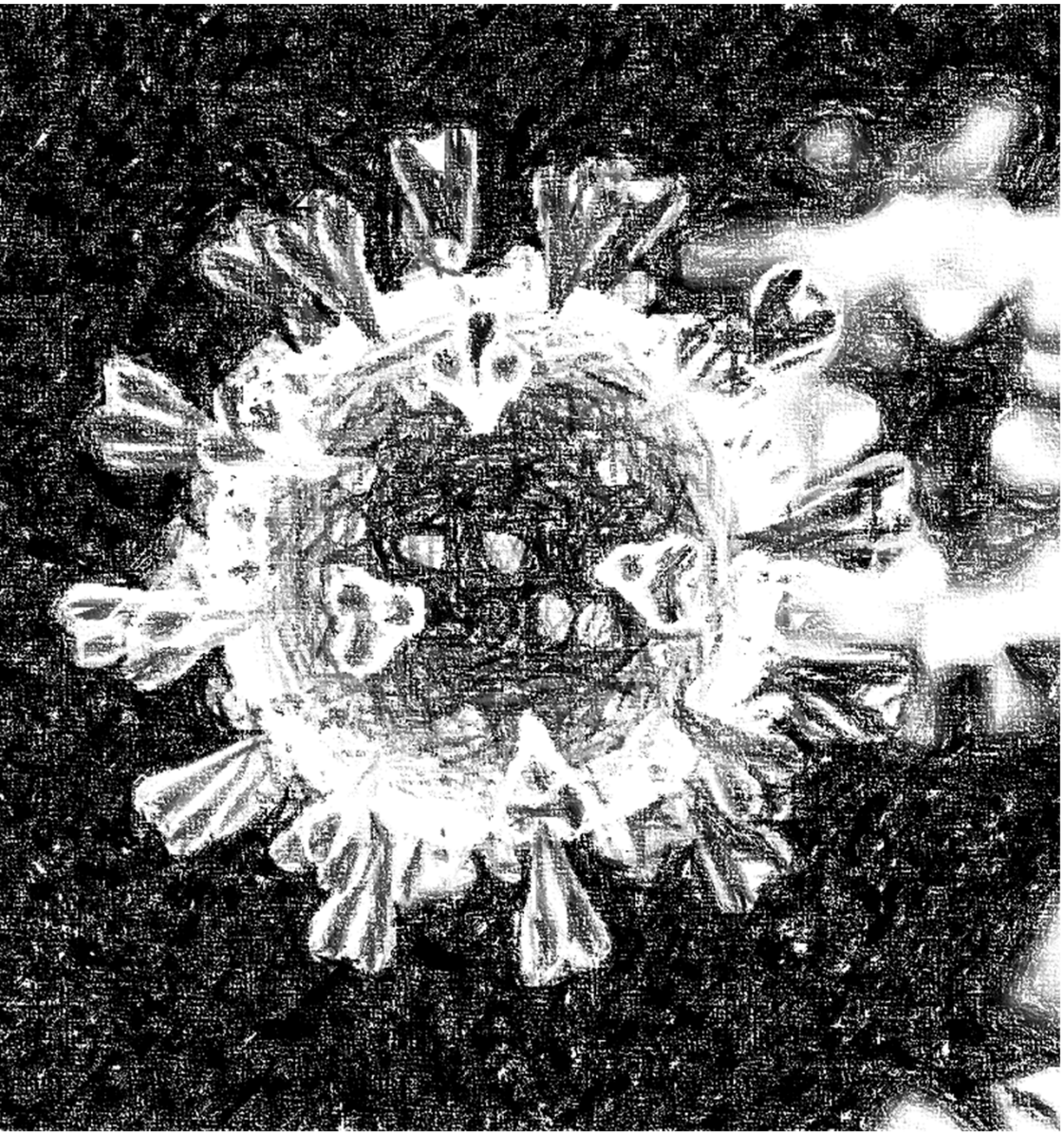
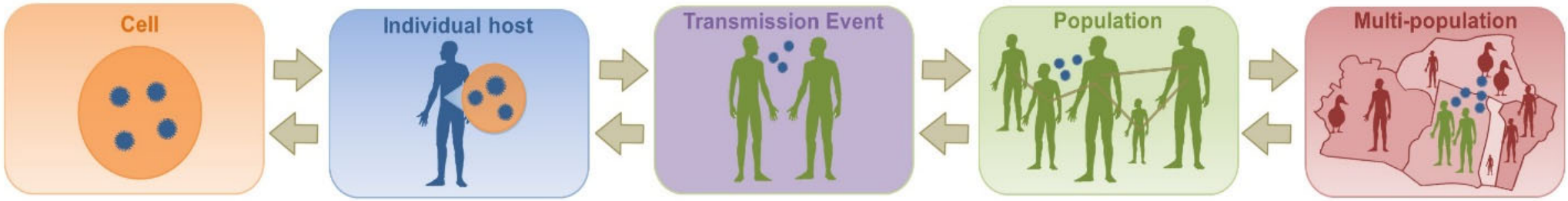


Virus evolution between transmission pairs

Katia Koelle
Department of Biology, Emory University

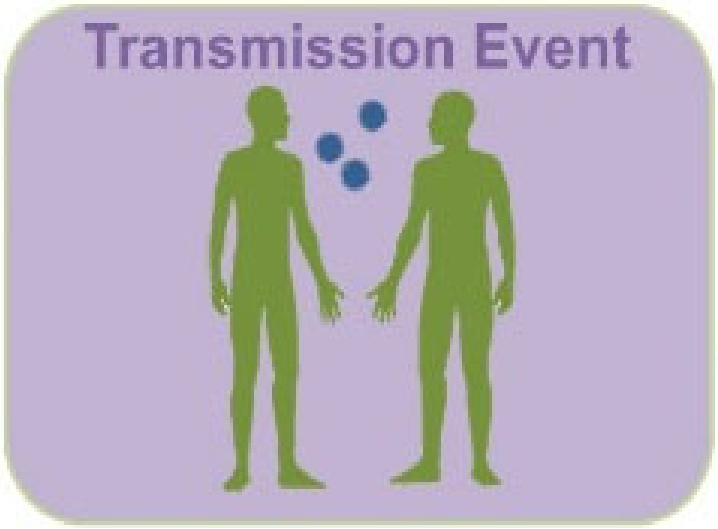
NORDITA
June 12, 2023

NORDITA program: Unifying the epidemiological and evolutionary dynamics of pathogens (Workshop 2: Evolutionary dynamics)

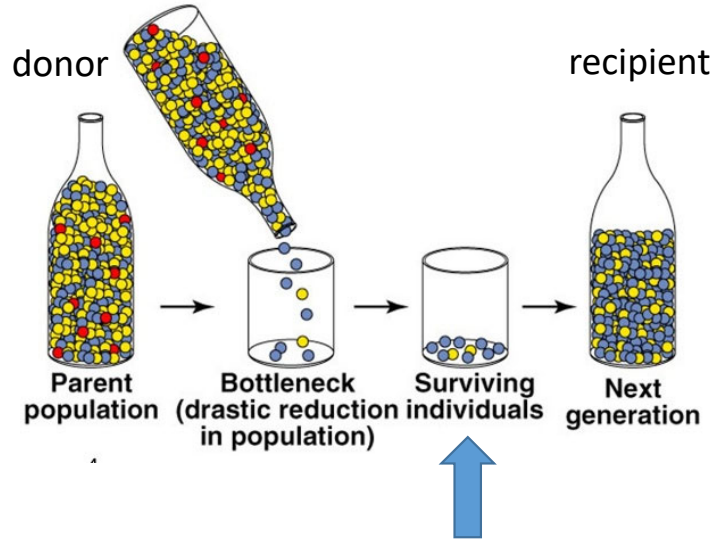


- Viral population dynamics occur across a range of biological scales
- Viral evolutionary dynamics occur across this same range of biological scales
- Patterns of viral evolution at these scales can also tell us about viral population dynamics at these scales

What are transmission bottlenecks?



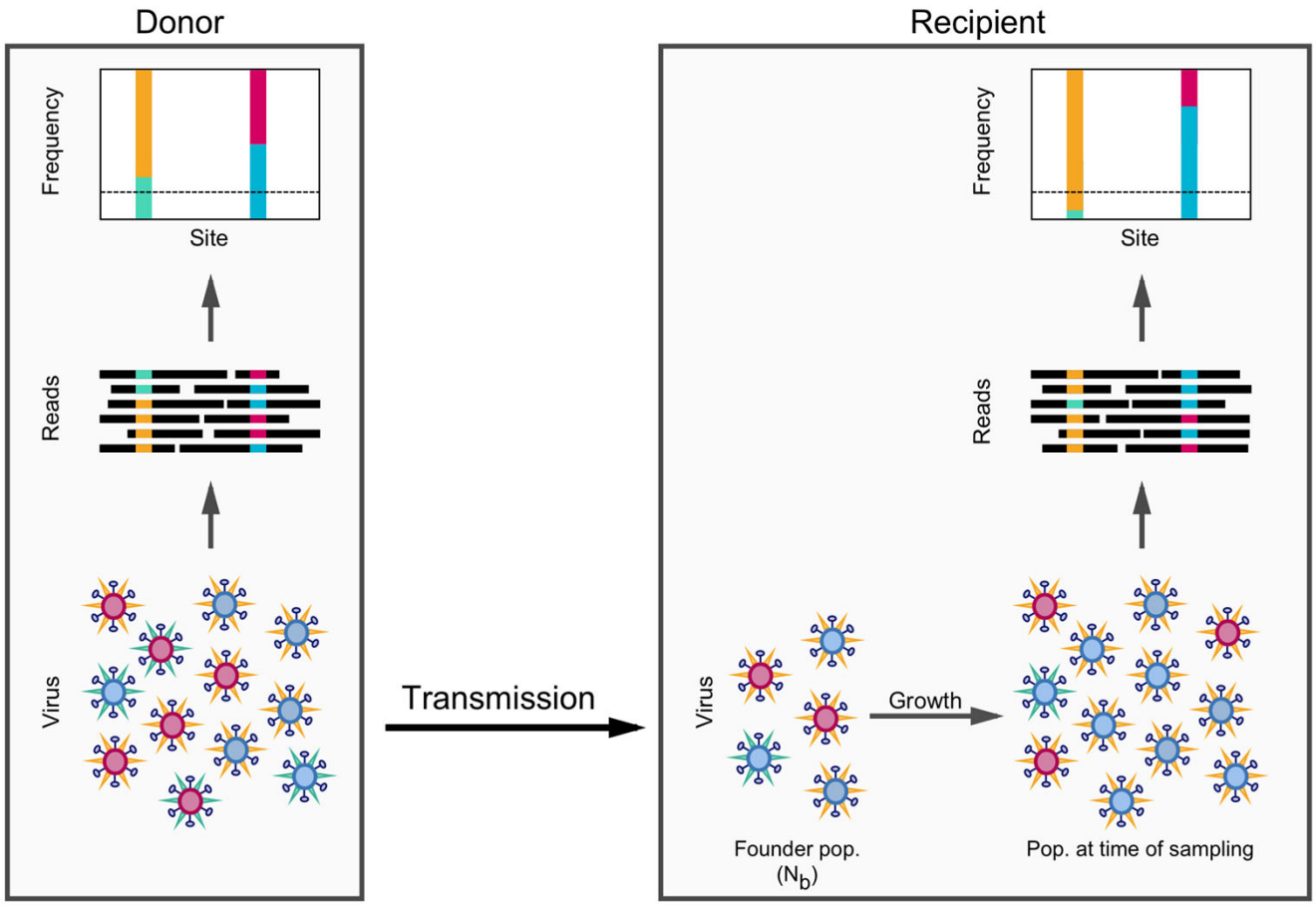
donor recipient



Transmission bottleneck size (N_b) = # of virions that go on to establish a genetic lineage in an infected individual = population bottleneck

- Not # of virions that "fall on" the recipient
- Not unique # of virus genotypes that establish infection = this is the genetic bottleneck

How are their magnitudes being estimated in natural infections?



