Modelling and inference of heterogeneity in bacterial growth

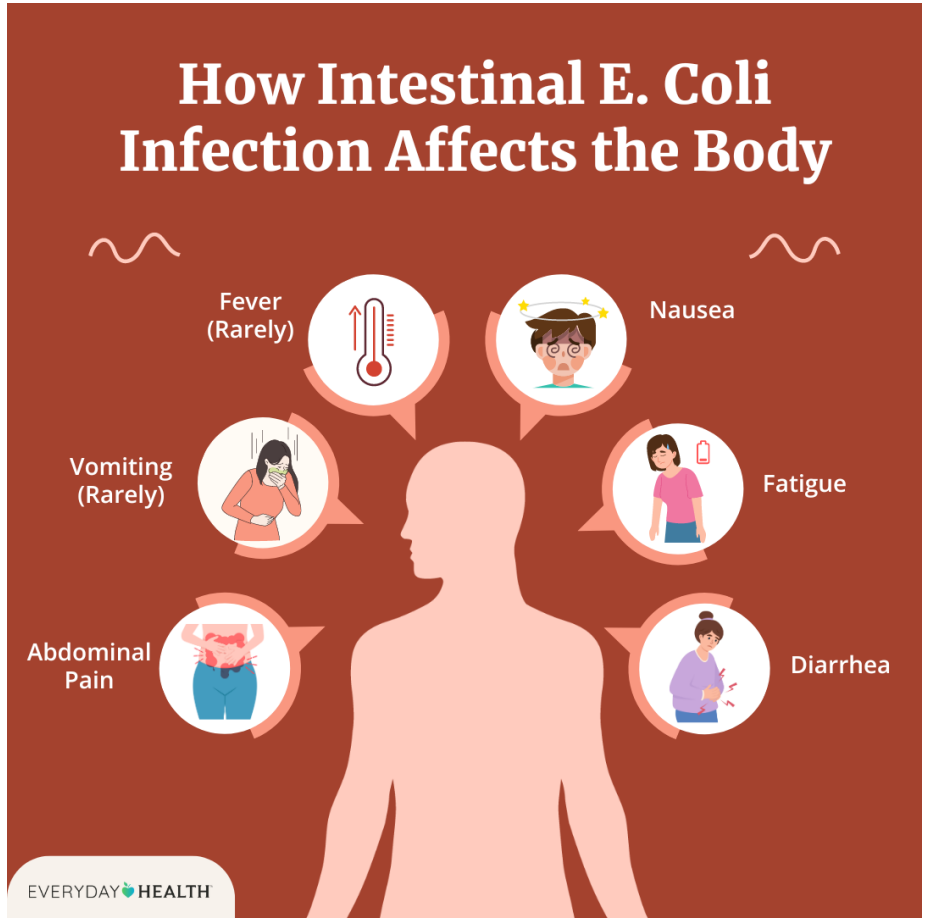
John Mackenzie, Gabriela Gomes, Adam Hillman
Department of Mathematics and Statistics

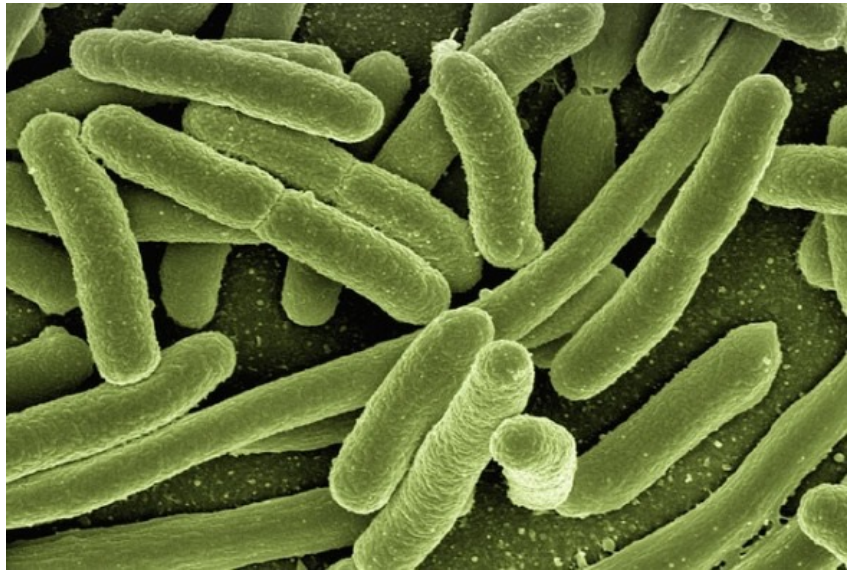
Michele Zagnoni and Rozan Vroman
EEE

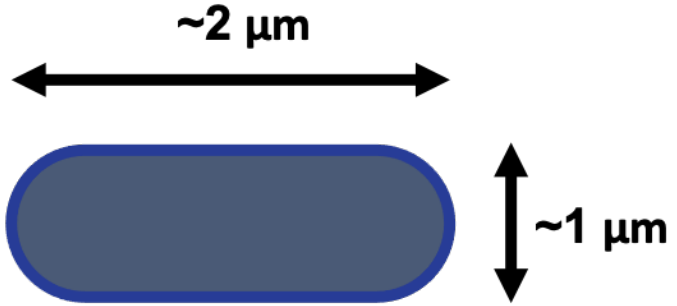
Paul Herron, Suzy Humphry, Liam Rooney
SIPBS

University of Strathclyde Glasgow

E. Coli


How Intestinal E. Coli Infection Affects the Body
Fever (Rarely)
Nausea
Vomiting (Rarely)
Fatigue
Abdominal Pain
Diarrhea
EVERYDAY HEALTH



