CDM CMM
Centre for Disease Modelling
Centre de la modélisation des maladies

YORK UNIVERSITÉ UNIVERSITY
redefine THE POSSIBLE.

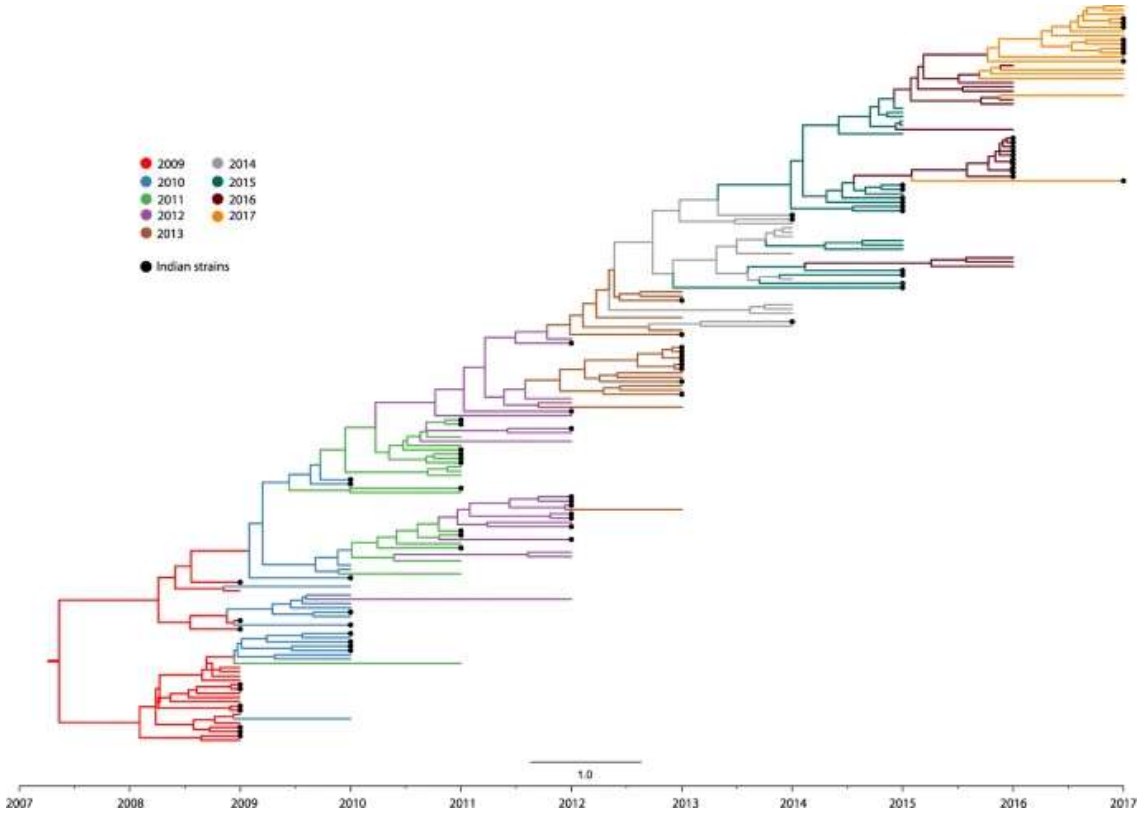
Evolution

Pathogen Evolution

- When pathogens replicate, they can incur mutation
  - Most mutations are deleterious or neutral
    - Mutant fitness is decreased or remains the same
  - Some rare mutations are advantageous
    - Can increase fitness for replication in the body, or transmission between individuals in the population

Multiple infectious diseases (or strains) spreading through one host species

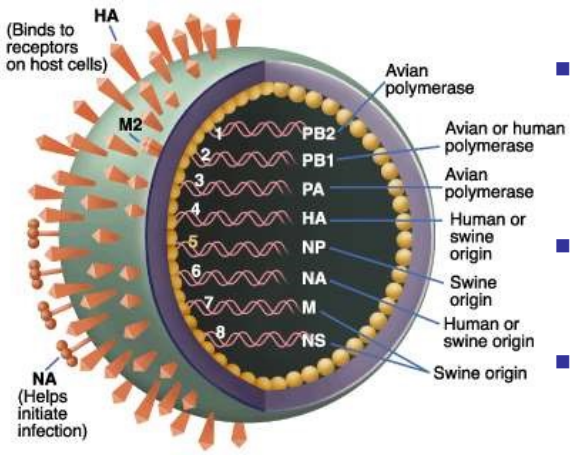
- Add mutation rate / evolution
- Strain structure
- Must consider effects of cross-immunity over longer timespans
- Must consider immune escape



https://www.nature.com/articles/s41598-019-51097-w Phylogenetic tree of H1N1 influenza A virus from Indian and global strains reported from 2009 till 2017 with branches colored by year of isolation

Influenza

- The flu virus evolves/mutates from year to year.
- Each year, the seasonal flu vaccine is updated based on a study of the previous year's strains.
- Phylogenetic analysis of virus strains can help determine evolutionary patterns.