Presence/absence method

Binomial sampling/Wright-Fisher single generation model

Beta-binomial approach

Sobel-Leonard et al. (2017) *JVI*

Haplotype-based extension of beta-binomial approach

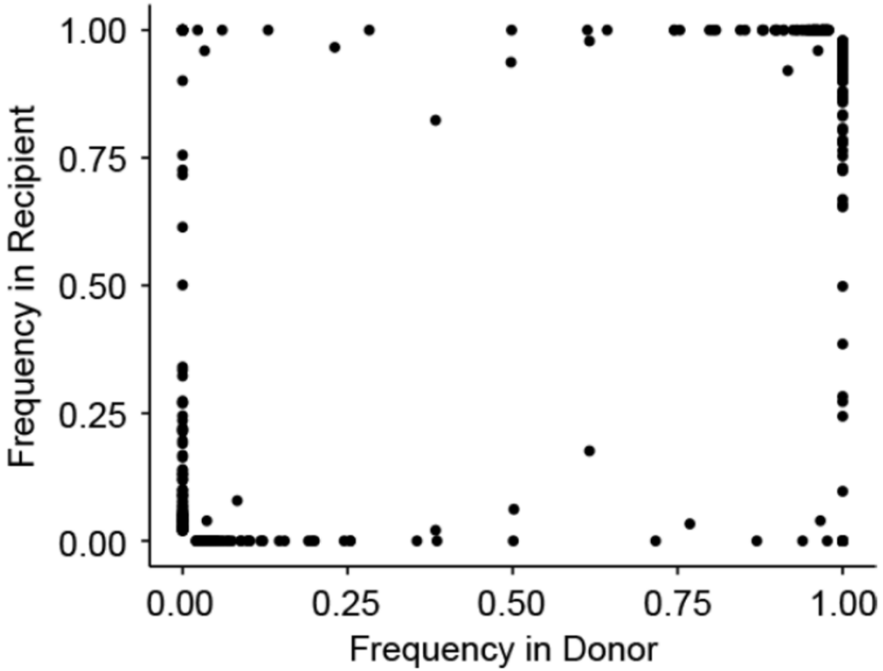
Ghafari et al. (2020) *JVI*

All existing approaches rely exclusively on the subset of loci in the donor that are polymorphic

Sobel-Leonard et al. (2017) *JVI*

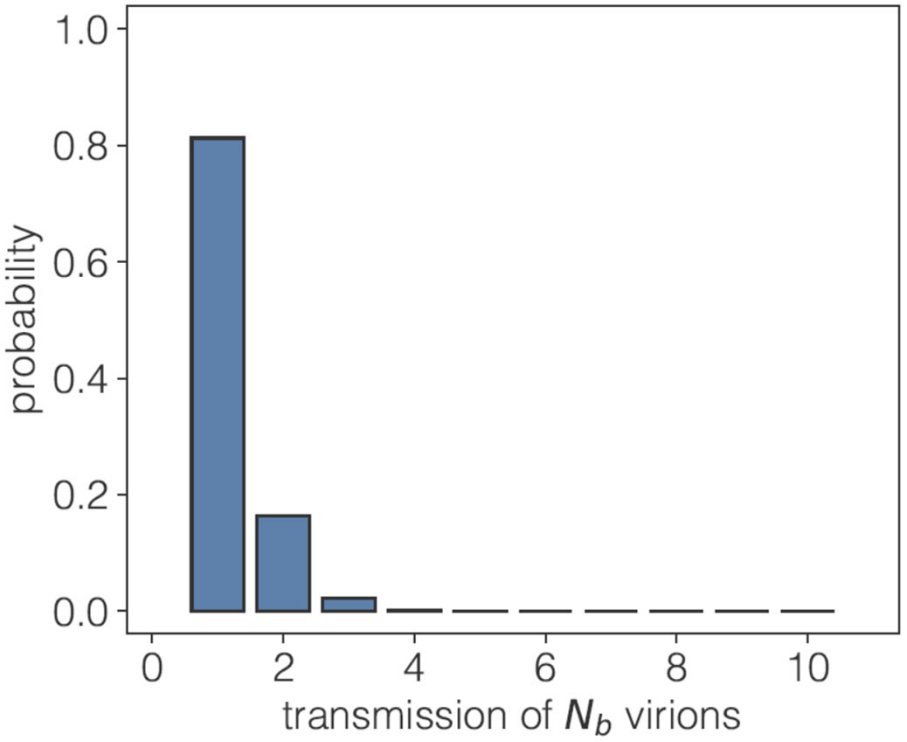
Viral sequence data point towards very tight transmission bottlenecks in acutely-infecting respiratory viruses

Influenza A virus
43 transmission pairs
N_b = 1-2 viral genomes



McCrone et al. (2018) *eLife*

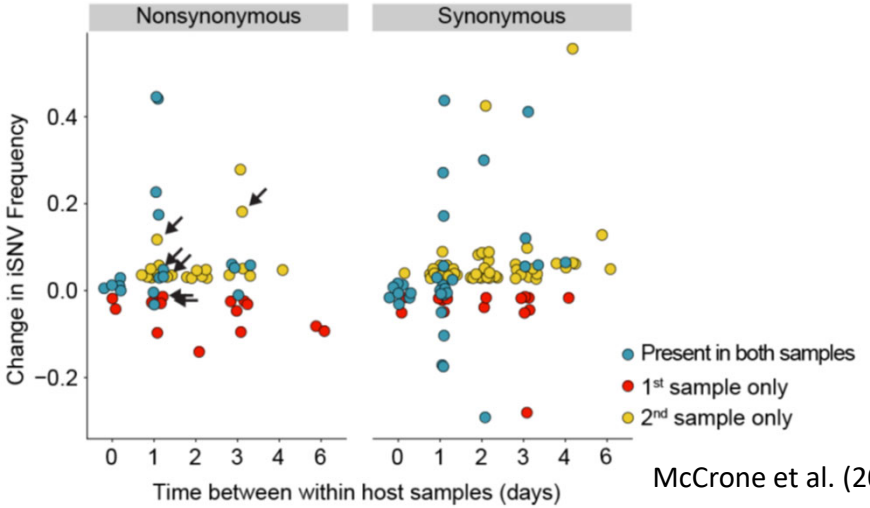
SARS-CoV-2
13 transmission pairs
N_b = 1-3 viral genomes



Martin & Koelle (2021) *Science Translational Medicine*

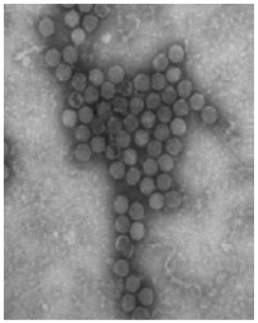
What can go (really) wrong with existing methods:

Sampling time from donor is not identical to the time of transmission, and iSNVs rapidly come and go due to very low within-host N_E



McCrone et al. (2018) *eLife*

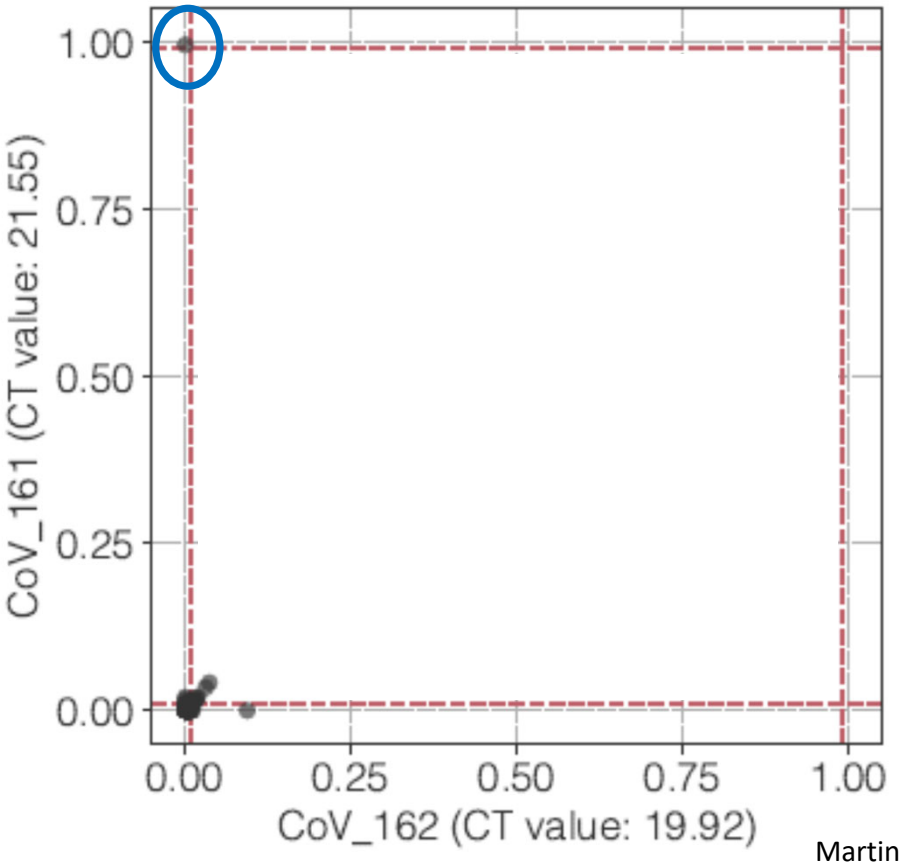
Nonrandom sampling from donor is likely due to, e.g., viral aggregation



Pradhan et al. (2022) *Viruses*

Estimating transmission bottleneck size based on methods that rely on genetic variation present in a donor can vastly underestimate bottleneck sizes

Estimating transmission bottleneck sizes from *de novo* genetic variation observed in recipients of transmission pairs



Clonal variants: variants that are absent from a donor and fixed in a recipient

Number of clonal variants

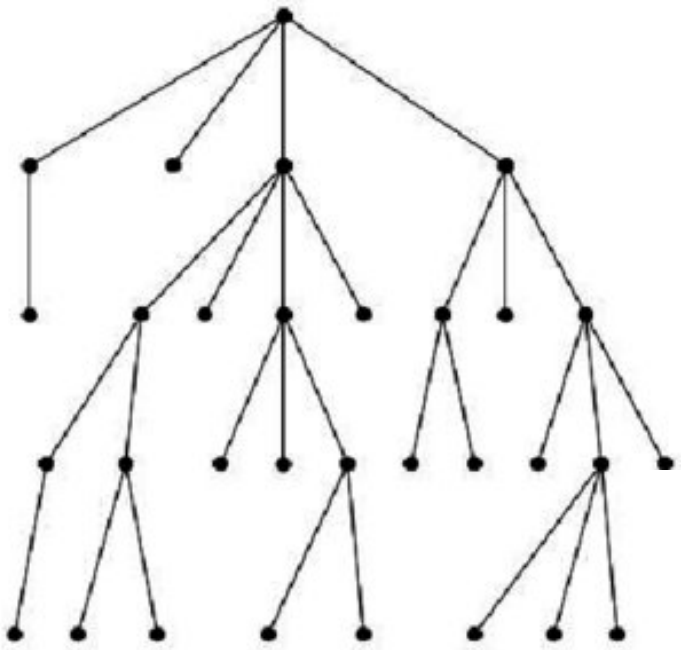
Influenza A virus data (McCrone et al. (2018) *eLife*):

