

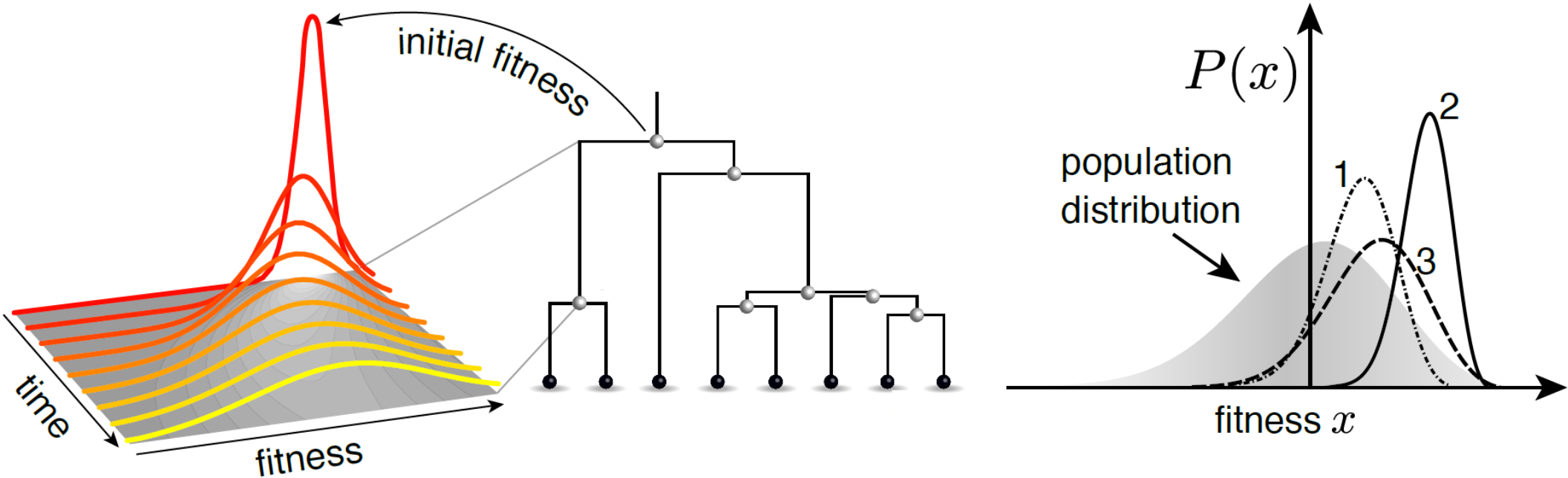


Predicting evolution from the shape of genealogical trees

with Boris Shraiman and Colin Russell

<http://arxiv.org/abs/1406.0789>

