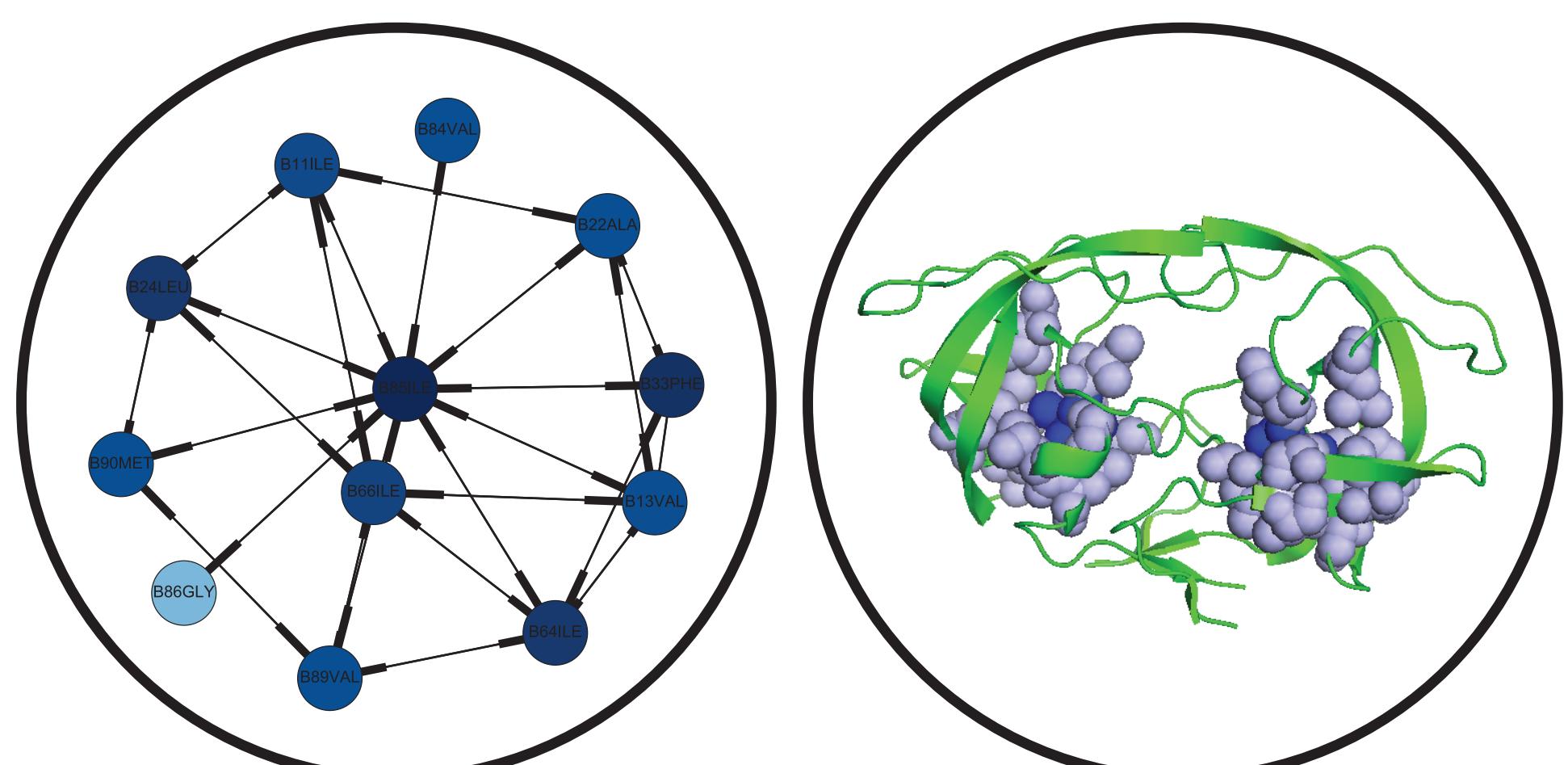
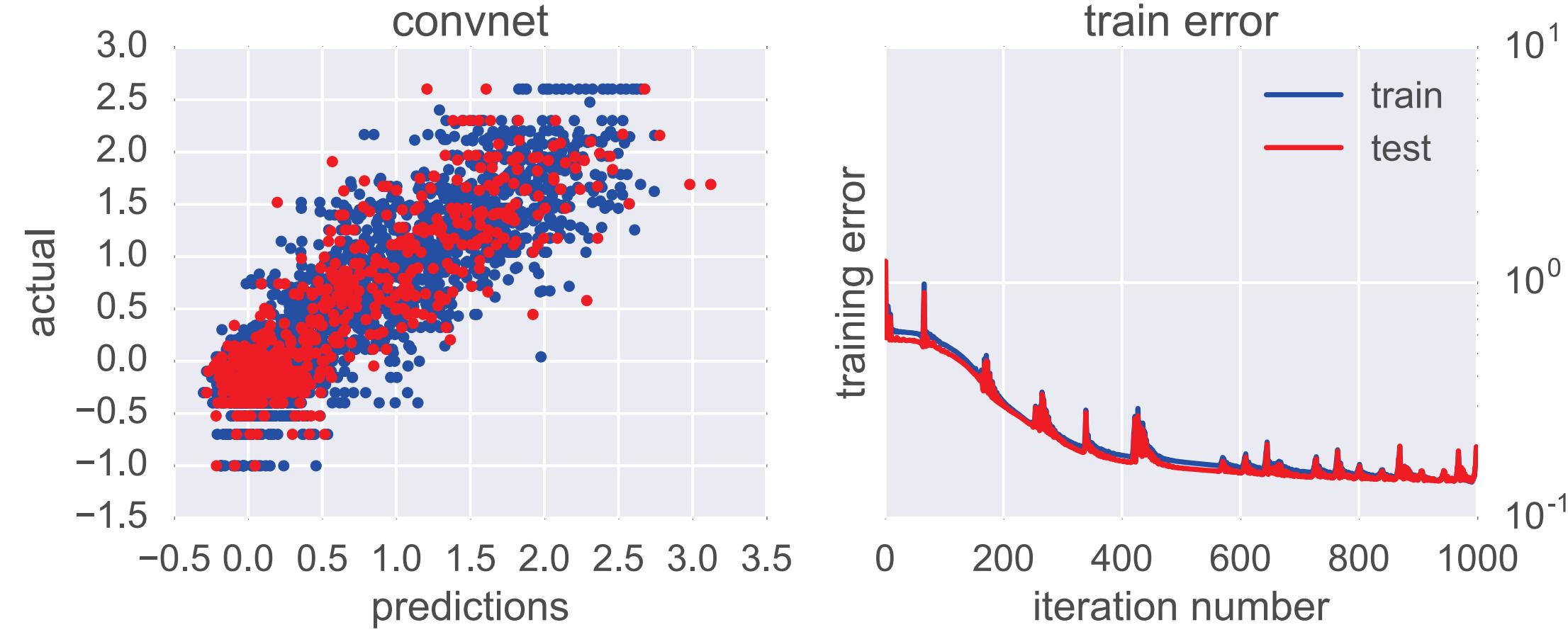
Input data: HIV sequence-drug resistance measurements, with sequence converted to 3D model & graph representation