0 clonal variants:	42 transmission pairs
1 clonal variant:	5 transmission pairs
2 clonal variants:	3 transmission pairs
3 clonal variants:	2 transmission pairs
4+ clonal variants:	0 transmission pairs

Martin & Koelle (2021) *Science Translational Medicine*

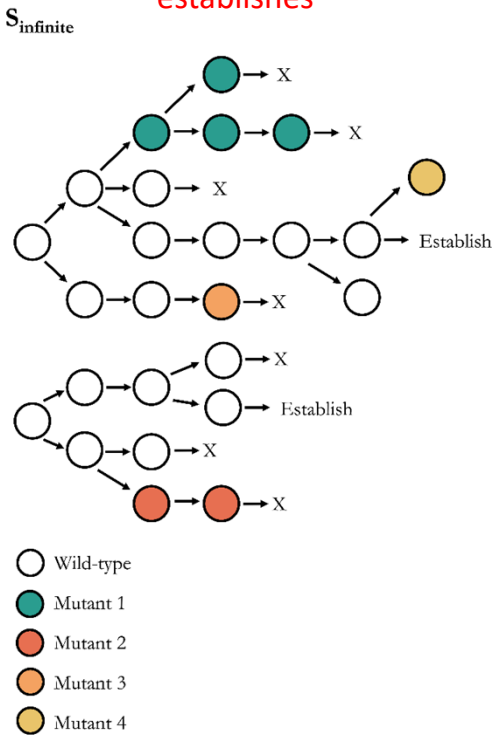
Branching process model

- Initial viral population size = N (wild-type viral particles)
- Each viral particle has a geometric offspring distribution with mean offspring number = within-host R_0
- mutations occur during the 'birth' of a viral particle: Poisson distribution with mean μ

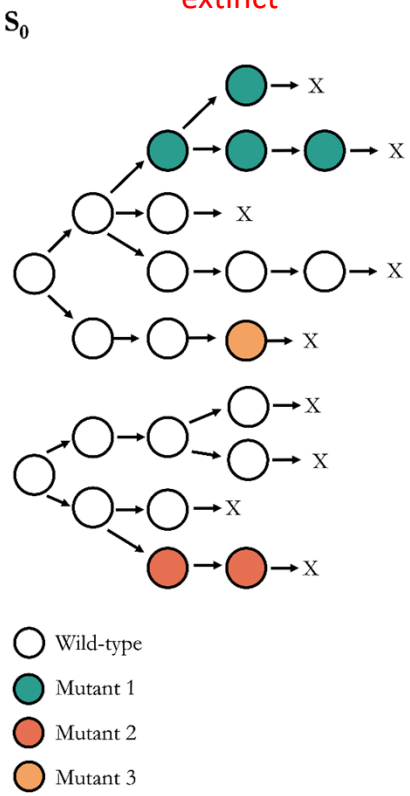


Approach (initial number of viral particles *N* = 2)

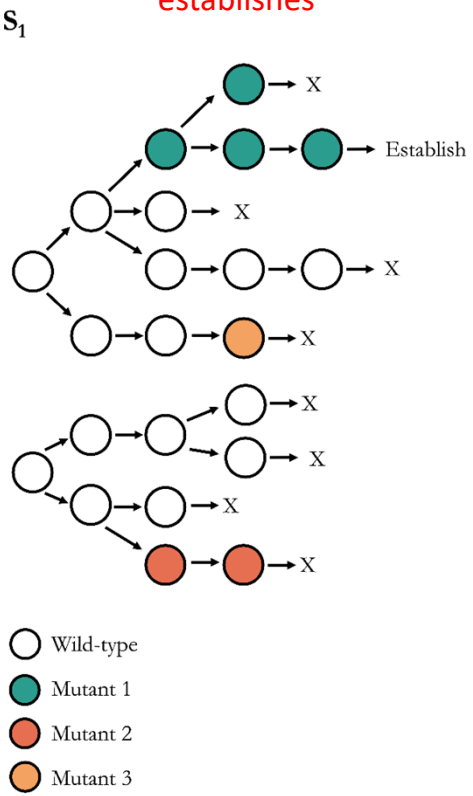
Wild-type viral population establishes



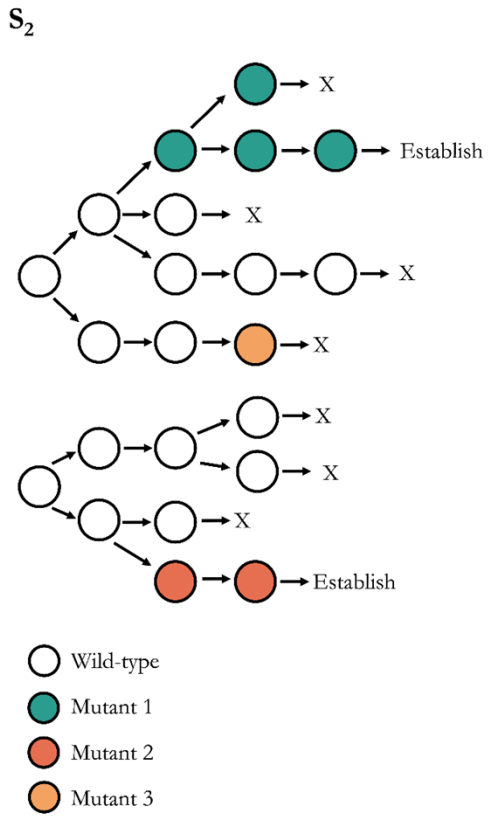
Viral population goes extinct



A single mutant lineage establishes

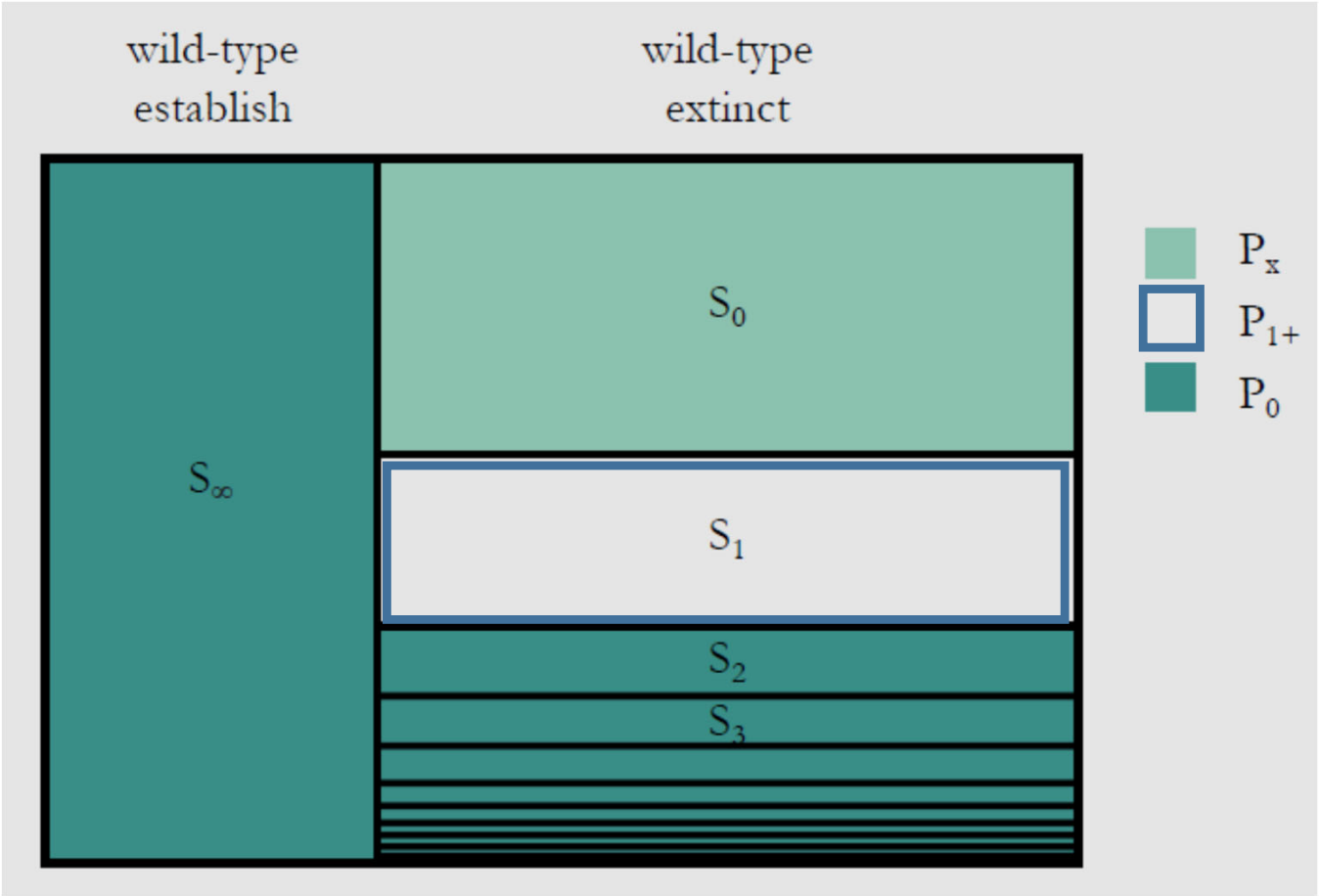


2(+) mutant lineages establish



s_k: *k* is the number of mutant lineages that establish in the recipient

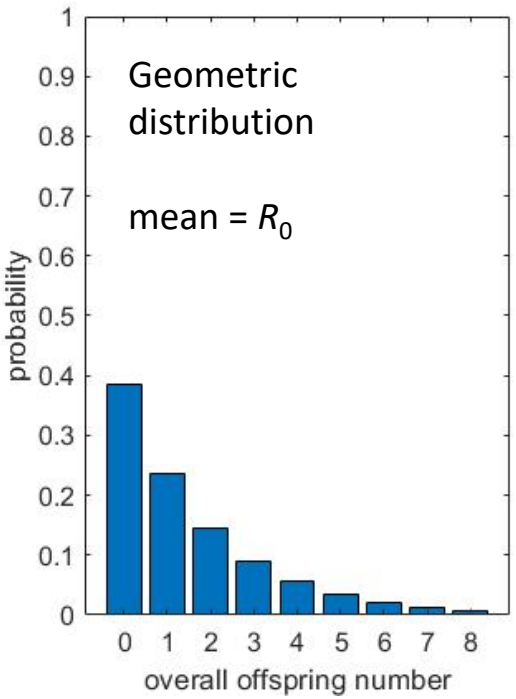
Possible outcomes



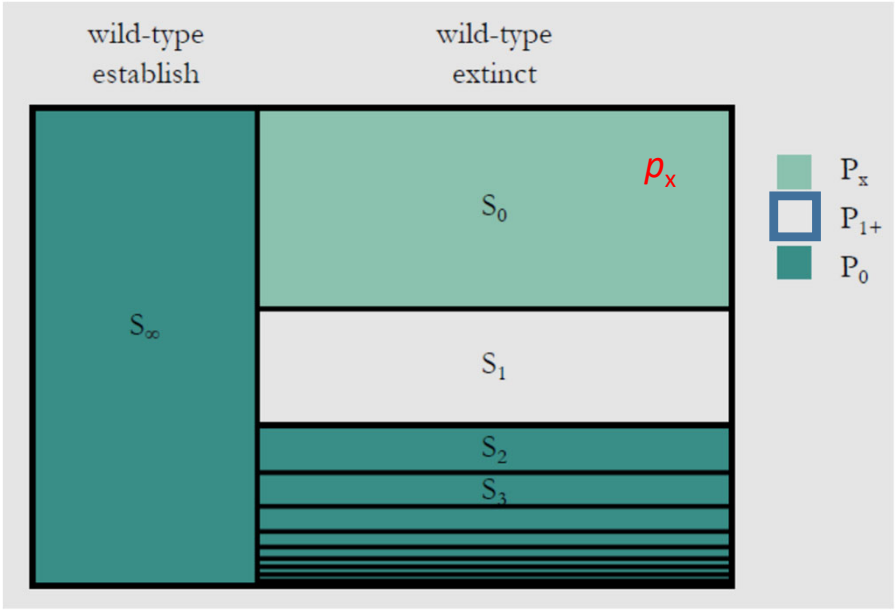
Goal:

Derive pmf for # of clonal variants, given a within-host R_0, N, and μ
Use this pmf to calculate likelihood of (N, μ), given within-host R_0

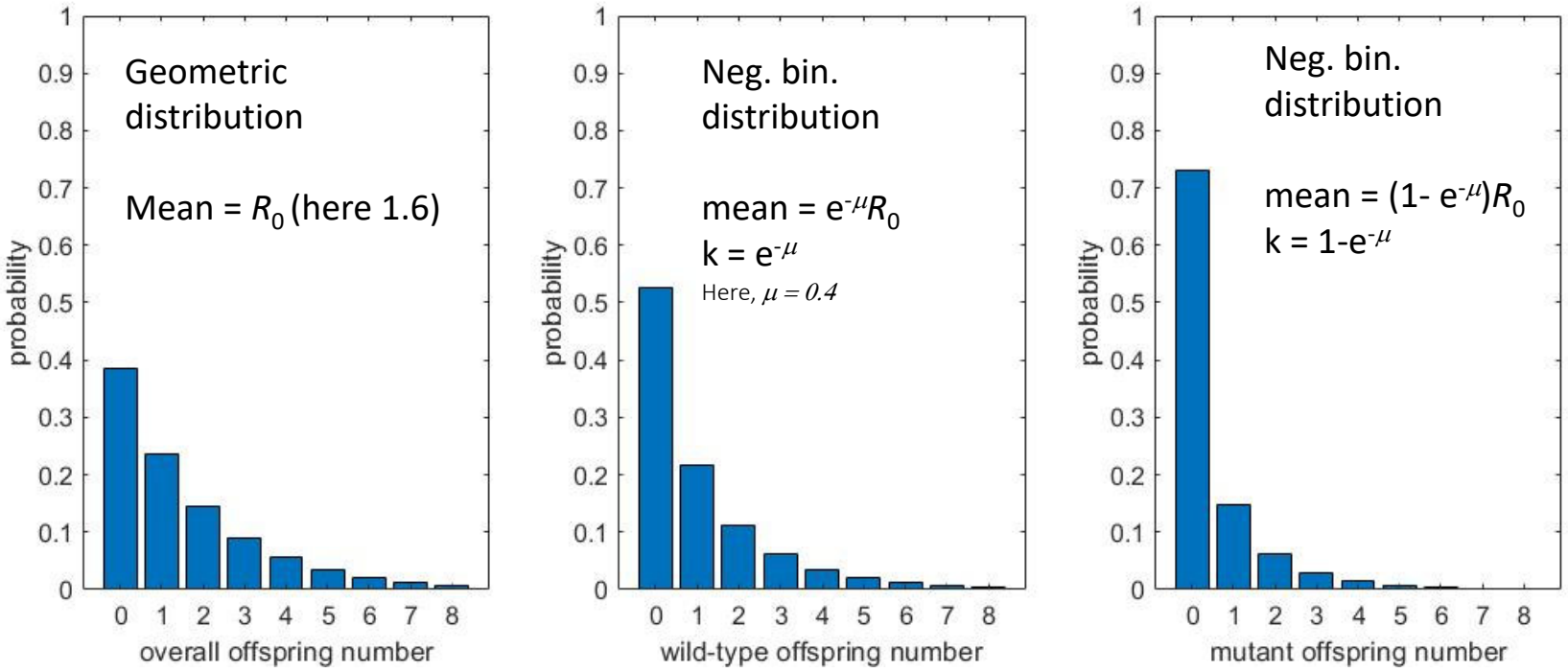
Calculating probability of extinction



$$p_x = 1-(1/R_0)^N$$

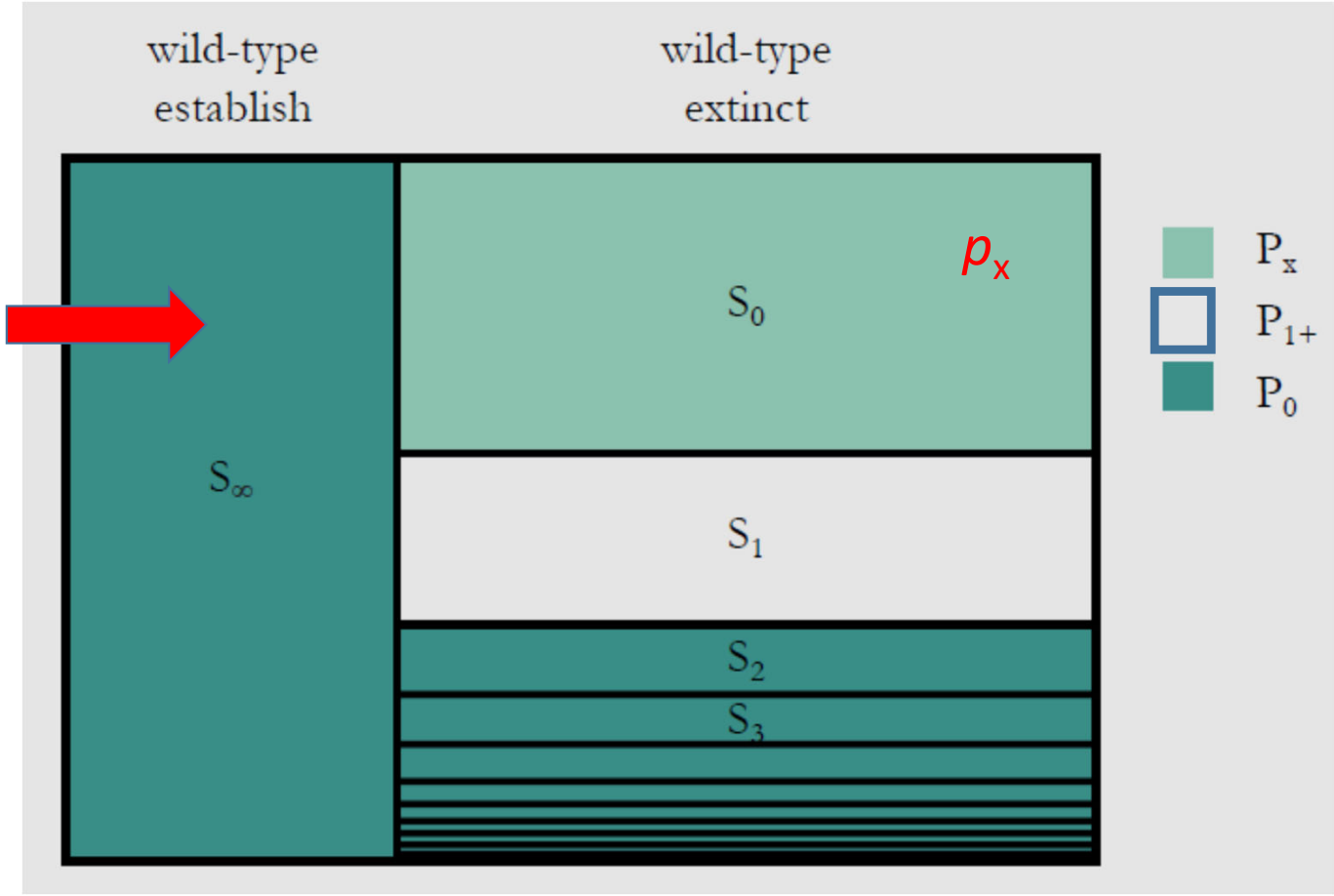


Offspring distribution of wild-type particles

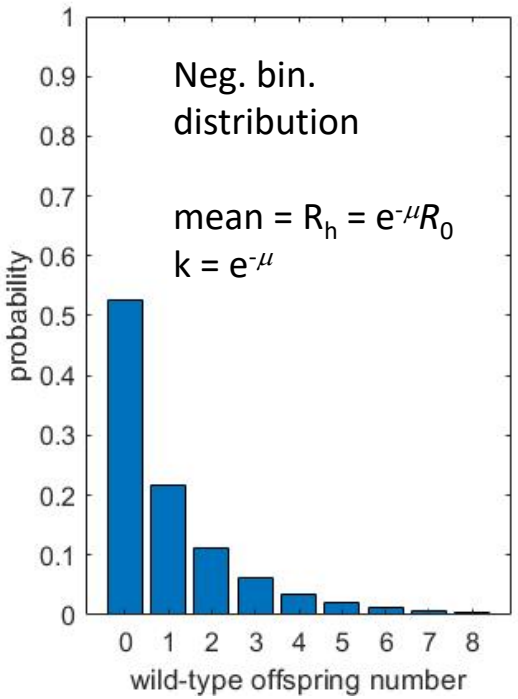


Possible outcomes

Calculate using pgf for branching process model with wild-type offspring distribution (neg. bin.)



Minor outbreak final size distribution of wild-type particles



Journal of Theoretical Biology 294 (2012) 48–55

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Estimating the transmission potential of supercritical processes based on the final size distribution of minor outbreaks

Hiroshi Nishiura [a,b,*], Ping Yan [c], Candace K. Sleeman [d], Charles J. Mode [d]

[a] School of Public Health, The University of Hong Kong, Level 6, Core F, Cyberport 3, 100 Cyberport Road, Pokfulam, Hong Kong
[b] PRESTO, Japan Science and Technology Agency, Saitama 332-0012, Japan
[c] Public Health Agency of Canada, Ottawa, Canada
[d] Department of Mathematics, Drexel University, Philadelphia, PA, USA

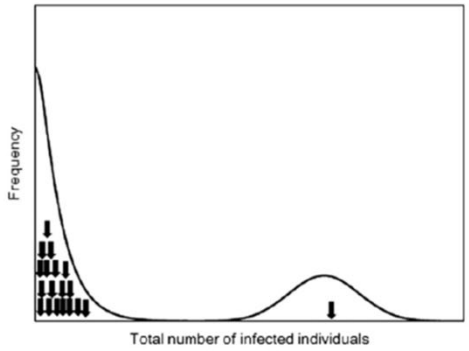


Fig. 1. The distribution of final epidemic size. The final size distribution of a homogeneously mixing population usually exhibits a bimodal shape where the first peak represents minor outbreaks and the second peak corresponds to the mode of major epidemics. Arrows indicate the available evidence for the final size of pneumonic plague (18 minor outbreaks and 1 major epidemic).