- The influenza type A virus is made up of a viral envelope wrapped around a central core
- The envelope contains two large proteins:
  - hemagglutinin (HA) mediates binding/entry to target cell
  - neuraminidase (NA) involved in release of progeny from infected cells
- H1/H2/H3 and N1/N2 subtypes are the ones most commonly found in humans
- The influenza A virus genome is 13,588 bases long and contains eight RNA segments that code for 11 proteins: segment 4 encodes HA; segment 6 encodes NA

Data description

Data from Influenza Research Database (IRD) www.fludb.org:

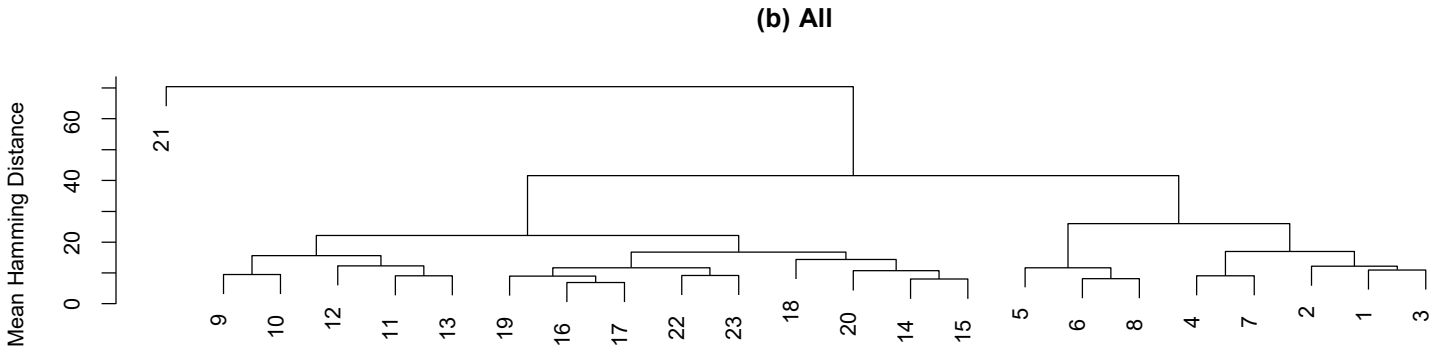
- Protein, A, H3N2, only complete segments
  - HA segment only
  - NA segment only
- from September 1998 to July 2012 (implies month of isolation available)
- Human host
- All geographic regions
- Remove duplicate sequences!
- Data translated into amino acids and aligned: 20 amino acids plus gap
- Date of isolation for each sequence was then used to assign each sequence into a flu season (Oct 1 - Sep 30)

This yields 1947 HA sequences of length 566 aa
2037 NA sequences of length 550 aa

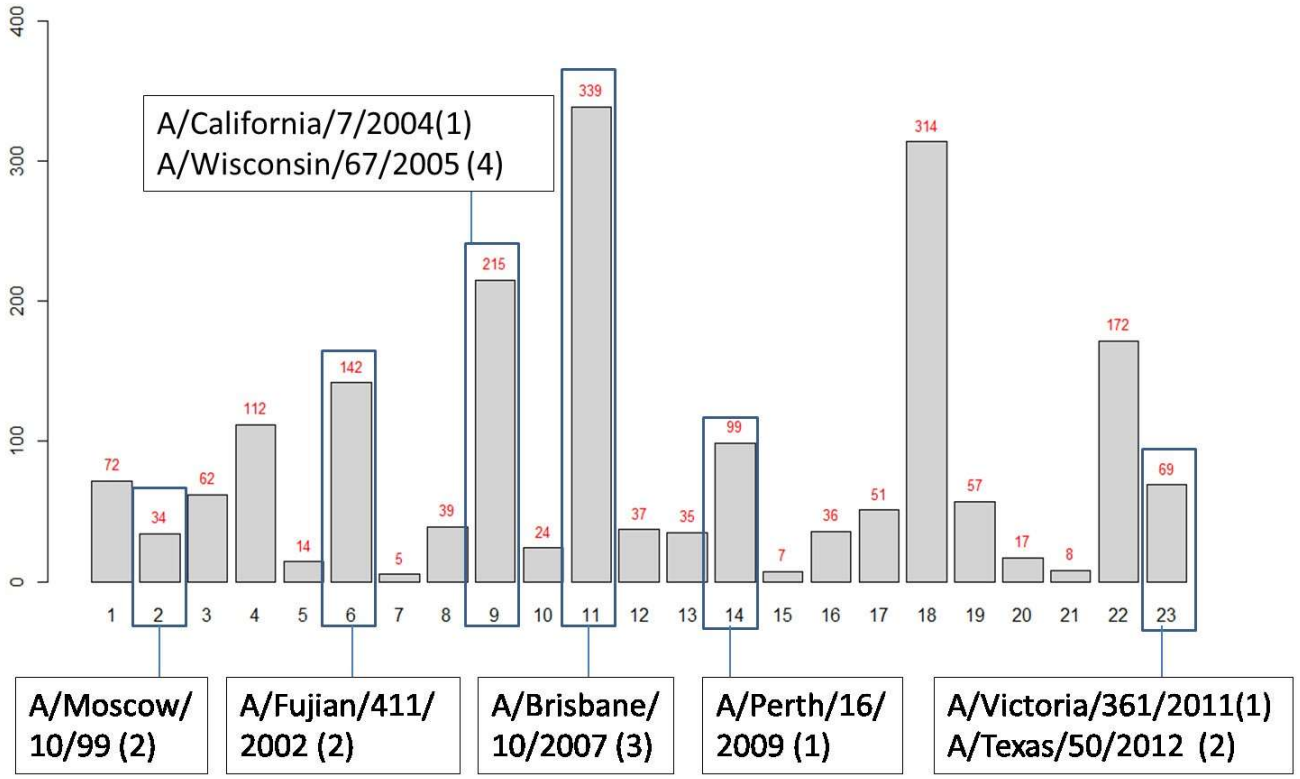
Vaccines for the time period (via WHO) were added manually.