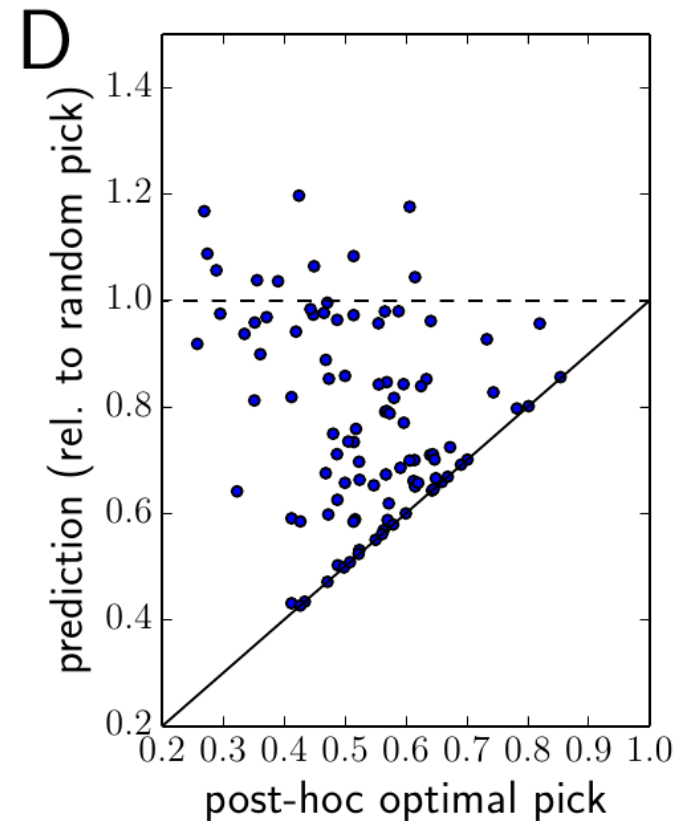
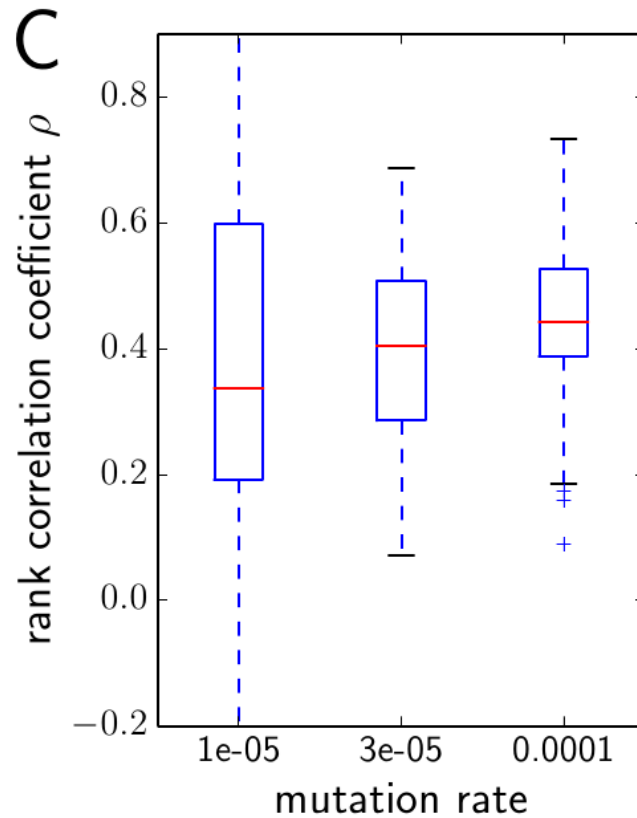
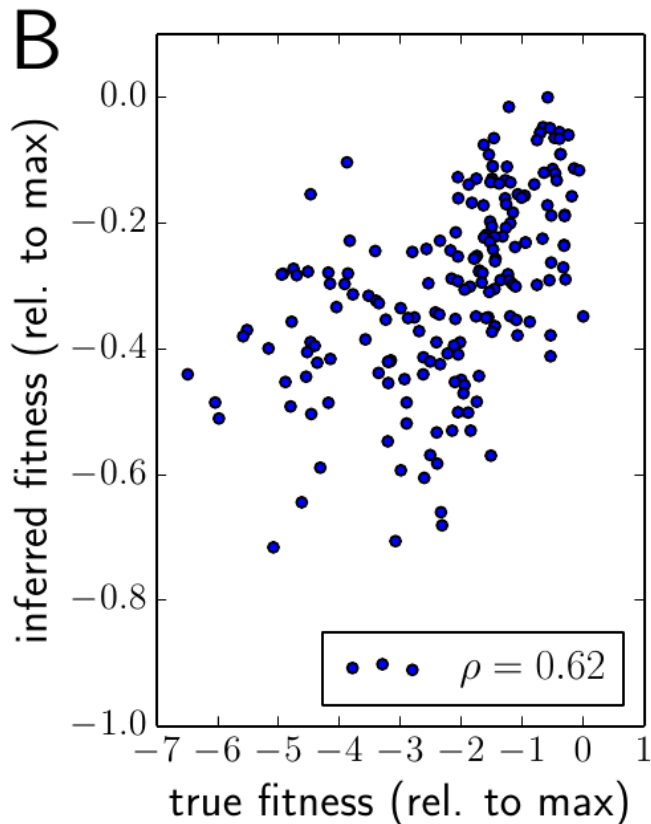
Inferring fitness