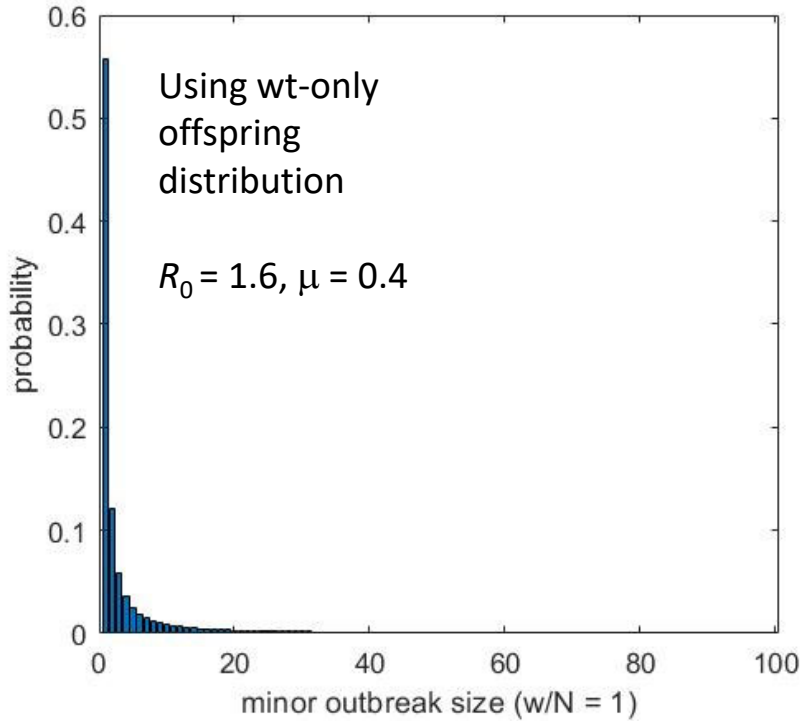
For $y=1$,

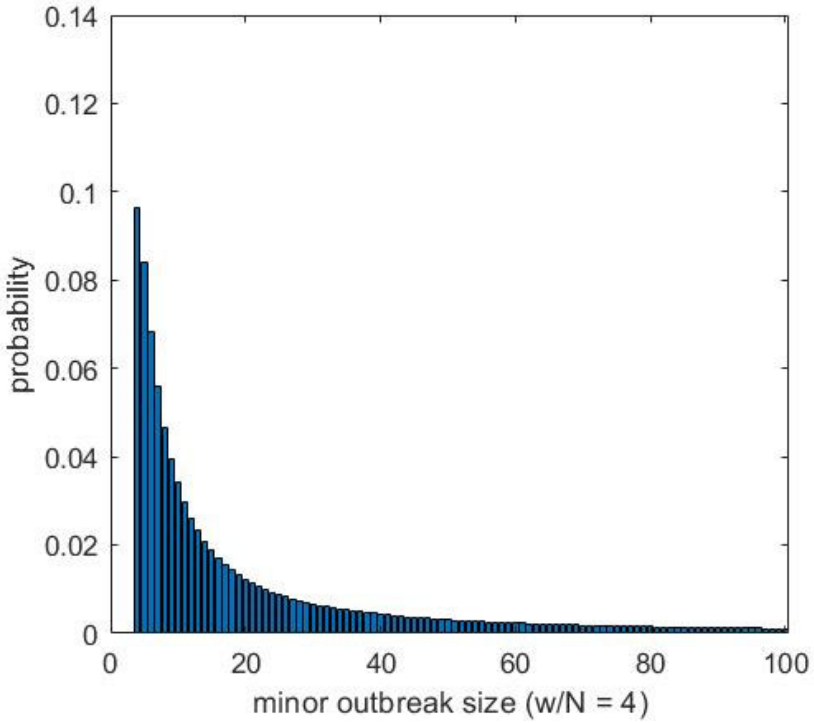
$$p_1 = \frac{1}{(1+(R_h/k))^k}$$

For $y \geq 2$, the distribution is recursively calculated as (Yan, 2008)

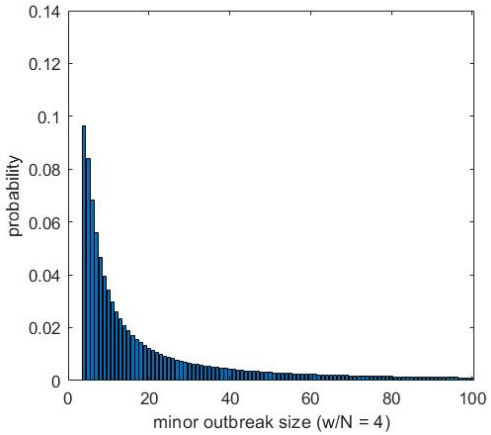
$$p_y = \frac{1}{y!}\frac{d^y}{ds^y}g(s)\Big|_{s=0} = \frac{\Pi_{j=0}^{y-2}((j/k)+y)}{y!}\left(\frac{k}{R_h+k}\right)^{ky}\left(\frac{R_h k}{R_h+k}\right)^{y-1}$$