Data Analysis for HA

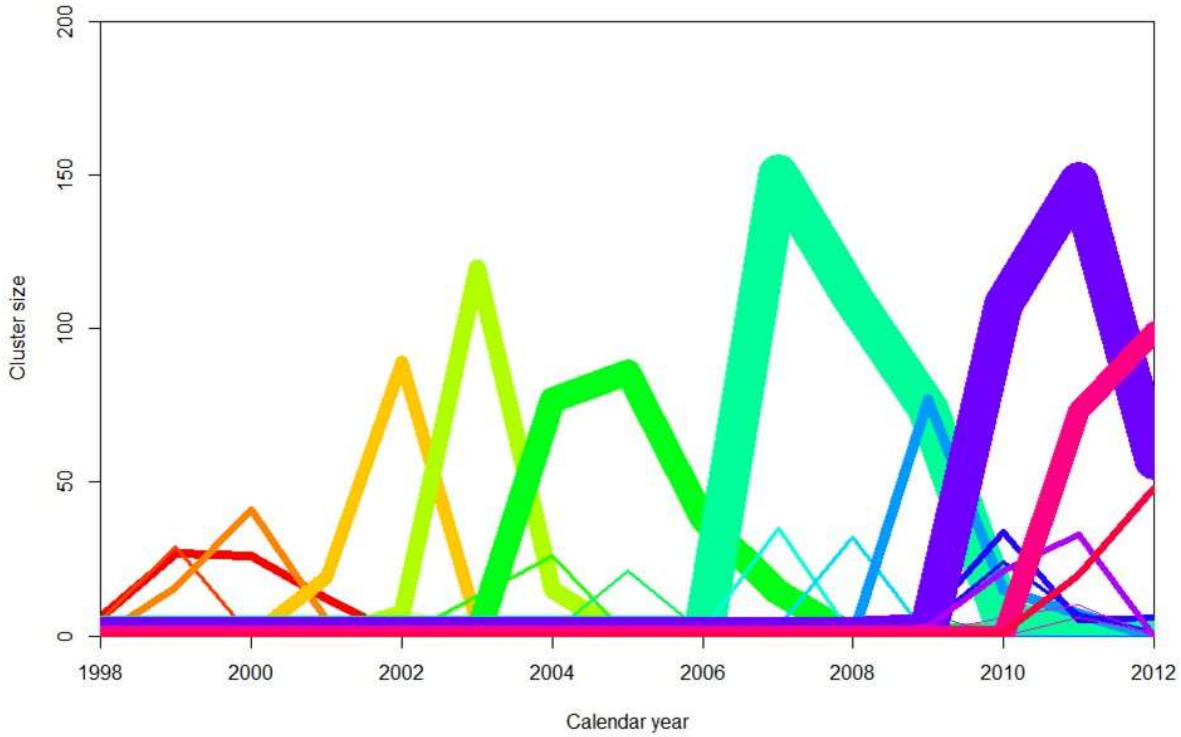
Dendrogram based on mean Hamming distance for all 566 sites;
on 62 sites the main change is a shift in cluster 21...



Results HA



Histogram of cluster size and vaccine location; ordered by earliest year of isolation.