Graph convolutions convert variably-sized, discrete inputs into a fix-length vector suitable for machine learning



Graph convolutions enable automatic learning of structural elements correlated with drug resistance



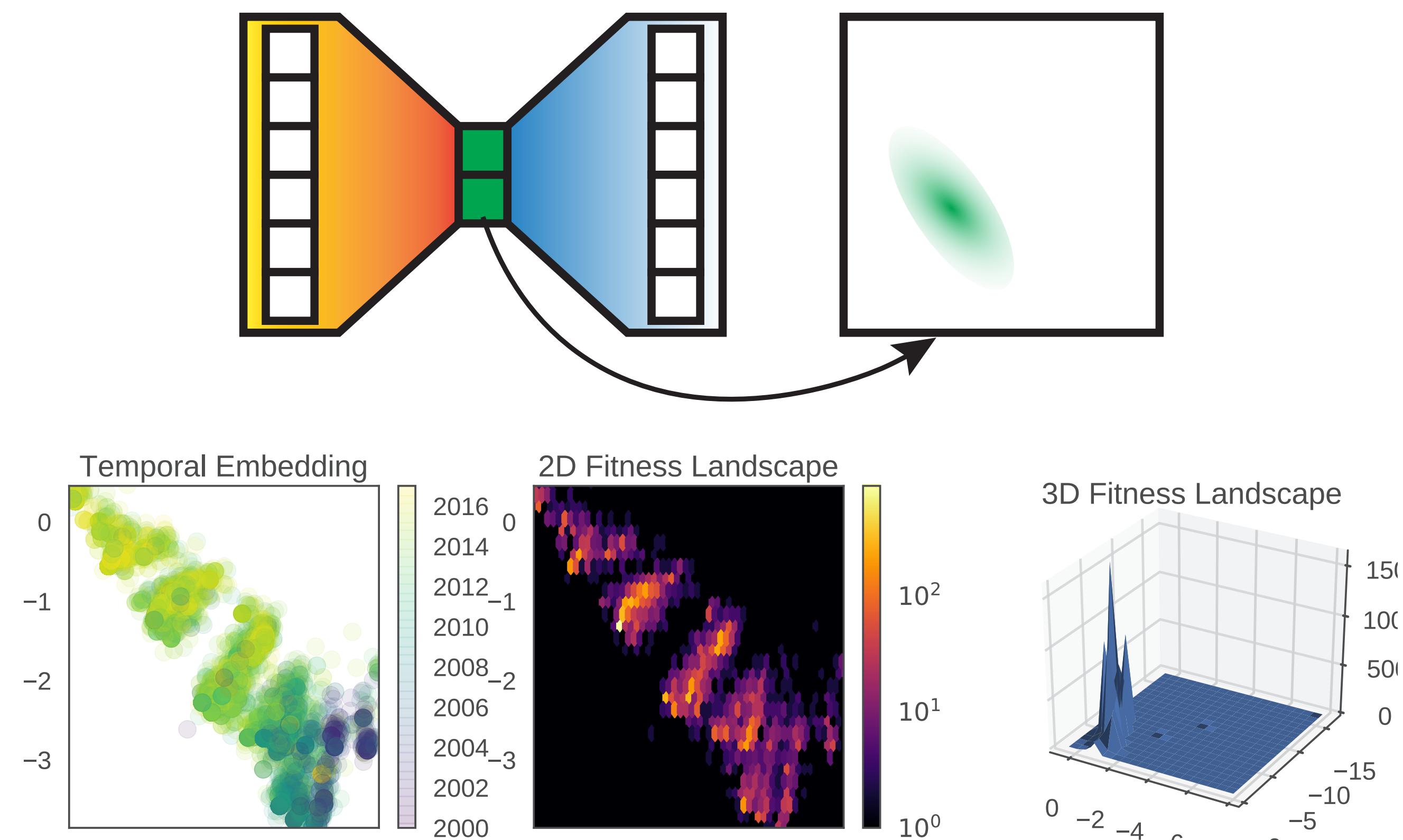
Sequence space forecasting using variational autoencoders & gaussian processes

- Can deep learning be combined with timeseries analysis to predict evolutionary trajectories?

Input data: un-aligned North American influenza H3 protein sequence (2000-2017) transformed to 1-of-K

Q	ACDEF	GHI	KLM	NPQR	STVWY
T					1
V					1
I				1	

Variational autoencoders learn continuous latent space from discretized sequence space.



Gaussian processes model time-series evolution with uncertainty.