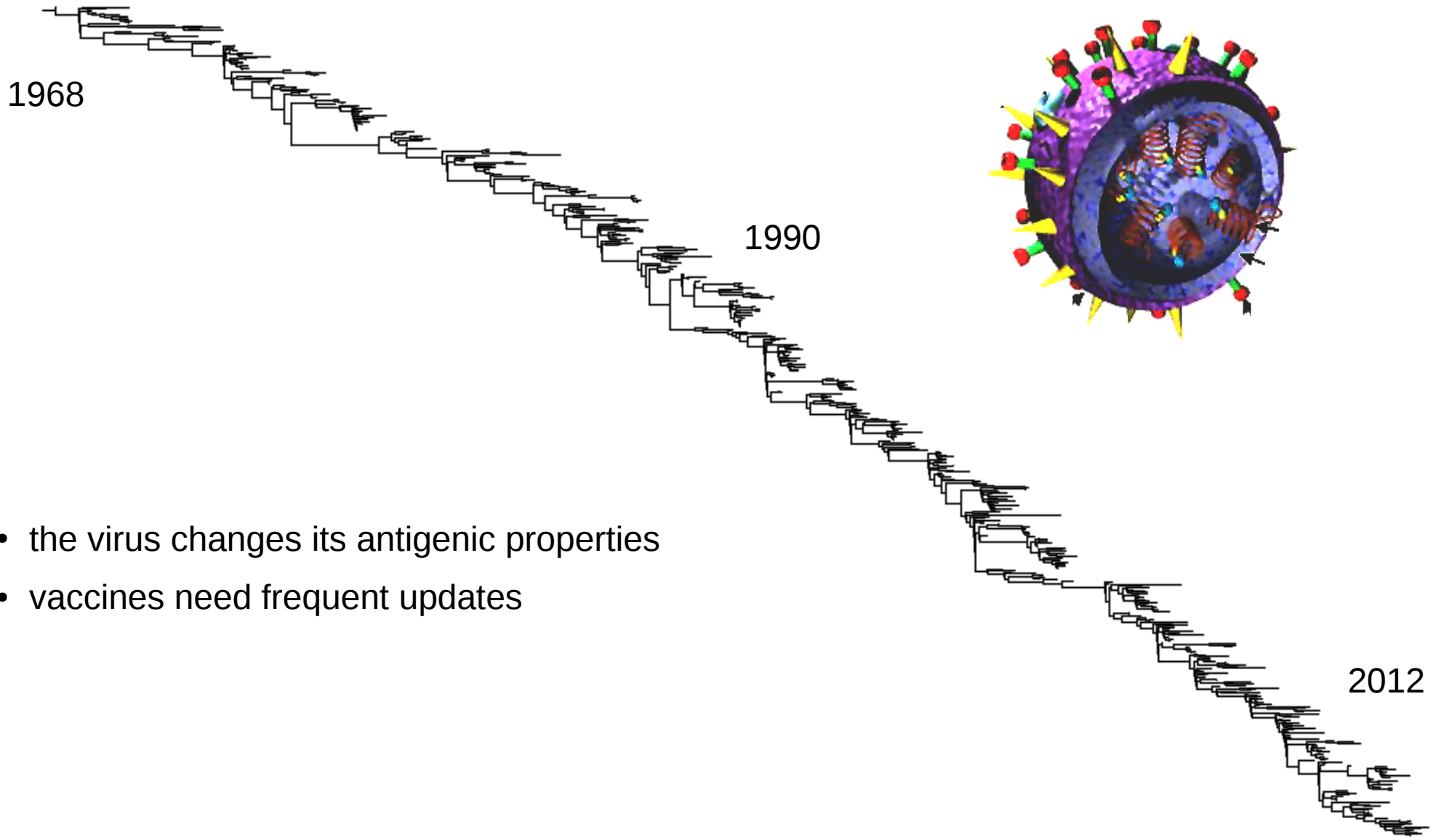
$$P(\mathbf{x}|T) = \frac{1}{Z(T)} p_0(x_0) \prod_{i=0}^{n_{int}} g(x_{i_1}, t_{i_1} | x_i, t_i) g(x_{i_2}, t_{i_2} | x_i, t_i)$$