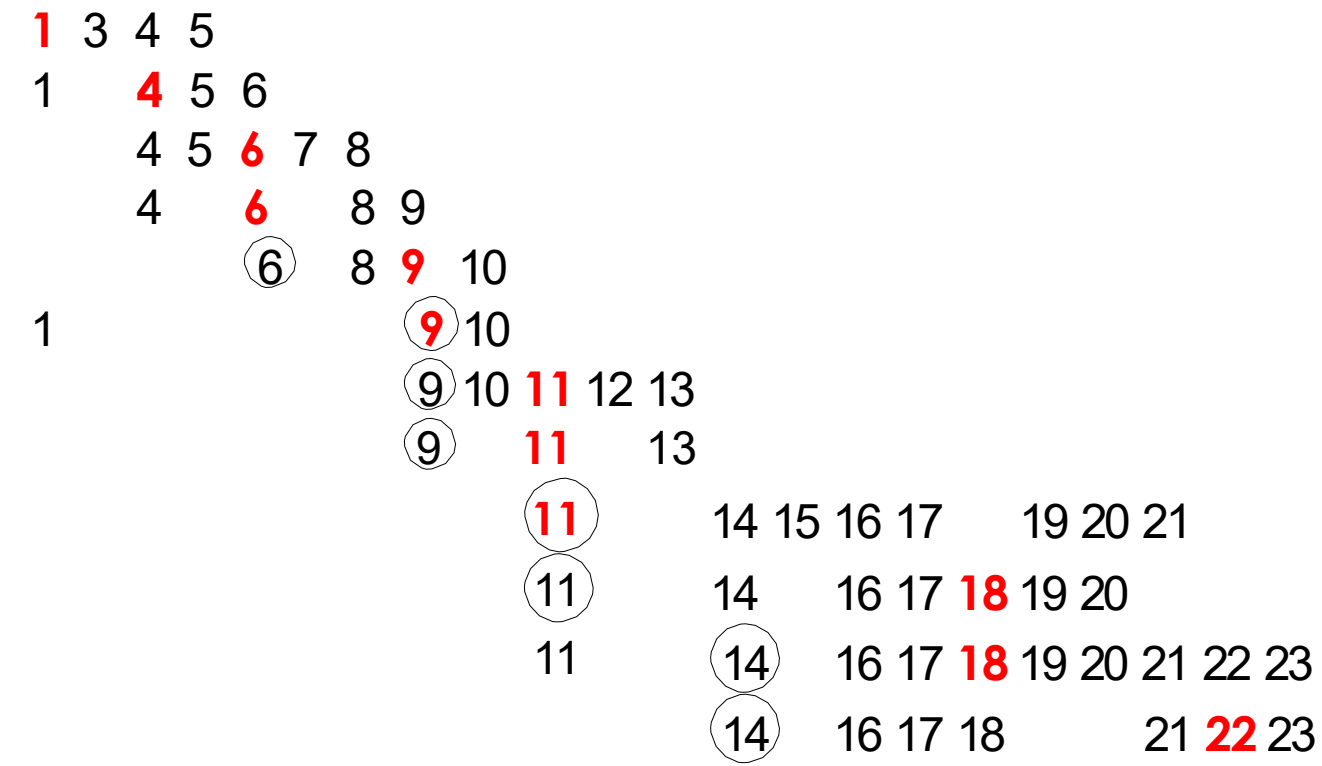


Number of sequences within each cluster vs calendar year by year of isolation.

Vaccines and clusters by year (2000/2001 to 2011/2012)

Season	Vaccine strain name	Clusters appearing	inant clu	er vaccine
2000-2001	A/Moscow/10/99	1,3,4,5	1	2
2001-2002	A/Moscow/10/99	1,4,5,6	4	2
2002-2003	A/Moscow/10/99	4,5,6,7,8	6	2
2003-2004	A/Moscow/10/99	4,6,8,9	6	2
2004-2005	A/Fujian/411/2002	6,8,9,10	9	6
2005-2006	A/California/7/2004	1,9,10	9	9
2006-2007	A/Wisconsin/67/2005	9,10,11,12,13	11	9
2007-2008	A/Wisconsin/67/2005	9,11,13	11	9
2008-2009	A/Brisbane/10/2007	11,14,15,16,17, 19,20,21	11	11
2009-2010	A/Brisbane/10/2007	11,14,16,17,18, 19,20	18	11
2010-2011	A/Perth/16/2009	11,14,16,17,18, 19,20,21,22,23	18	14
2011-2012	A/Perth/16/2009	14,16,17,18,21, 22,23	22	14

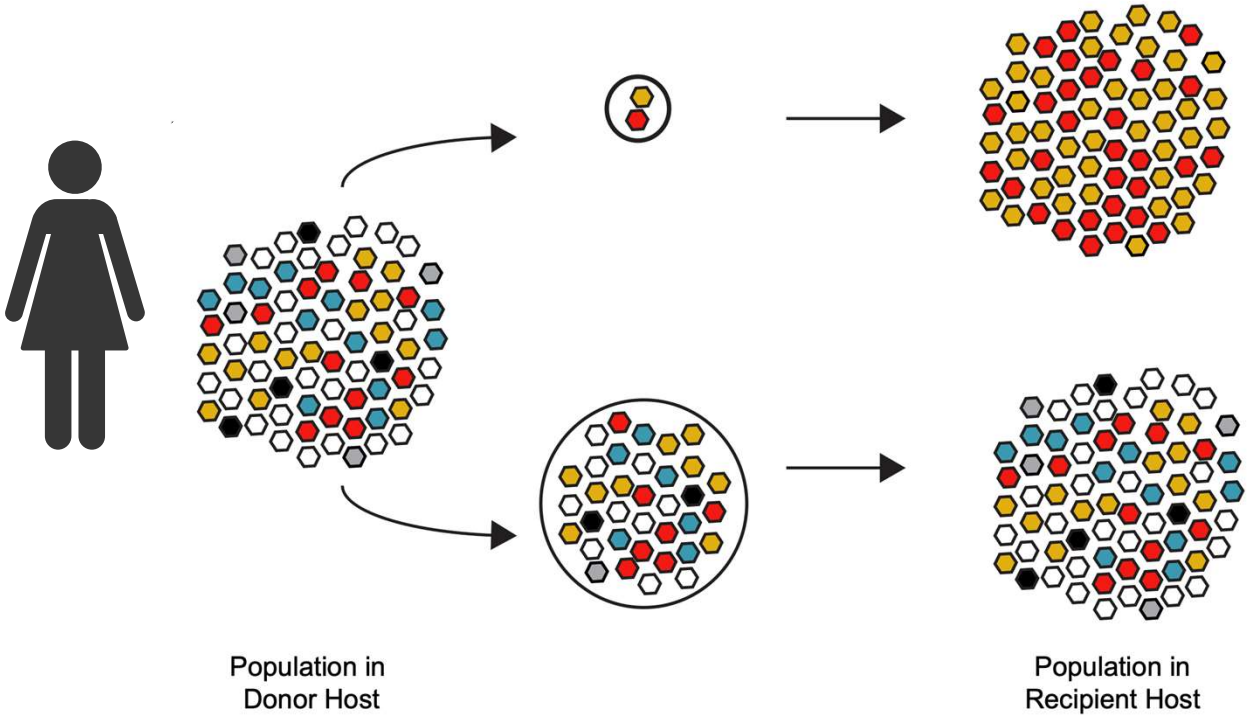
HA: Vaccines and clusters by year (2000/01 to 2011/12)