Minor outbreak final size distribution of wild-type particles

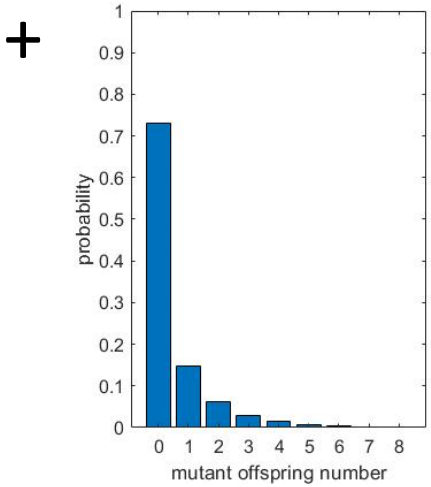


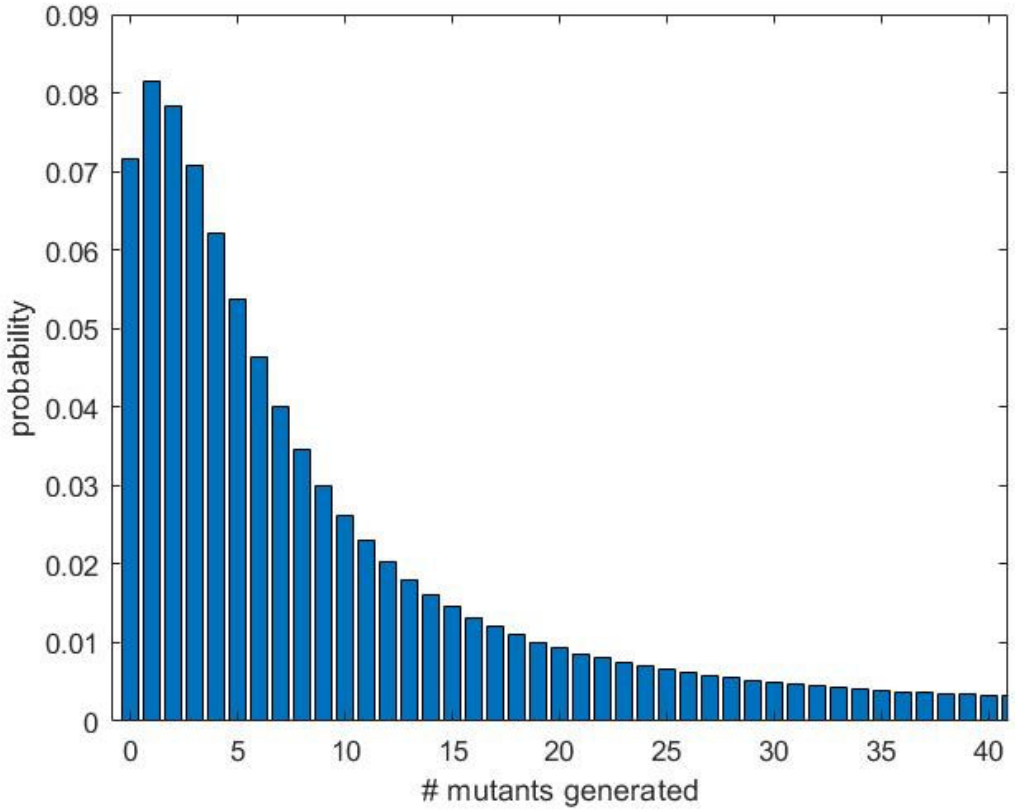


Distribution of the number of mutant offspring generated

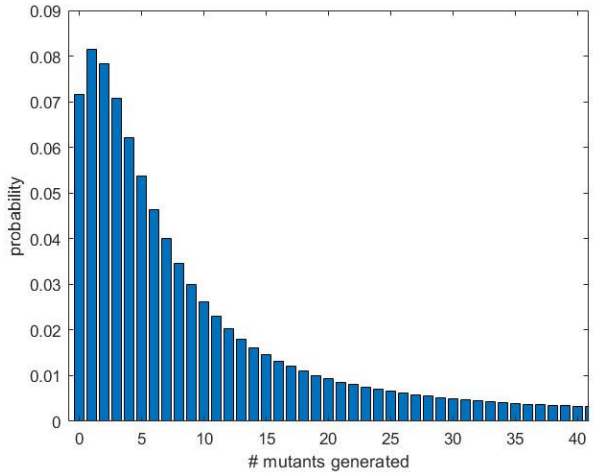


+



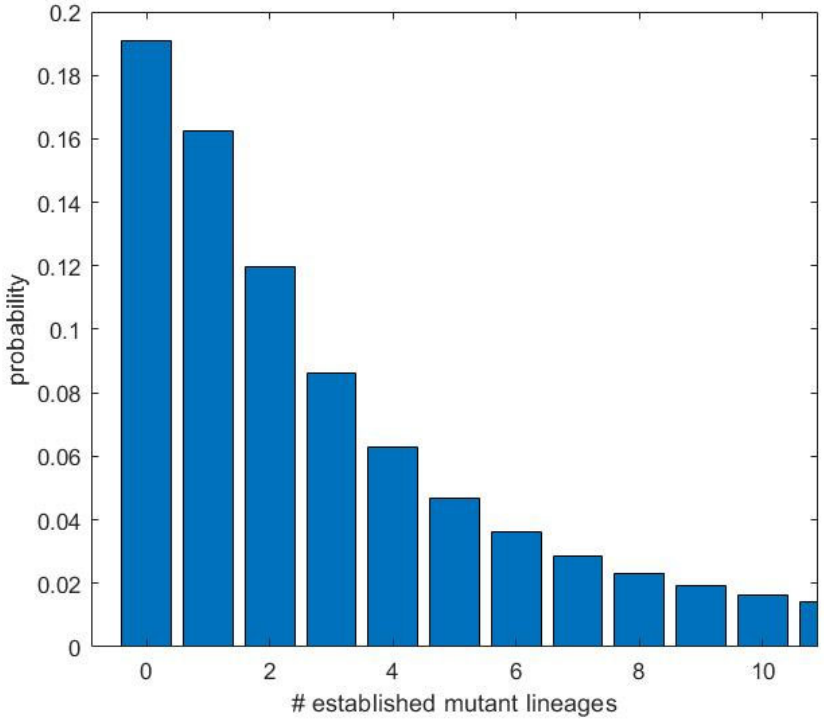


Distribution of the number of mutant lineages that establish

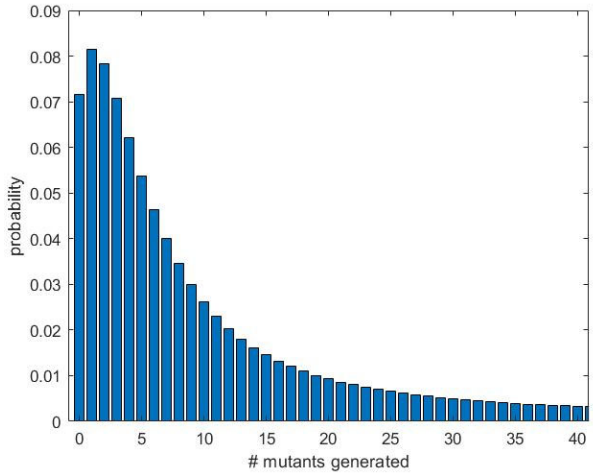


+

prob. of a given generated mutant establishing a lineage: $1-(1/R_0)$

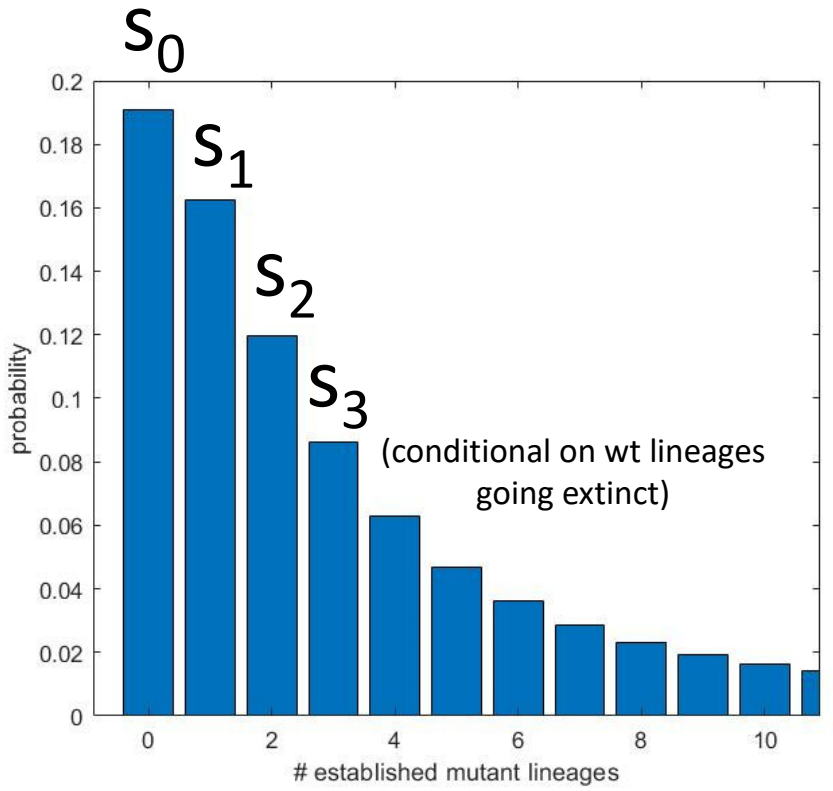


Distribution of the number of mutant lineages that <u>establish</u>

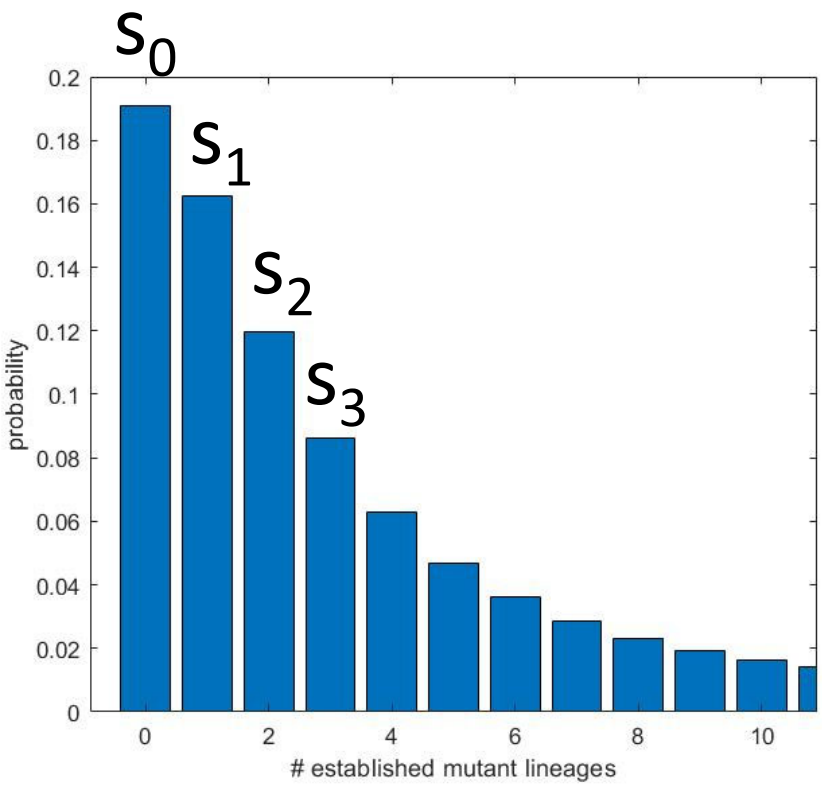


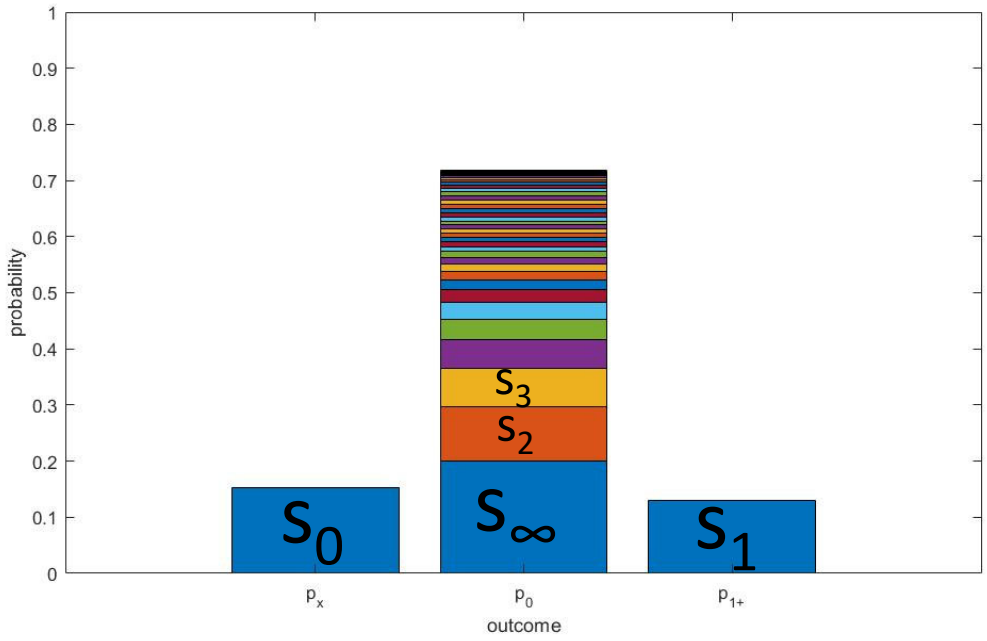
+

prob. of a given generated mutant establishing a lineage:
$1-(1/R_0)$



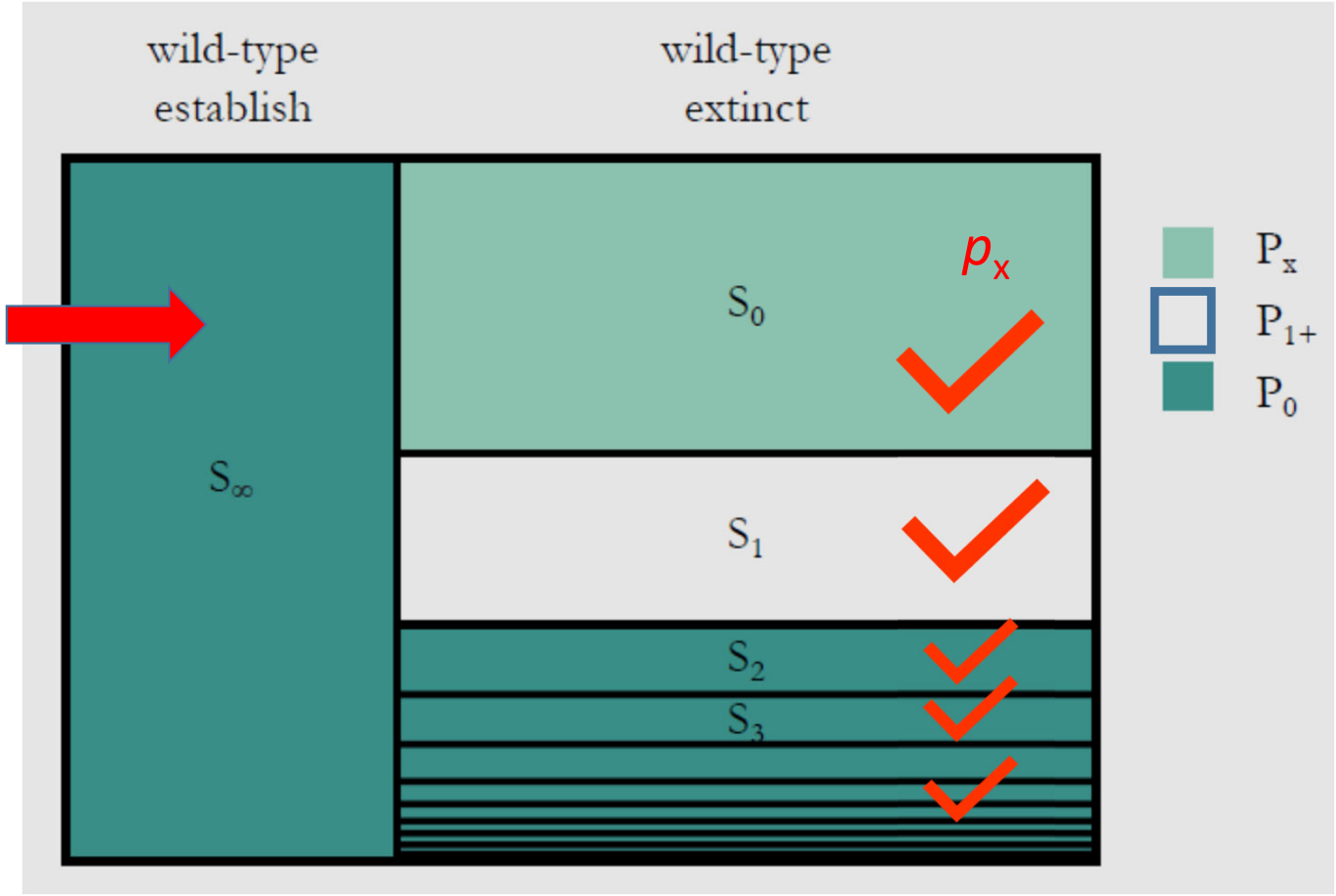
Outcomes


S_0
S_1
S_2
S_3
probability
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
2
4
6
8
10
established mutant lineages


probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
S_3
S_2
S_0
S_∞
S_1
p_x
p_0
p_{1+}
outcome