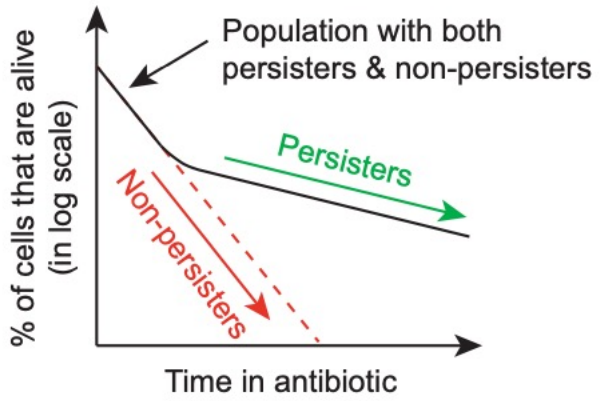

~2 µm
~1 µm

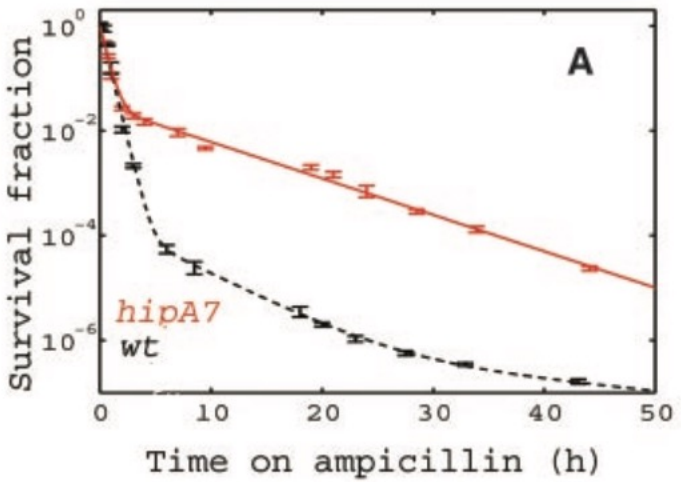
Cell persistence

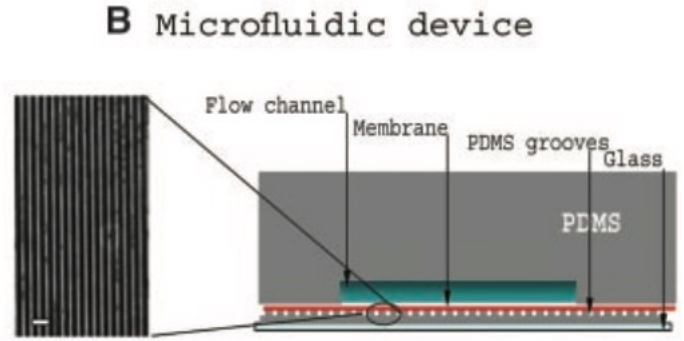
Bacterial Persistence as a Phenotypic Switch

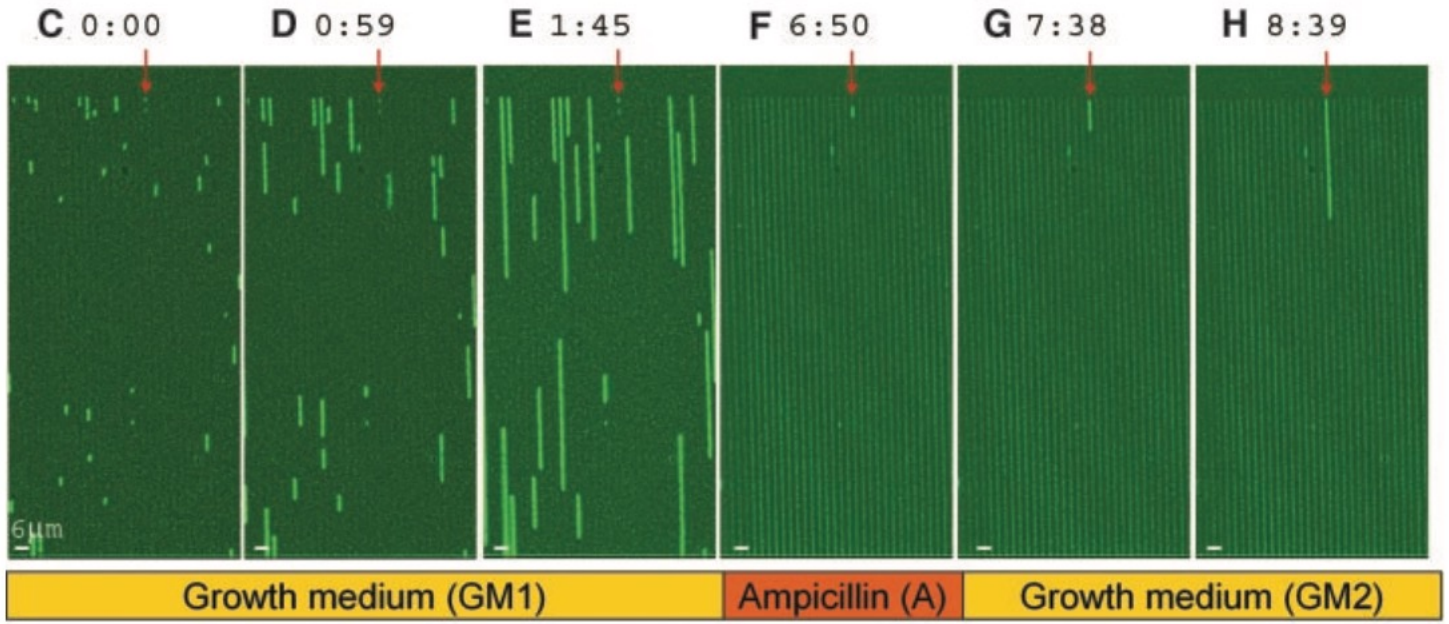
Nathalie Q. Balaban,[1,2*] Jack Merrin,[1] Remy Chait,[1] Lukasz Kowalik,[1] Stanislas Leibler[1]

A fraction of a genetically homogeneous microbial population may survive exposure to stress such as antibiotic treatment. Unlike resistant mutants, cells regrown from such persistent bacteria remain sensitive to the antibiotic. We investigated the persistence of single cells of *Escherichia coli* with the use of microfluidic devices. Persistence was linked to preexisting heterogeneity in bacterial populations because phenotypic switching occurred between normally growing cells and persister cells having reduced growth rates. Quantitative measurements led to a simple mathematical description of the persistence switch. Inherent heterogeneity of bacterial populations may be important in adaptation to fluctuating environments and in the persistence of bacterial infections.