6 = dominant cluster, (6) = cluster containing vaccine

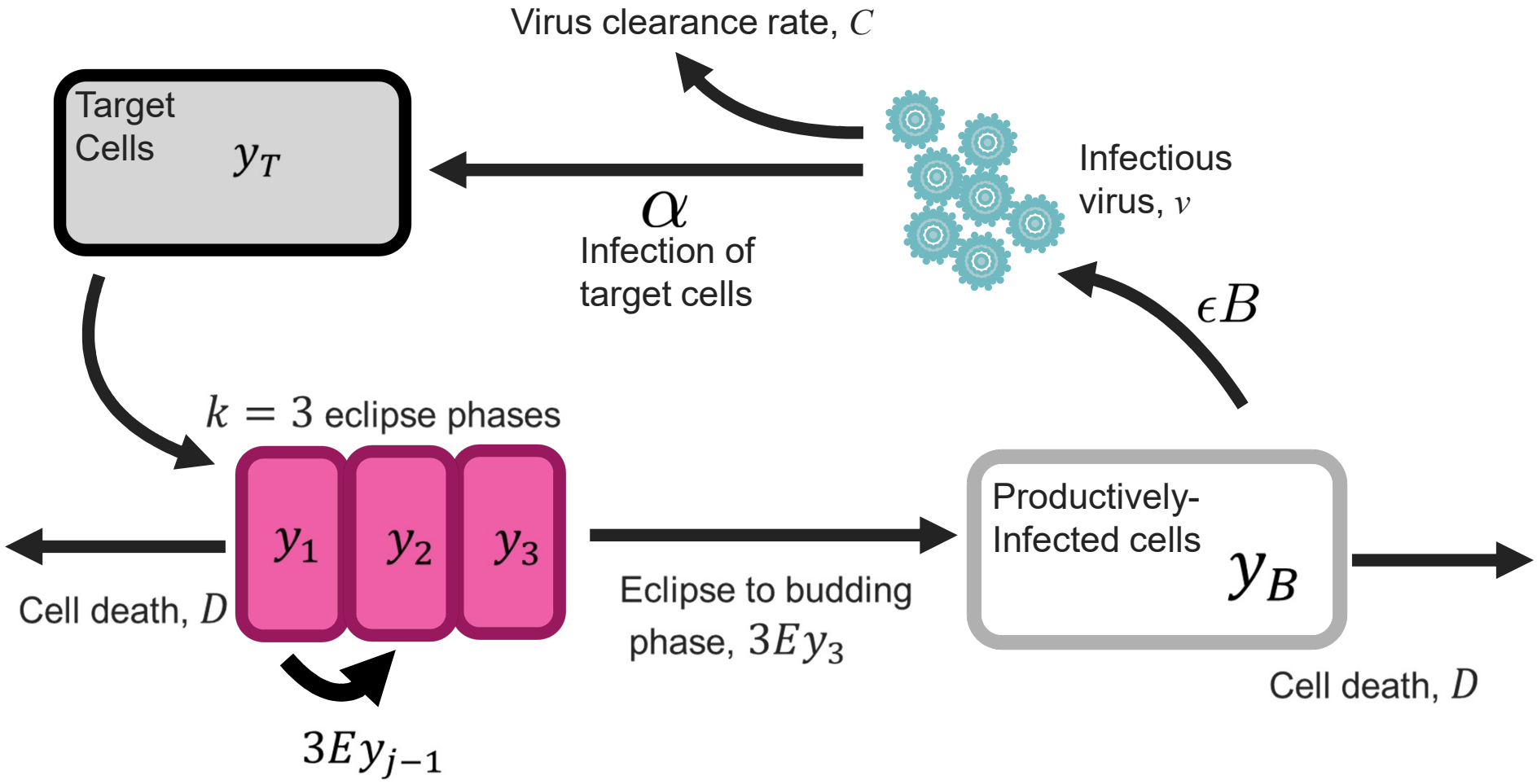
Background: Transmission bottleneck size

- The number of viral particles transmitted from one host to another.
  - It is an important factor in determining how the evolutionary dynamics of the population play out, restricting the extent to which the evolved diversity of the population can be passed from one host to another