Simulated data

- simulate evolution with beneficial and deleterious mutations
- take samples every 100 generations
- reconstruct trees and infer fitness



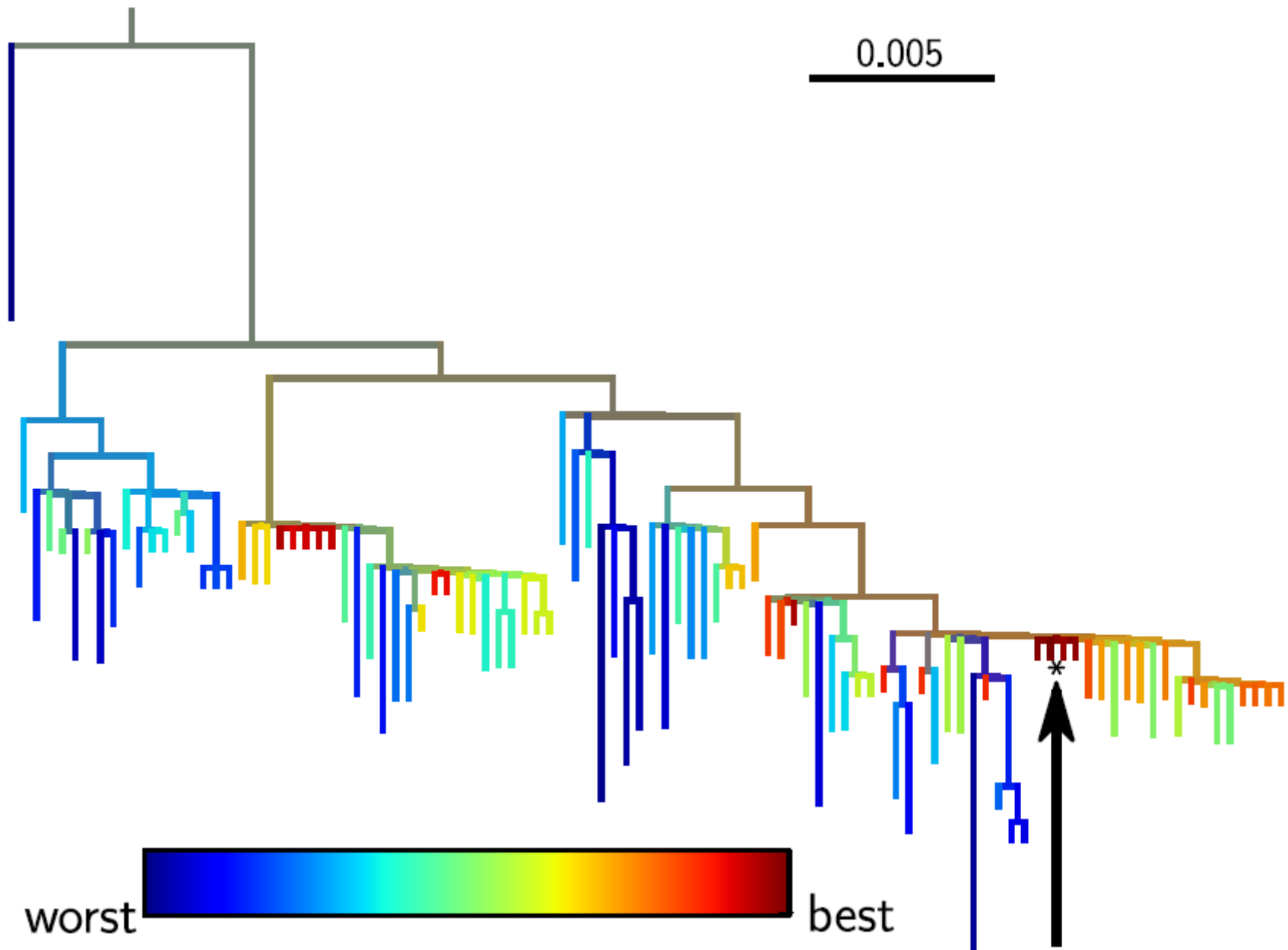