Mathematical and statistical modelling

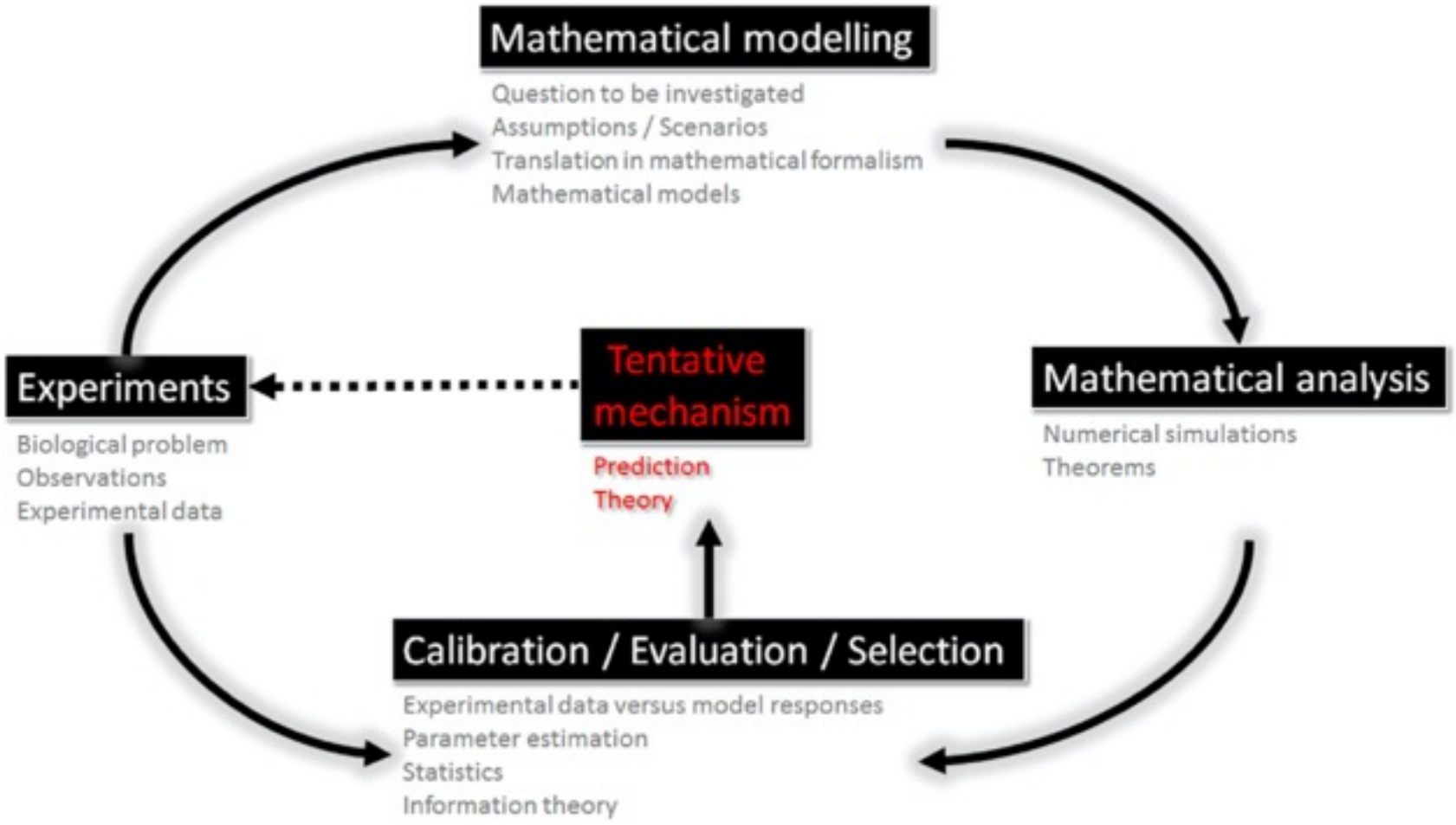


Fig. 1. Mathematical biology workflow; adapted from (Portet, 2015).

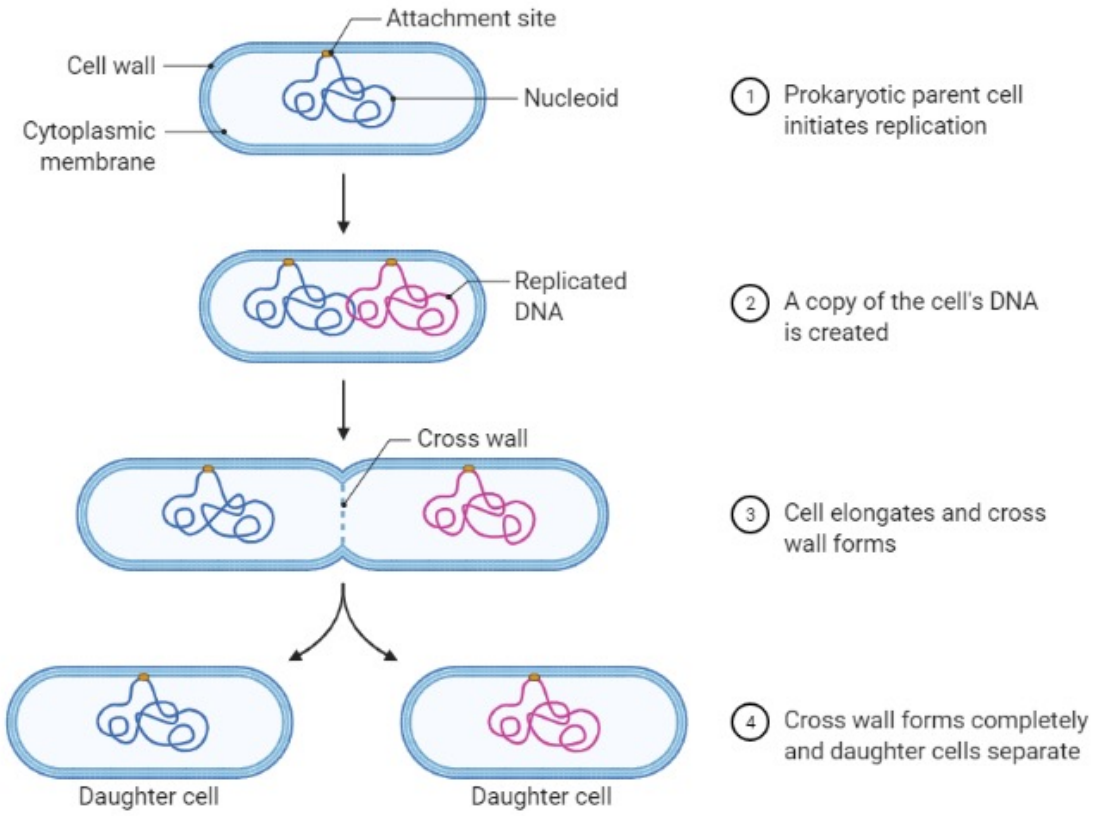
Simple ODE model: homogeneous population