J.T. McCrone *et al. Current Opinion in Virology,* 28 (Feb), 2018

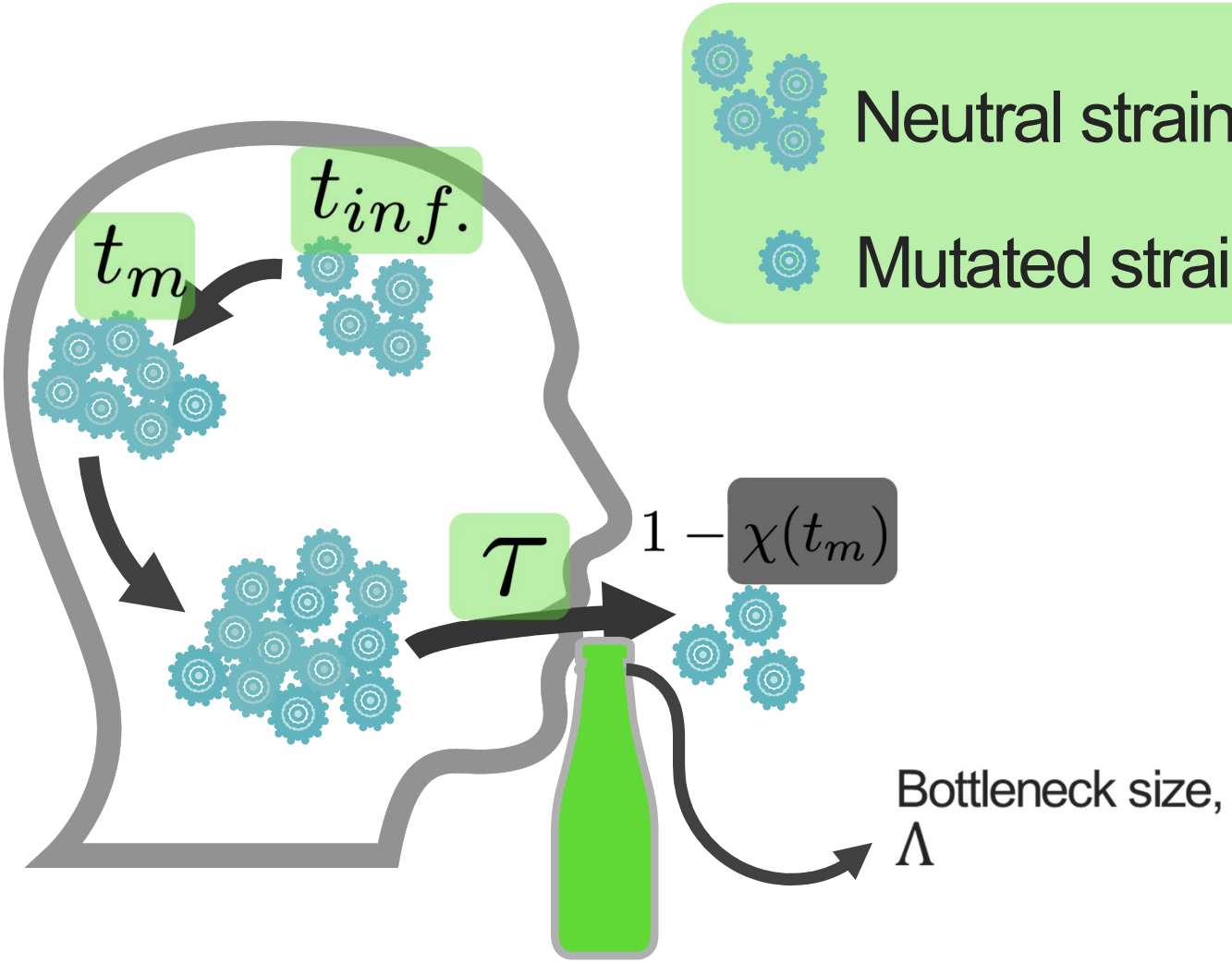
Estimate SARS-CoV-2 Within-host dynamics



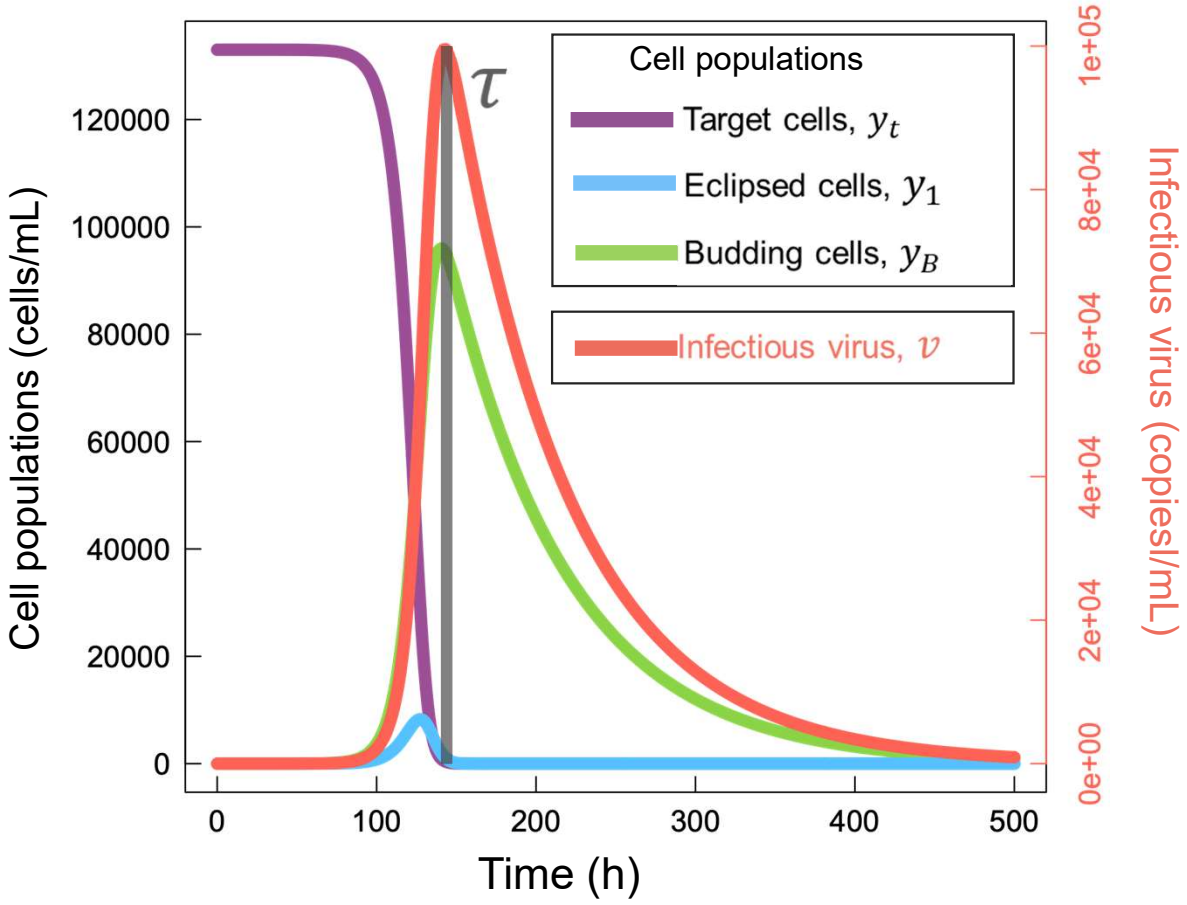
Stochastic life-history bottleneck model