Predicting influenza evolution



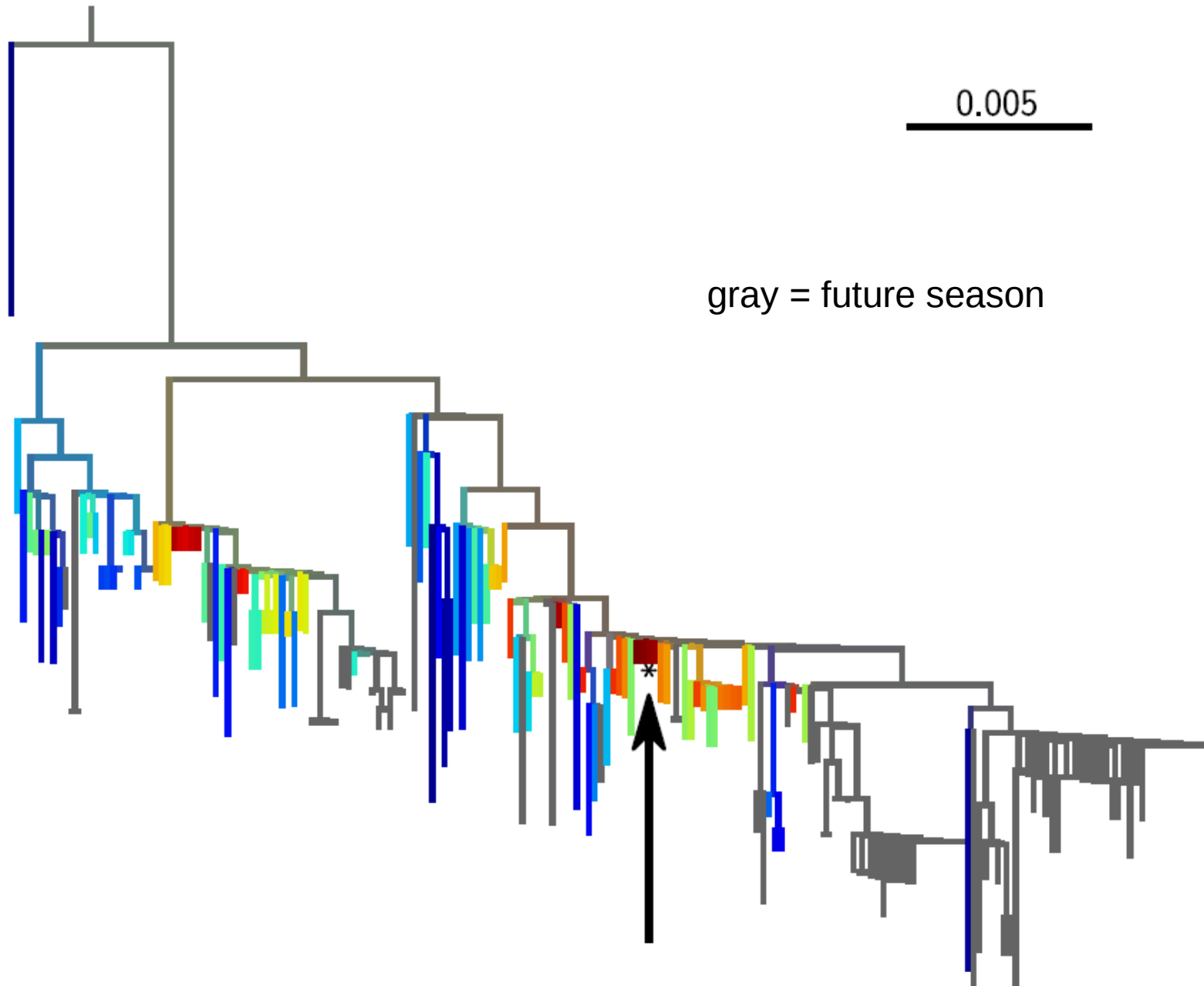
- the virus changes its antigenic properties
- vaccines need frequent updates

Progenitor of the future?