$$\frac{dN}{dt} = 2\mu N - \mu N$$

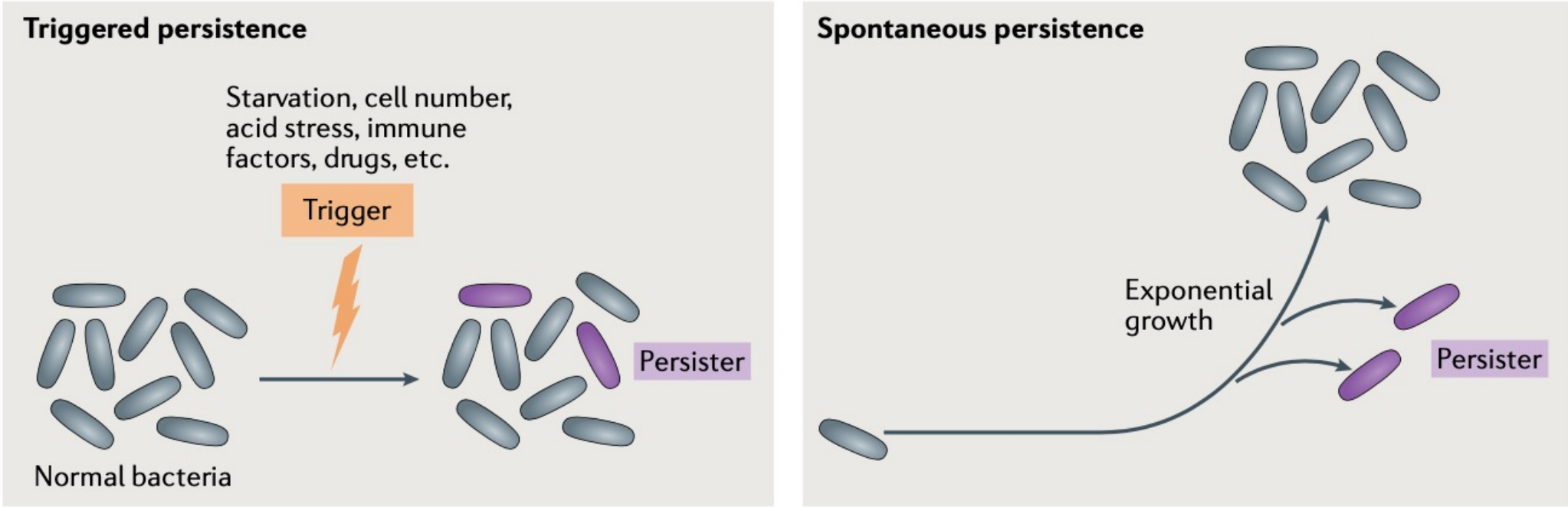
$$N(0) = n_0$$

$$N(t) = n_0 e^{\mu t}$$

$$\ln(N(t)) = \ln(n_0) + \mu t$$



Models of persistence



2-phenotype ODE system

- Assume we have two phenotypes N_1 and N_2 with division rates μ_1 and μ_2.
- Probability p that any newly born cell belongs to phenotype 1 and (1-p) to phenotype 2.
- Antibiotic stress reduces cell viability from new births.

$$\frac{dN_1}{dt} = \beta p(\mu_1 N_1 + \mu_2 N_2) - \mu_1 N_1,$$

$$\frac{dN_2}{dt} = \beta(1-p)(\mu_1 N_1 + \mu_2 N_2) - \mu_2 N_2,$$

$$\beta = 2(1-\sigma) \text{ and } 0 \leq \sigma \leq 1$$

Gomes MGM, King JG, Nunes A, Colegrave N, Hoffmann AA (2019) The effects of individual nonheritable variation on fitness estimation and coexistence. *Ecol Evol* **9**, 8995-9004.

2-phenotype model

$$\frac{d}{dt}\begin{bmatrix} N_1 \\ N_2 \end{bmatrix} = \begin{bmatrix} (\beta p - 1)\mu_1 & \beta\, p\, \mu_2 \\ \beta(1-p)\mu_1 & (\beta(1-p)-1)\mu_2 \end{bmatrix}\begin{bmatrix} N_1 \\ N_2 \end{bmatrix}.$$

$$\boldsymbol{N}(t) = C_1 e^{\lambda^+ t}\boldsymbol{v}^+ + C_2 e^{\lambda^- t}\boldsymbol{v}^-,$$

Growth rate determined by λ^+

Equivalent homogeneous model

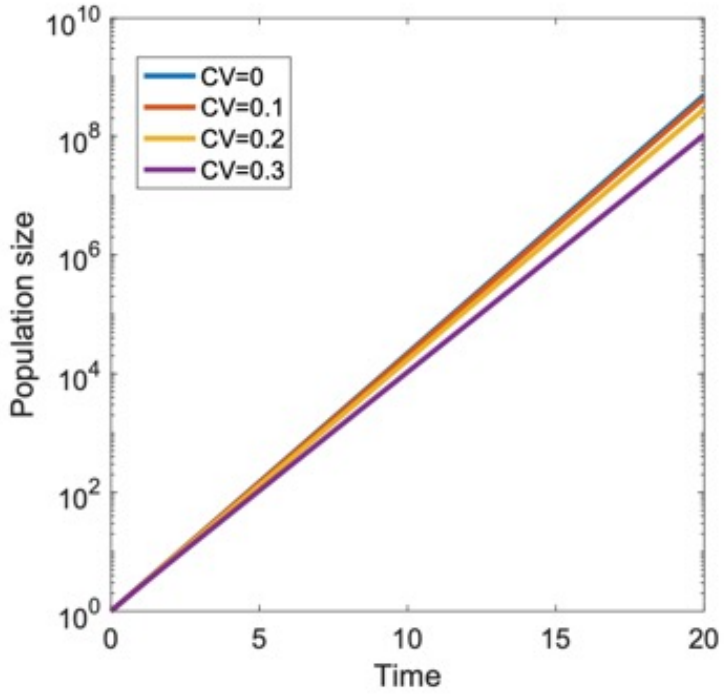
Equivalent one phenotype (homogeneous) system