Input

Output

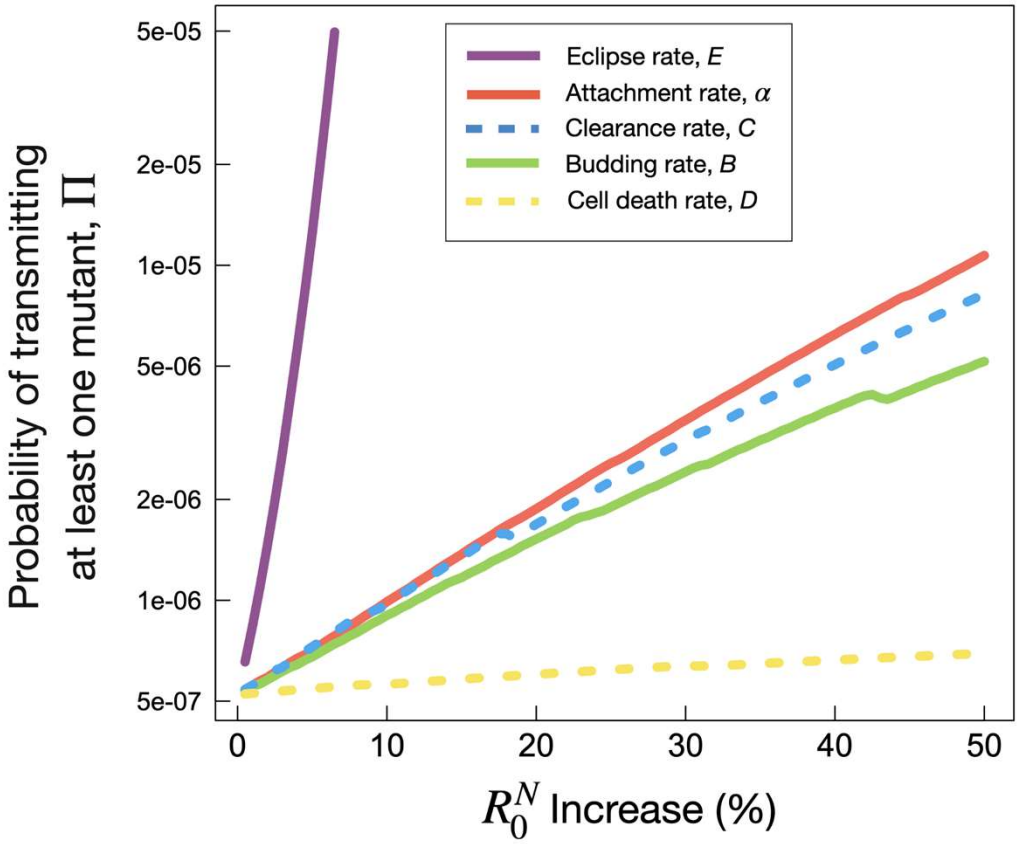


Estimate SARS-CoV-2 Within-host dynamics



What parameter space governs immune escape?

- To achieve a selective coefficient increase of 50%, mutations in C are highly favourable for transmission as compared to mutations in D