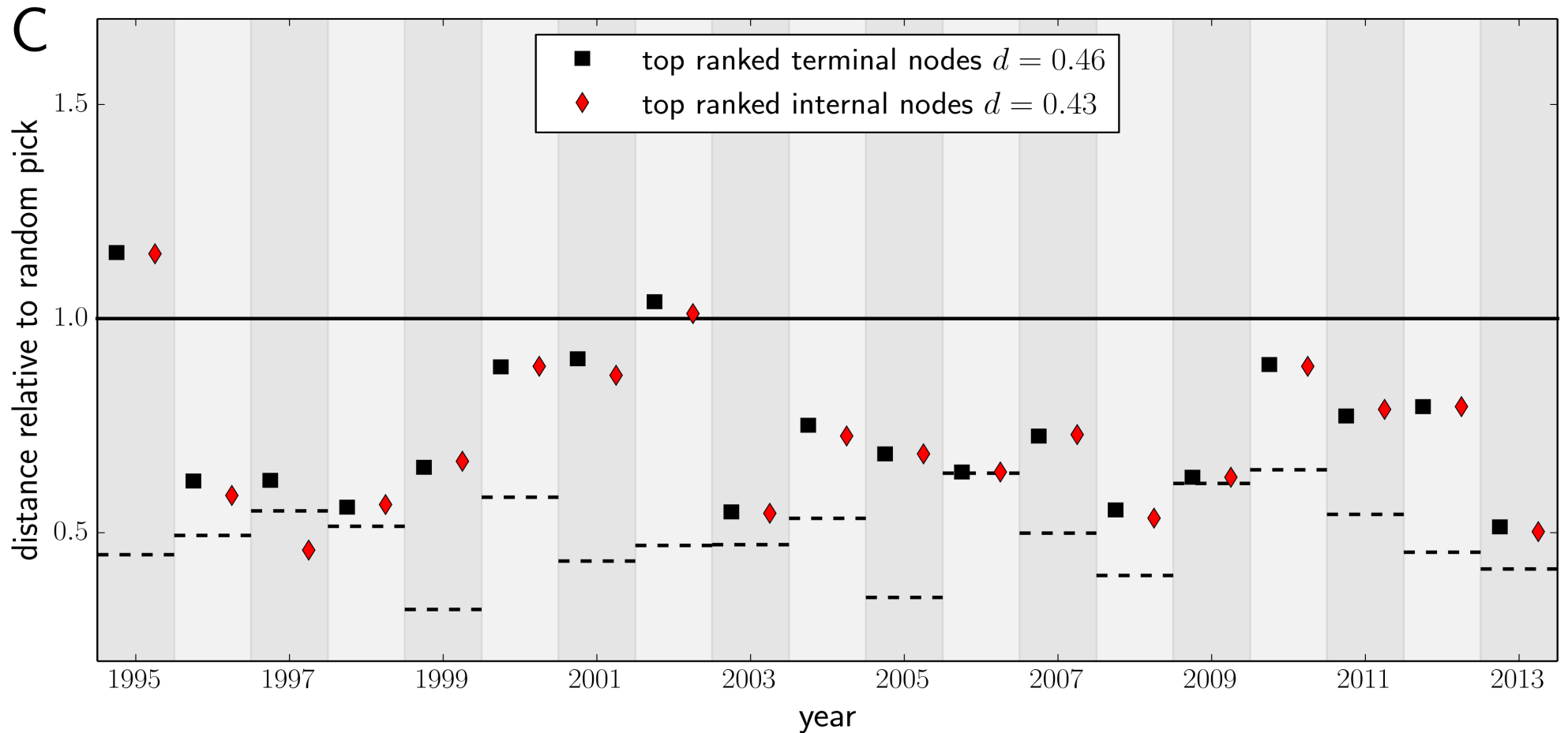
Predicting seasonal influenza



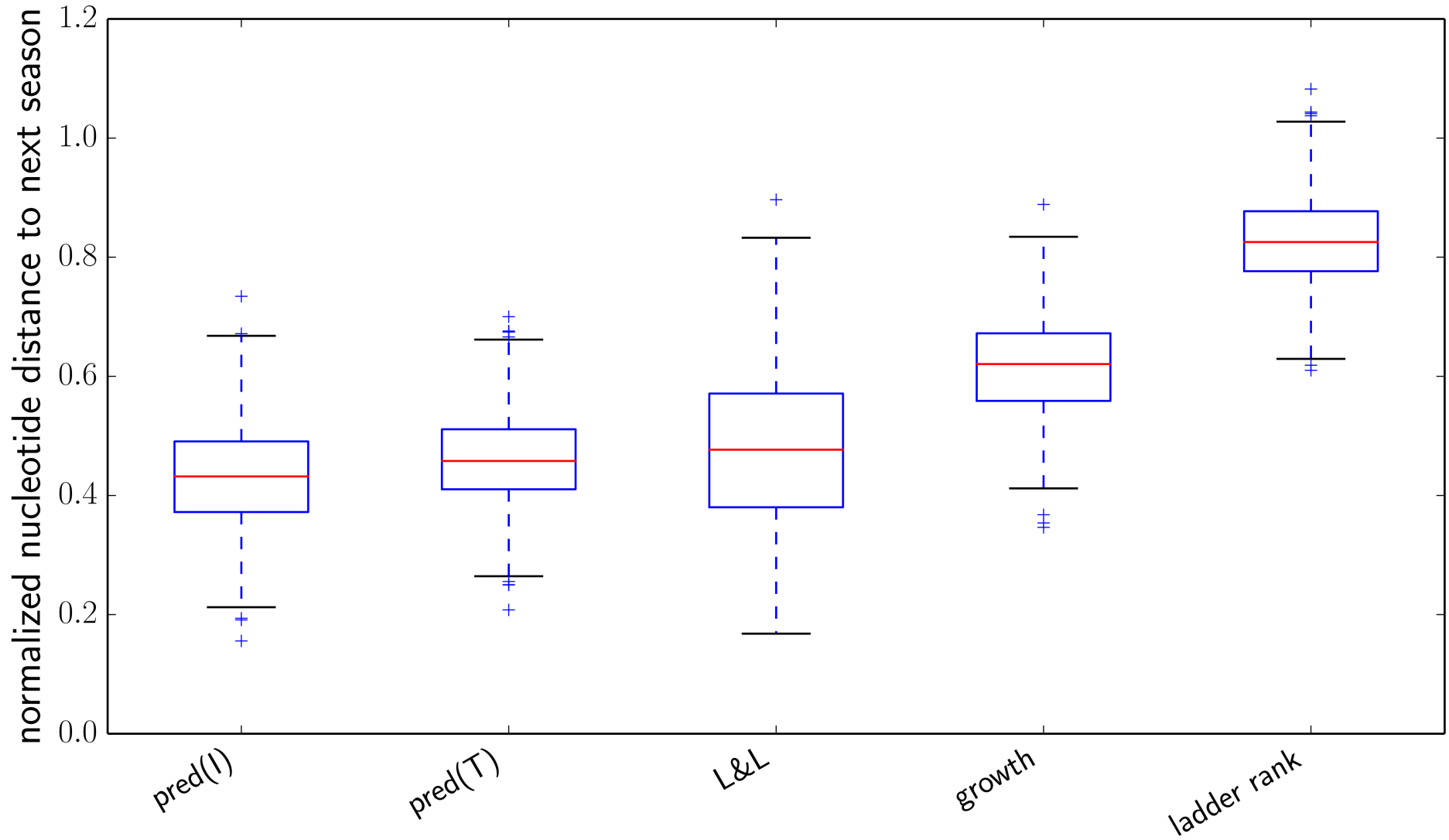
Predicting seasonal influenza



Predicting seasonal influenza (A/H3N2)



Predicting seasonal influenza (A/H3N2)