$$\bar{\mu} = p\mu_1 + (1-p)\mu_2.$$

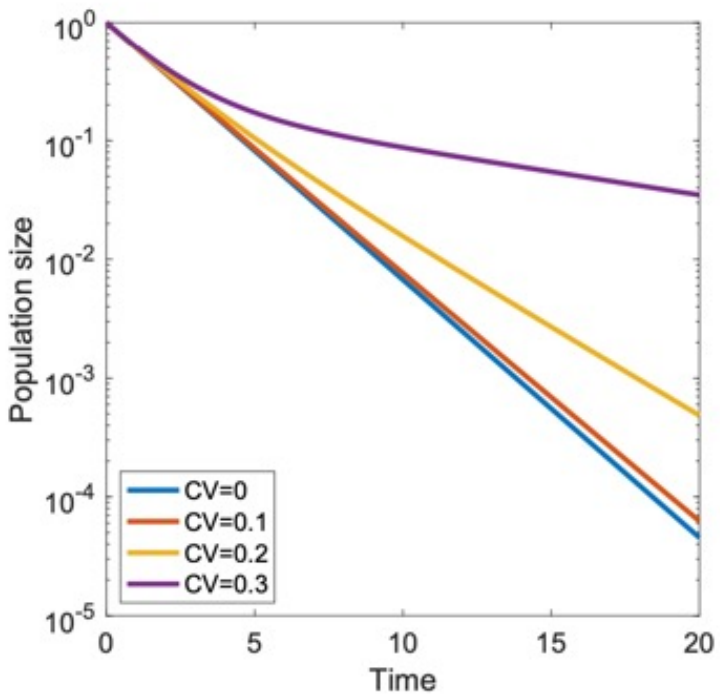
$$\frac{dN}{dt} = \beta\bar{\mu}N - \bar{\mu}N, \quad N(0) = 1,$$

$$N(t) = \exp((\beta - 1)\bar{\mu}t)$$

Solution behaviour in zero and high stress



(a) Zero stress ($\beta = 2$).



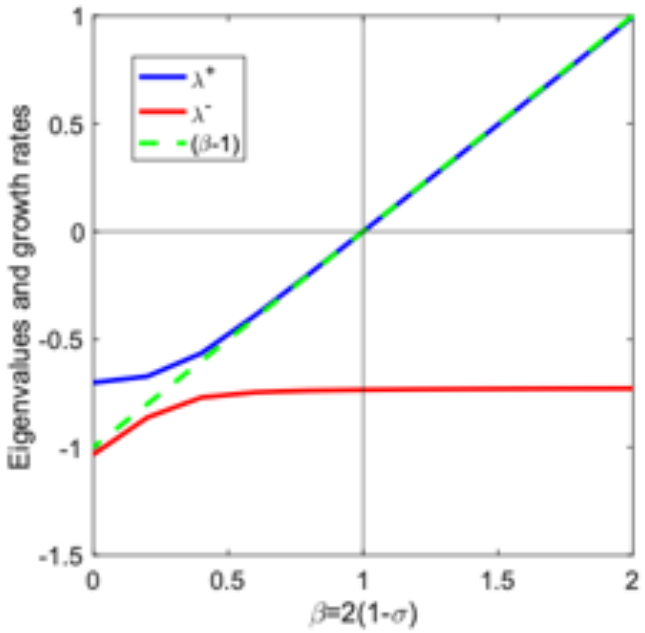
(b) High stress ($\beta = 0.25$).

$$\mathrm{CV} = \frac{\sqrt{p(\mu_1 - \bar{\mu})^2 + (1-p)(\mu_2 - \bar{\mu})^2}}{\bar{\mu}}.$$

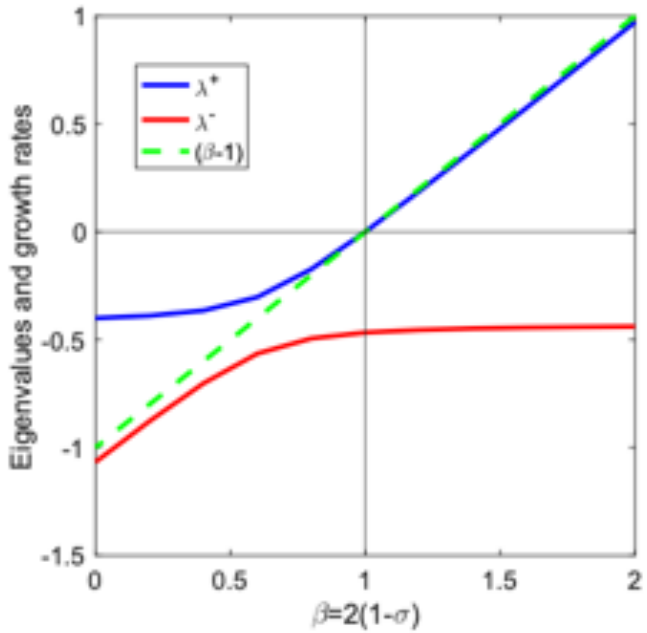
Theoretical growth rate bounds

$$
\begin{aligned}
(\beta-1)\mu_1 < \lambda^+ < (\beta-1)\bar{\mu}, \quad & 1 < \beta \leq 2 \\
\lambda^+ = 0, \quad & \beta = 1 \quad \text{and} \\
(\beta-1)\mu_1 > \lambda^+ > (\beta-1)\bar{\mu}, \quad & 0 \leq \beta < 1.
\end{aligned}
$$