Increases by a factor of 15.1

Increases by a factor of 1.3

Epi Model Accounting for Many Strains

- Model with n strains has 2n equations

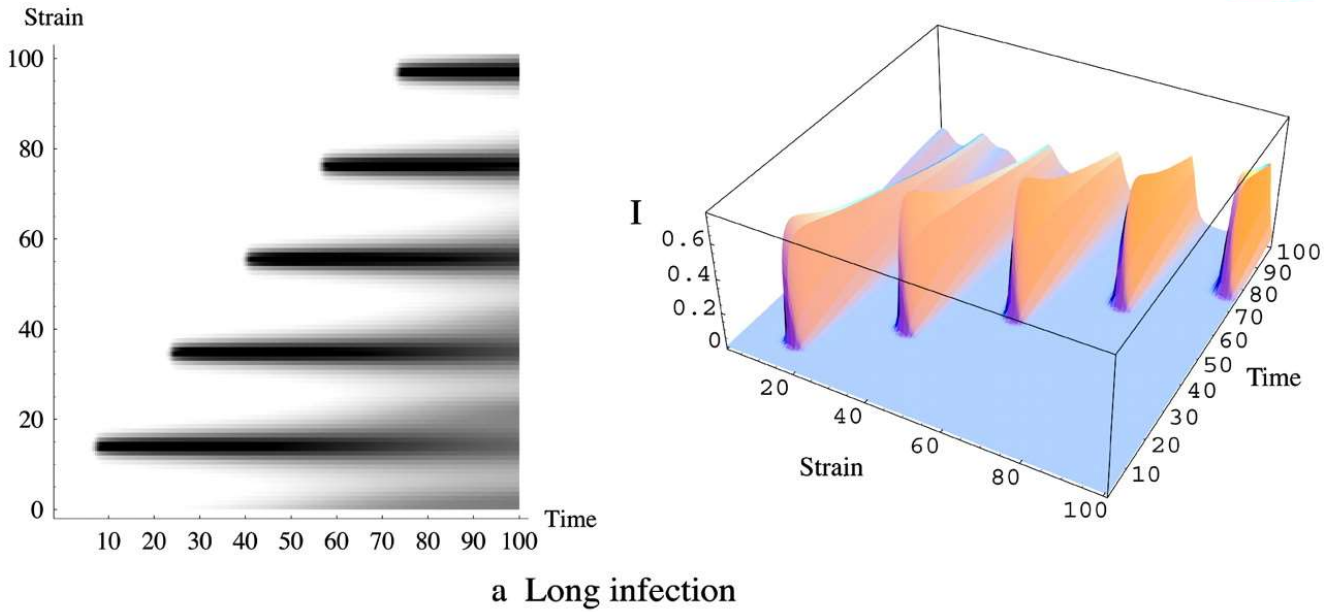
$$\frac{dS_i}{dt} = \nu - \sum \beta_j c_{ij} S_i I_j - \mu S_i,$$

$$\frac{dI_i}{dt} = \beta S_i I_i - \gamma I_i - \mu I_i - \varepsilon I_i + \frac{1}{2}\varepsilon I_{i+1} + \frac{1}{2}\varepsilon I_{i-1}$$

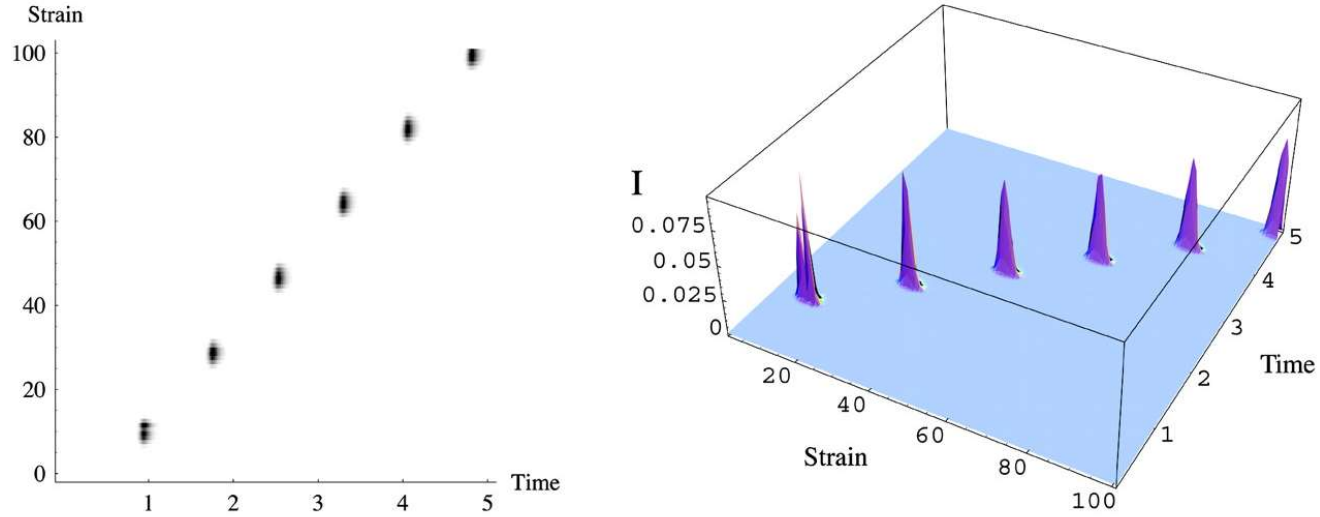
- where $0 \leq c_{ij} \leq 1$ changes susceptibility, and $c_{ii} = 1$
- Note: susceptibility classes are not mutually exclusive
- Gog/Grenfell: strains on 1D line, spontaneous mutation (ϵ)

$$c_{ij} = \exp(-A[i-j]^2)$$

- Numerical results for a linear strain space. Time is in years in both examples; d = 10, β = 3v. In a, μ = 1, v = 0.5, m = 0.02. In b, μ = 1/80, v = 52, m = 0.1.
- (a) For a long infection, clusters arise sequentially and persist, so that at a given time there are several clusters at high prevalence.
- (b) For a short infection, the clusters are narrow and appear only for a short time before vanishing again. There is at most one cluster at any given time.



a Long infection



b Short infection