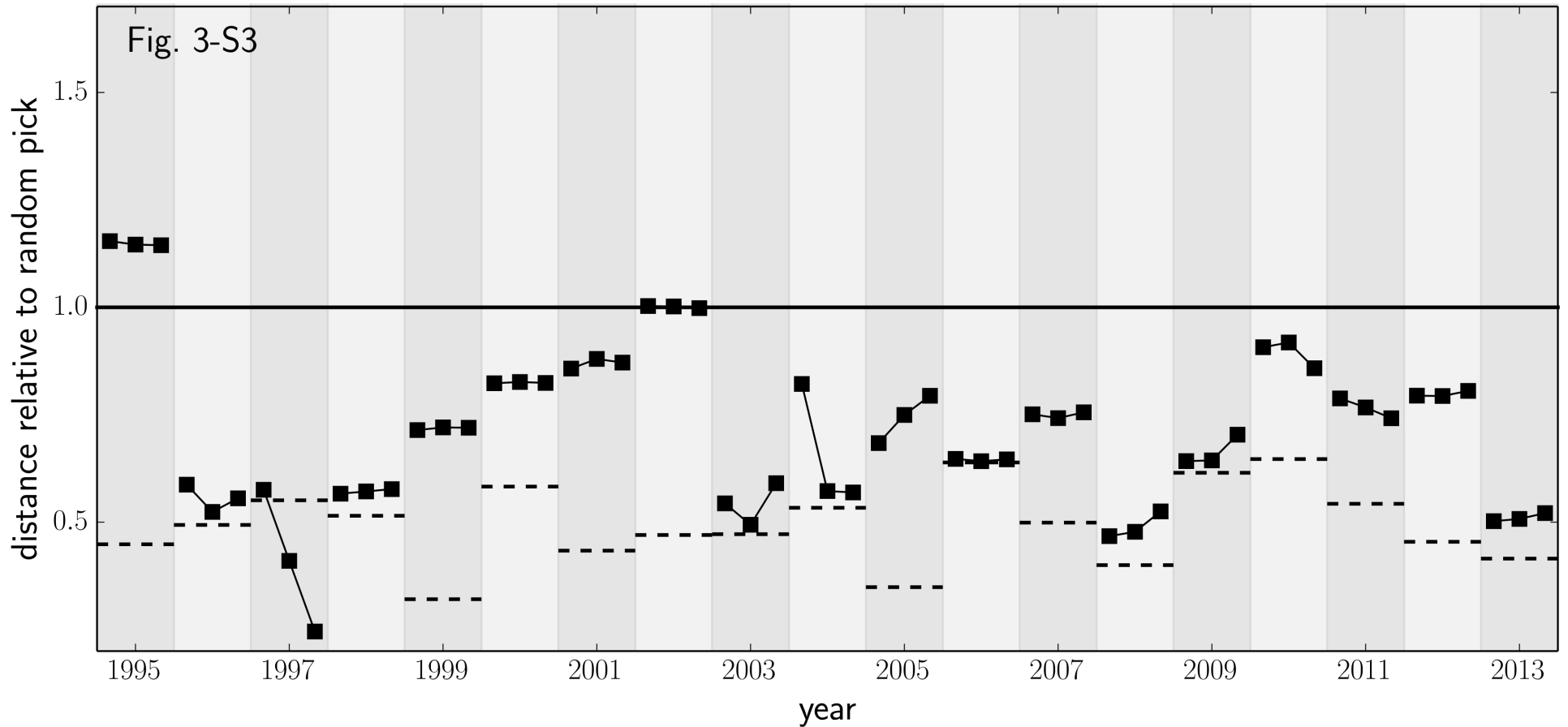
Predicting seasonal influenza (A/H3N2)

Enrichment of epitope and Koel sites among high fitness branches:

Quartile	# non-syn	# syn	# epi	# Koel
1	413	485	155	34
2	97	96	47	8
3	3	4	1	1
4	152	134	90	19
total	665	719	293	62

Comparison	enrichment	<i>p</i> -value
non-syn vs syn	1.3	0.04
epi vs syn	2.1	1e-05
Koel vs syn	2.0	0.025
epi vs non-syn	1.6	0.005
Koel vs non-syn	1.5	0.2

Including Koel mutations



Predicting MERS