Possible outcomes

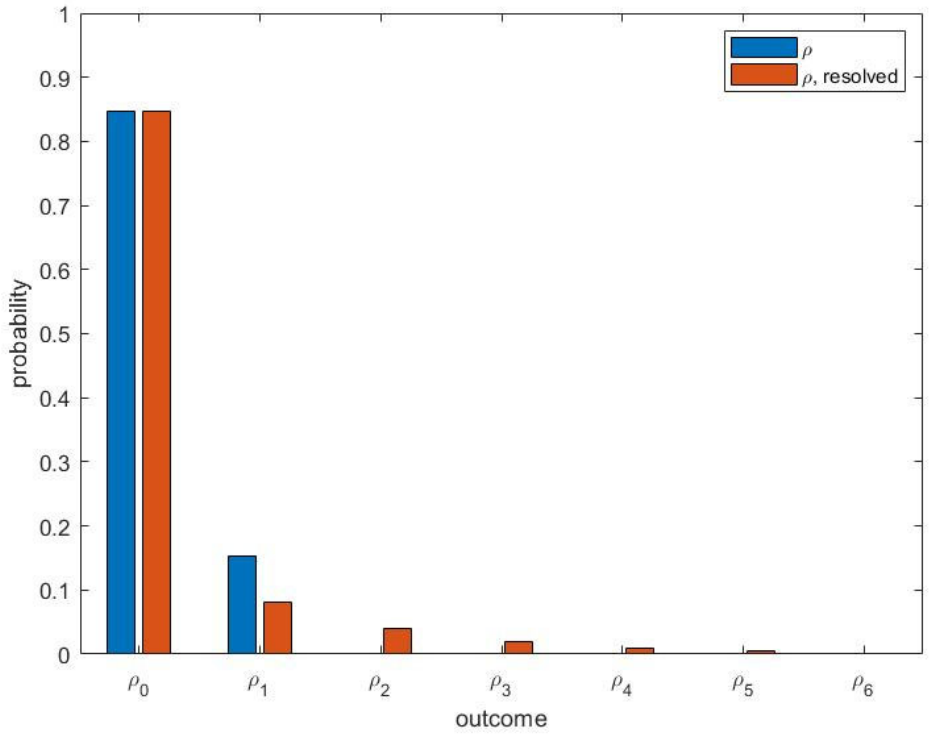
Calculate using pgf for branching process model with wild-type offspring distribution (neg. bin.)



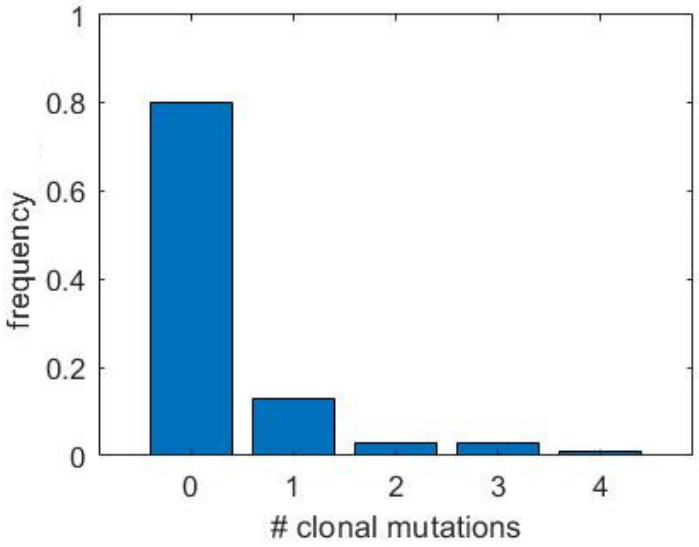
Calculating pmf of number of clonal variants (ρ) and resolving ρ_{1+}

Need to condition on infection of recipient ($1-p_x$) → ρ

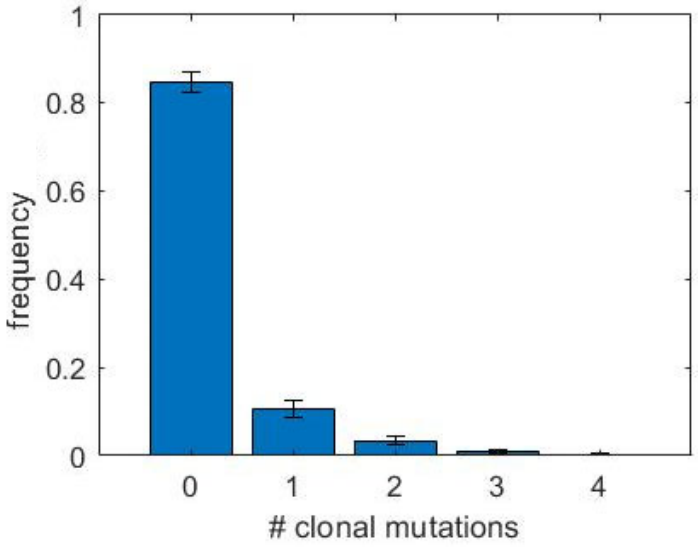
Need to 'resolve' ρ_{1+} into ρ_1, ρ_2, ρ_3, etc.



Testing on simulated data



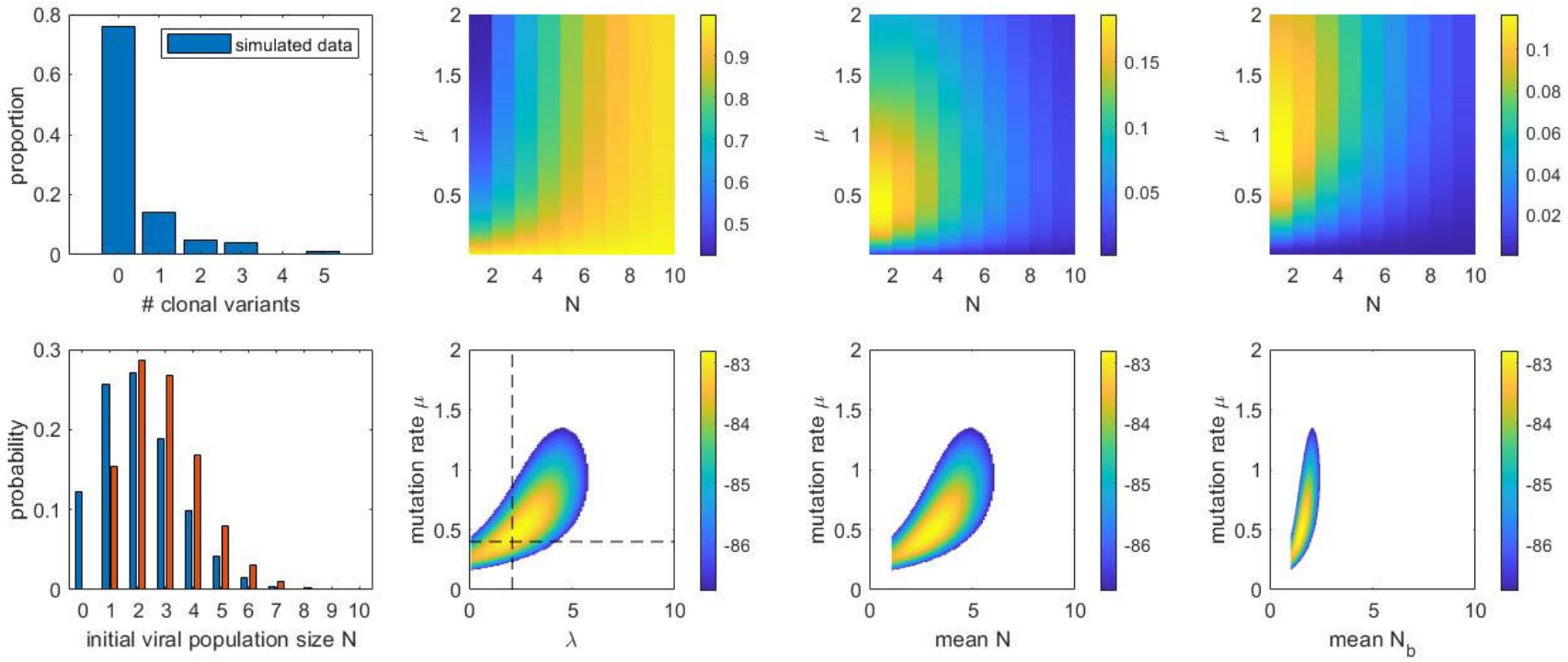
1000 simulated infections
R_0 = 1.6
μ = 0.4
N = 4



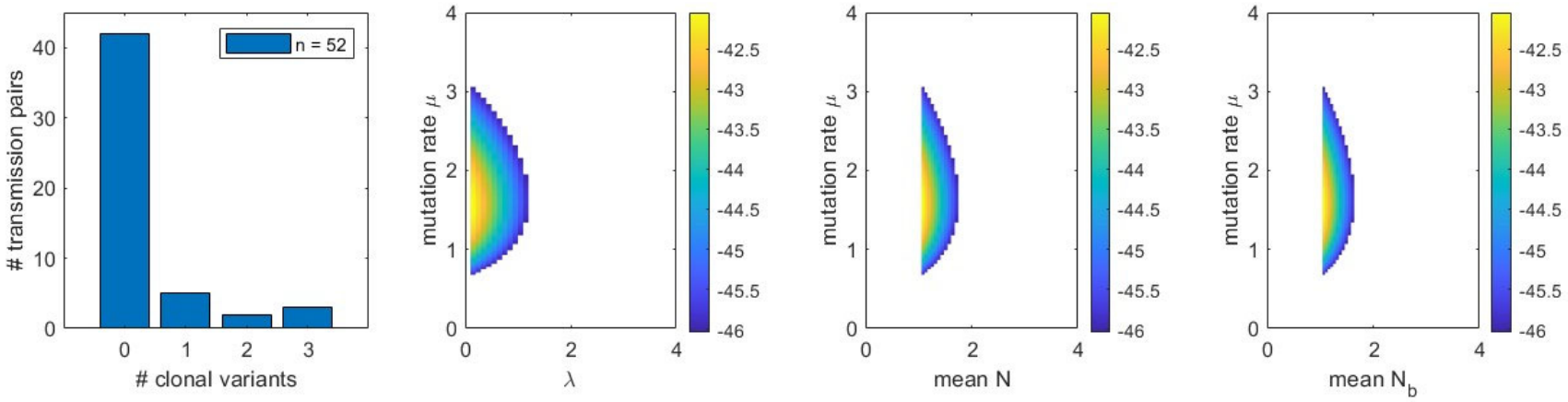
Analytical results
R_0 = 1.6
μ = 0.4
N = 4

Comparison against simulated data

100 simulated infections
R_0 = 1.6
μ = 0.4
λ = 2.1 (initial viral population *N* is Poisson distributed with mean λ)



Estimates for influenza A virus



Assumed within-host R_0 = 11.1 (Baccam et al. (2006) *Journal of Virology*)

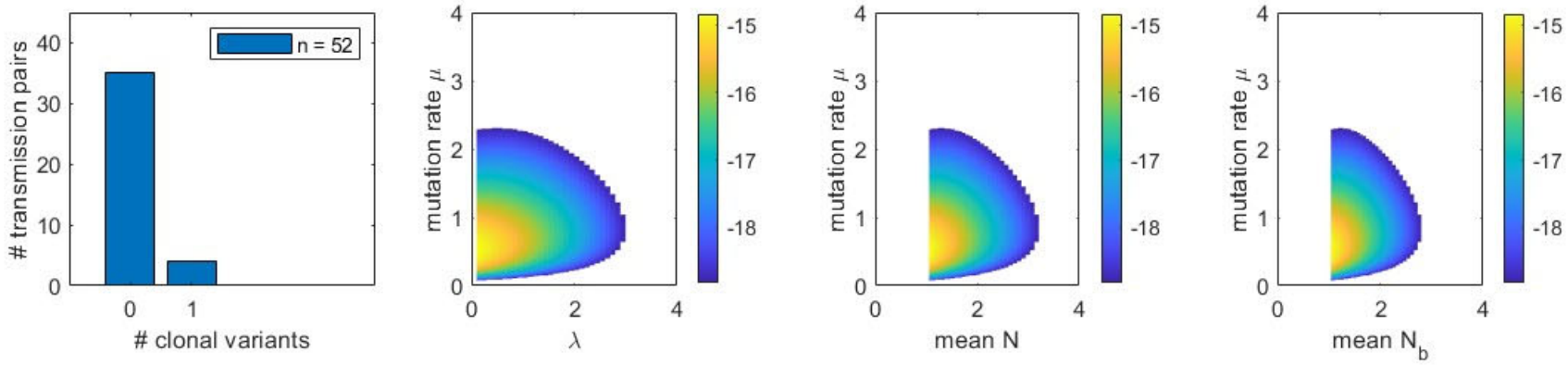
Mean N_b is very small (~1-2 viral particles)

Mutation rate μ is ~0.68-3.05 per genome per replication cycle.