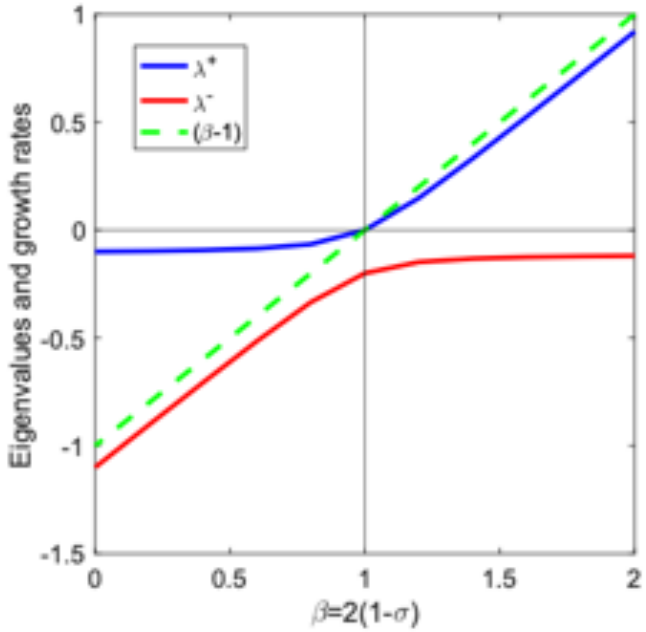
Growth rates over stress gradient



CV=0.1



CV=0.2



CV=0.3

n-phenotype ODE model

$$\{N_i(t)\}_{i=1}^{n} \qquad \mu_{i+1} > \mu_i, \text{ for } i = 1, \ldots, n-1$$

$$\frac{d}{dt}\begin{bmatrix} N_1 \\ N_2 \\ \vdots \\ N_n \end{bmatrix} = \begin{bmatrix} (2p_1 - 1)\mu_1 & 2p_1\mu_2 & \cdots & 2p_1\mu_n \\ 2p_2\mu_1 & (2p_2 - 1)\mu_2 & \cdots & 2p_2\mu_n \\ \vdots & \vdots & \ddots & \vdots \\ 2p_n\mu_1 & 2p_n\mu_2 & \cdots & (2p_n - 1)\mu_n \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \\ \vdots \\ N_n \end{bmatrix}.$$

If the eigenvalues of A are ordered such that $\lambda_1 > \lambda_2 > \ldots > \lambda_n$, then

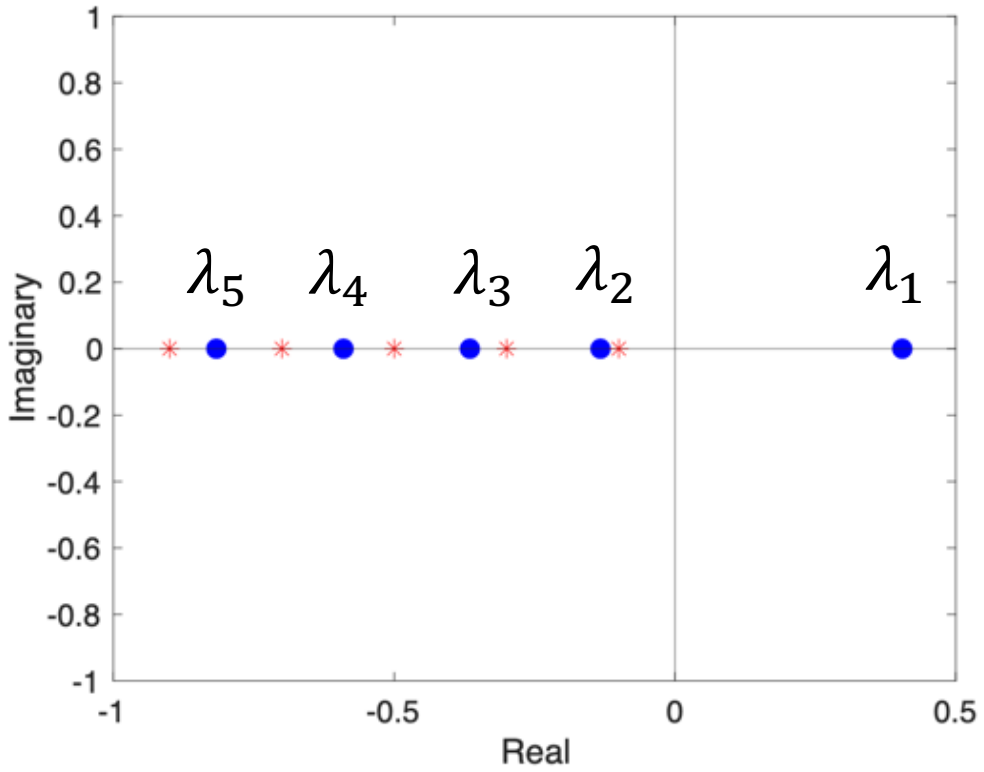
$$-\mu_n < \lambda_n < -\mu_{n-1} < \lambda_{n-1} < \cdots < -\mu_2 < \lambda_2 < -\mu_1 < 0$$

$$0 < \max(\mu_1, 2\bar{\mu} - \mu_n) < \lambda_1 < \bar{\mu}.$$

Example of n-phenotype system

$$\mu_1 = \frac{1}{2n}, \quad \text{and} \quad \mu_i = \mu_{i-1} + \frac{1}{n}, \quad i = 2, \ldots, n.$$

$$p_i = 1/n,\ i = 1, \ldots, n \qquad \bar{\mu} = \sum_{i=1}^{n} p_i \mu_i = 1/2.$$



$n = 5$

Continuous phenotype model

$$\frac{\partial N}{\partial t} = 2p(\mu) \int_{\mathcal{V}} \mu' N(\mu', t)\, \mathrm{d}\mu' - \mu N(\mu, t),$$

$$N(\mu, t) = T(t) S(\mu)$$

$$T(t) = Ae^{ct} \qquad S(\mu) = \int_{\mathcal{V}} \frac{2p(\mu)}{c + \mu} \mu' S(\mu')\, \mathrm{d}\mu'.$$

$$S(\mu) = 2c_1 b(\mu), \quad \text{where} \quad b(\mu) = \frac{p(\mu)}{c + \mu}.$$

Growth rate of continuous model

Growth rate equation

$$\int_{\nu} \mu' \frac{p(\mu')}{c+\mu'} \, \mathrm{d}\mu' = \frac{1}{2}.$$

There exists a unique growth rate $0 < c < \bar{\mu}$.

