

Viral protein language model

HEV1 ... HEV4 ... HEV7

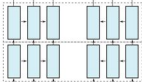
0.1 ... 0.2 ... 0.01

Genotype

Super-
vised



Semantic
embedding z



BiLSTM
layers



Input context
 $x_{[N]}$

V

L

S

_

K

A

A

Unsupervised