Consistent with Pauly et al. (2017) *eLife*

2-3 mutations per genome per replication cycle with ~40% being lethal deleterious

Estimates for SARS-CoV-2



Assumed within-host R_0 = 7.4 (Ke et al. (2021) *PNAS*)

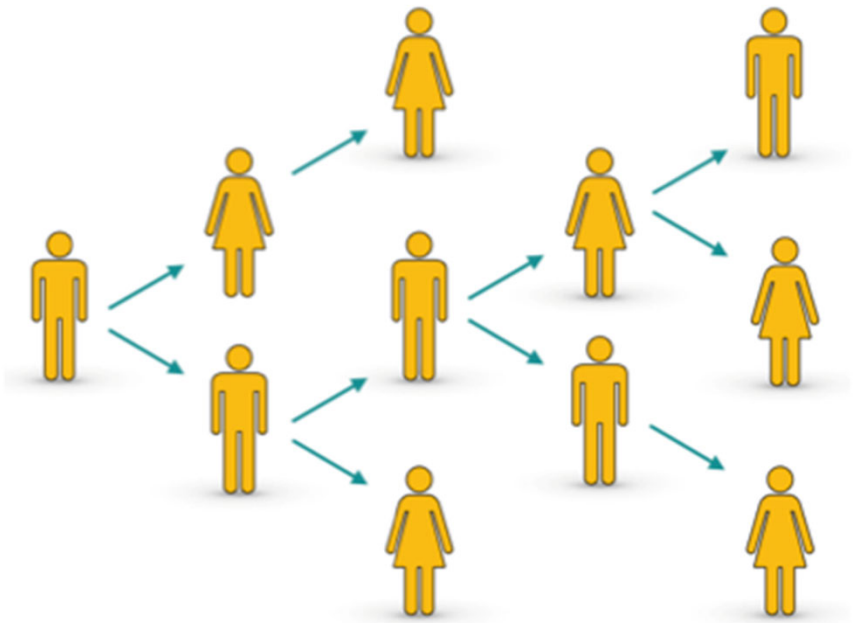
Mean N_b is very small (~1-4 viral particles)

Mutation rate μ is ~0.2-2.2 per genome per replication cycle.

Consistent with literature estimate of 0.9 per genome per replication cycle

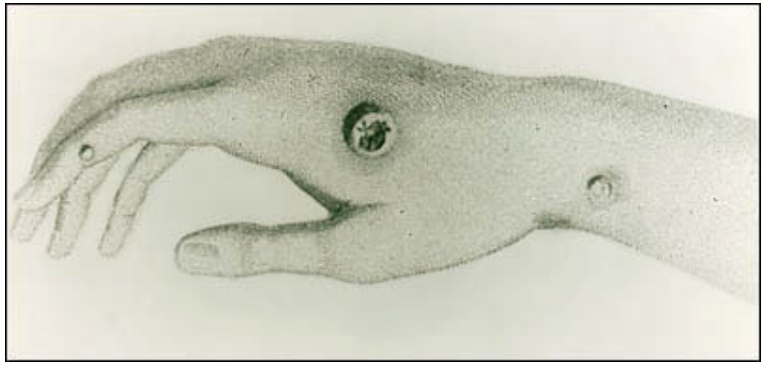
SO WHAT?

Feasibility of reconstructing who-infected-whom based on minor genetic variation



Gauge potential impact of control measures

Variolation



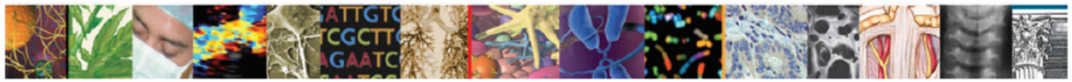
https://www.nlm.nih.gov/exhibition/smallpox/sp_variolation.html

SARS-CoV-2 “variolation” hypothesis

Masks Do More Than Protect Others During COVID-19: Reducing the Inoculum of SARS-CoV-2 to Protect the Wearer

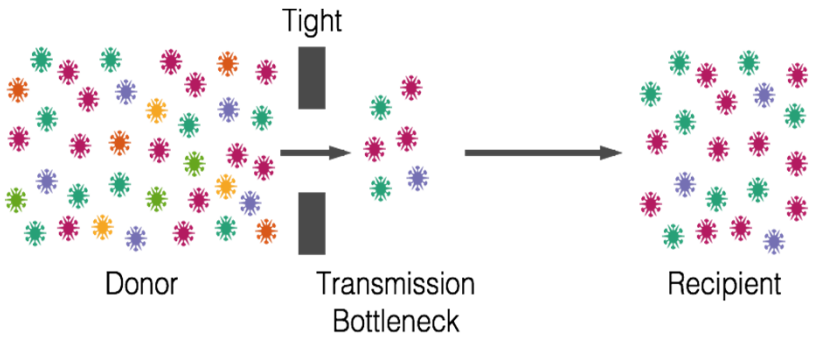
Monica Gandhi, MD, MPH[1], Chris Beyrer, MD, MPH[2], and Eric Goosby, MD[1]

Facial Masking for Covid-19 — Potential for “Variolation” as We Await a Vaccine



The NEW ENGLAND JOURNAL of MEDICINE

Implications for viral adaptation at the population level



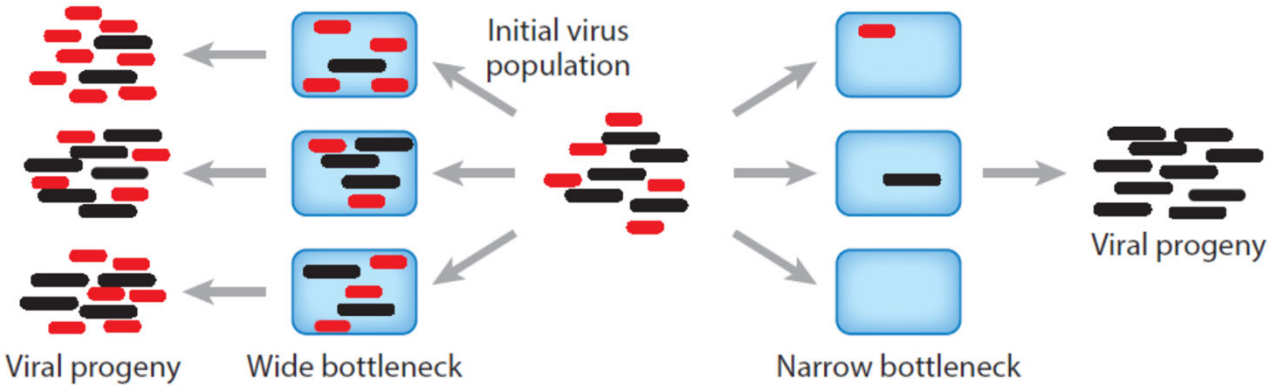
Genetic drift is important process during the transmission process

(explains why viral populations generally carry a deleterious mutation load/why purifying selection is incomplete)

Will tight bottlenecks always act to slow down viral adaptation at the population-level?

Implications for viral adaptation at the population level

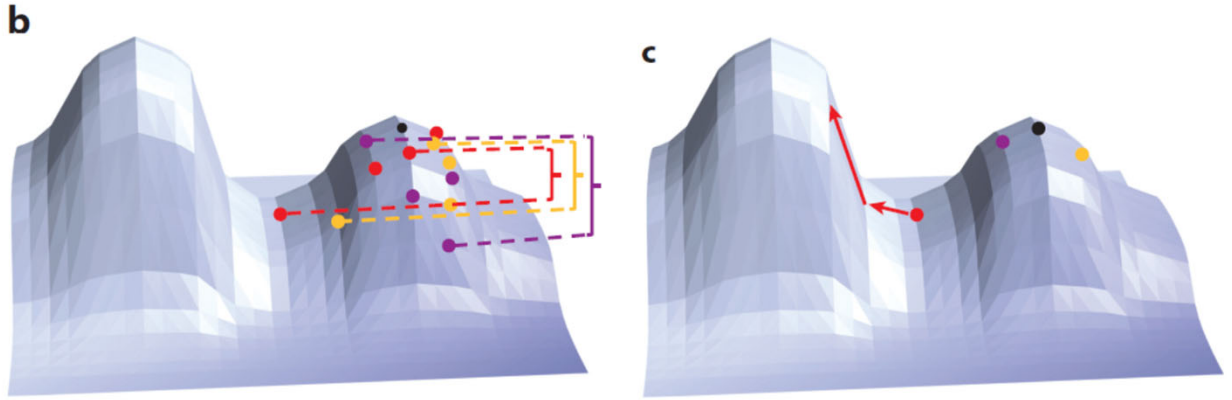
Narrow bottlenecks can purge viral populations of cheaters (e.g., defective interfering particles)



Zwart & Elena (2015) *Annual Review of Virology*

Implications for viral adaptation at the population level

Narrow bottlenecks can aid in crossing fitness valleys



Zwart & Elena (2015) *Annual Review of Virology*

SCIENCE ADVANCES | RESEARCH ARTICLE

EPIDEMIOLOGY

Genomic epidemiological models describe pathogen evolution across fitness valleys

Pablo Cárdenas[1]*, Vladimir Corredor[2], Mauricio Santos-Vega[3]

(2022)

Conclusions

Various approaches exist to estimate the size of the transmission bottleneck (that is, the population bottleneck) using viral deep-sequencing data

Existing approaches rely on genetic variation present in the donor and how it is shared with the recipient. These have the potential to bias Nb estimates low.

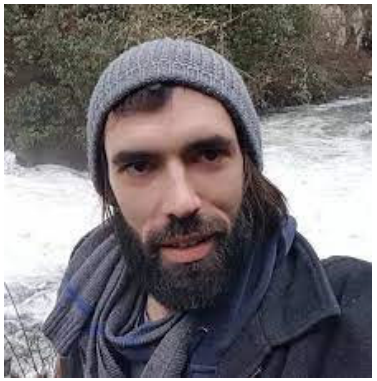
An alternative approach is to use the distribution of clonal variants observed between transmission pairs

Transmission bottlenecks appear to be very tight for respiratory pathogens of humans

Whether these bottlenecks impede or facilitate adaptive evolution at the population level depends on the fitness landscape the viral population is faced with and the impact of viral cheaters

Acknowledgments (for *de novo* clonal variant work)

Teresa Shi



Jeremy Harris

Mike Martin

National Institutes of Health

Emory CEIRR