Simulated genomic windows
(N base pairs)

Aligned haplotypes
(M base pairs)

Feature maps

CNN

P(S_w)

b

mutation rate
[per virus]

viral replication
[per generation]

r_0

μ

neutral phenotype

μp

beneficial phenotype

d

viral death
[per generation]

Sliding window analysis on full simulated
non-recombining genomes

$P(S_{w1})$ . . . . $P(S_{wz})$

LSTM

Final positive selection estimates
accounting for genetic linkage
Detecting and quantifying strength of
positive selection, P(Selection) and (1 + s)

c

CNN

Reference

Positive selection: 37 (18.5%) | 181 (90.5%)

Neutral: 163 (81.5%) | 19 (9.5%)

Neutral | Positive selection
Prediction

d

LSTM

Reference

No beneficial mutation in window: 2177 (13.9%) | 80594 (93%)

Beneficial mutation in window: 13446 (86.1%) | 6083 (7%)

Beneficial mutation in window | No beneficial mutation in window
Prediction