Example continuous phenotype

$\mathcal{V} = [0,1]$ and a uniform distribution $\mu \sim \mathcal{U}_{[0,1]}$ and hence $p(\mu) = 1$

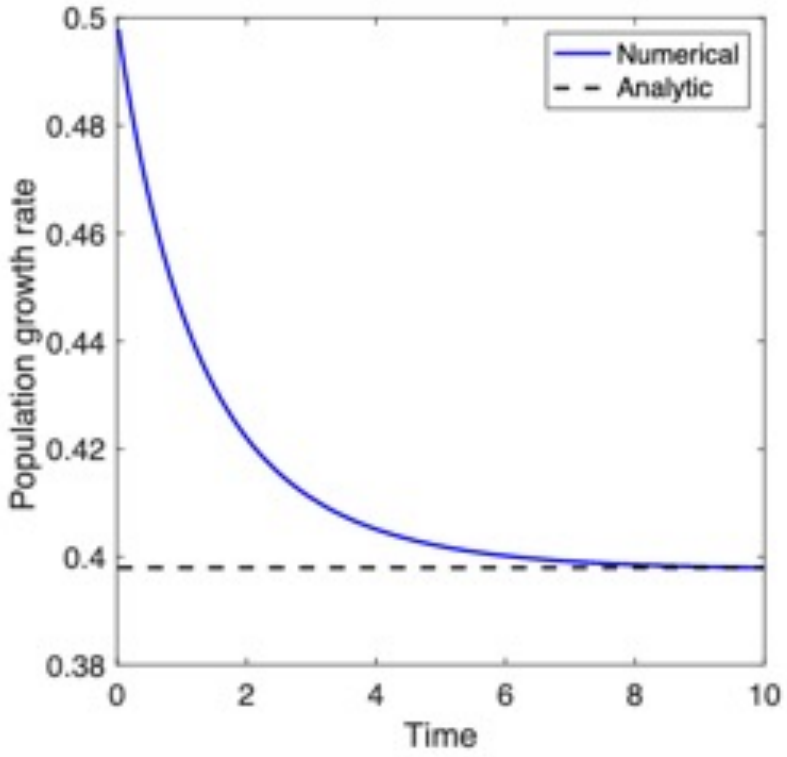
$$c \ln\left(1 + \frac{1}{c}\right) = \frac{1}{2}$$

$$c = 0.3980 \; < \; \bar{\mu} = 0.5$$

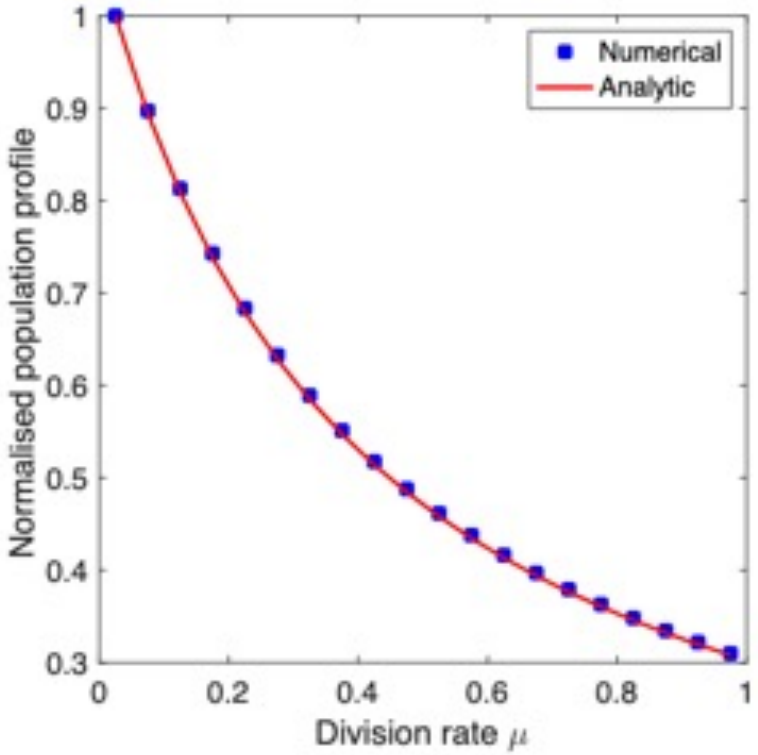
$$N(\mu, 0) = p(\mu), \quad \mu \in \mathcal{V},$$

$$\left.\frac{dN_p}{dt}\right|_{t=0} = \frac{d}{dt}\int_{\mathcal{V}} N(\mu,0)\,\mathrm{d}\mu = \int_{\mathcal{V}} \mu N(\mu,0)\,\mathrm{d}\mu = \int_{\mathcal{V}} \mu\, p(\mu)\,\mathrm{d}\mu = \bar{\mu},$$

Numerical simulations



(a) Population growth rates.



(b) Normalised population profiles.

Normalised profile

$$\frac{p(\mu)}{(c+\mu)} \Big/ \left(\int_{\mathcal{V}} \frac{p(\mu)}{c+\mu} \, \mathrm{d}\mu \right)$$

Time to subdivision data

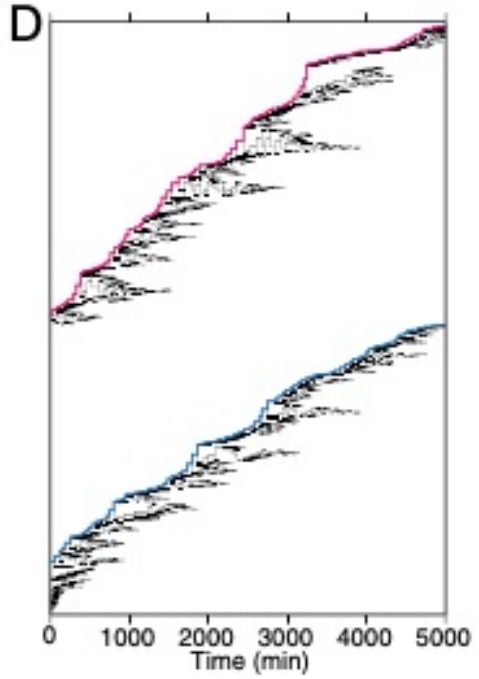
PNAS

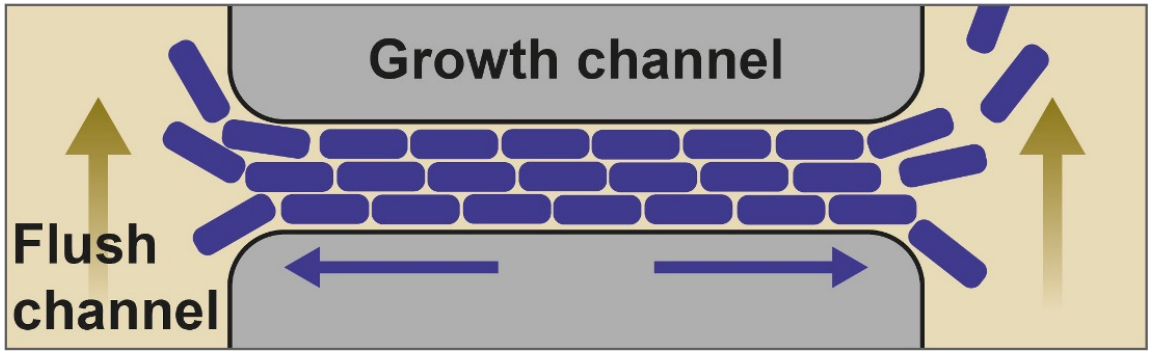
Noise-driven growth rate gain in clonal cellular populations

Mikihiro Hashimoto[a], Takashi Nozoe[a], Hidenori Nakaoka[a], Reiko Okura[a], Sayo Akiyoshi[a], Kunihiko Kaneko[a,b], Edo Kussell[c,d], and Yuichi Wakamoto[a,b,1]

[a]Department of Basic Science, Graduate School of Arts and Sciences, University of Tokyo, Tokyo 153-8902, Japan; [b]Research Center for Complex Systems Biology, University of Tokyo, Tokyo 153-8902, Japan; [c]Department of Biology, Center for Genomics and Systems Biology, New York University, New York, NY 10003; and [d]Department of Physics, New York University, New York, NY 10003

Edited by Daniel L. Hartl, Harvard University, Cambridge, MA, and approved February 5, 2016 (received for review October 19, 2015)





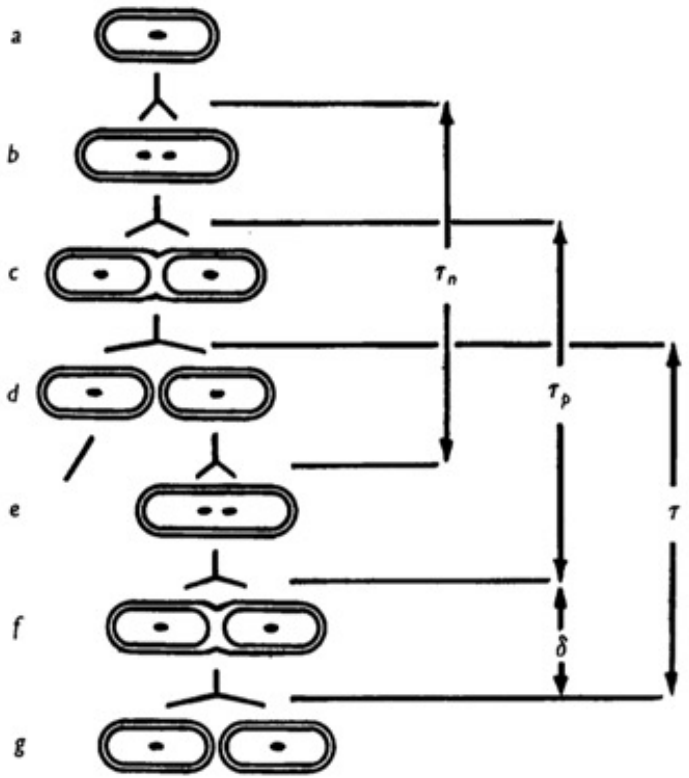
E.O Powell 1958

POWELL, E. O. (1958). *J. gen. Microbiol.* **18**, 382–417

An Outline of the Pattern of Bacterial Generation Times

BY E. O. POWELL

Microbiological Research Establishment (Ministry of Supply) Porton, Wiltshire



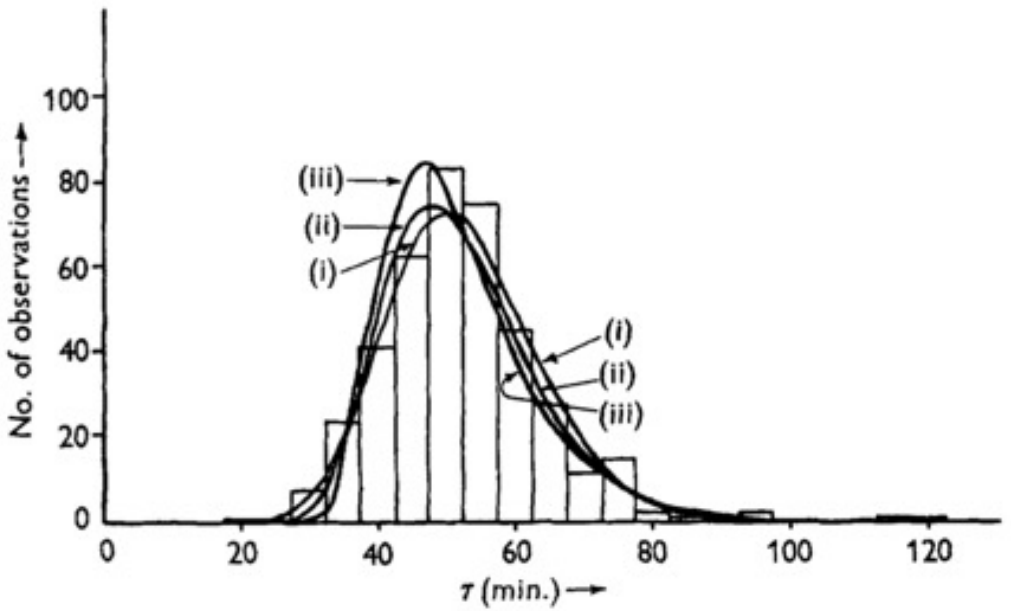
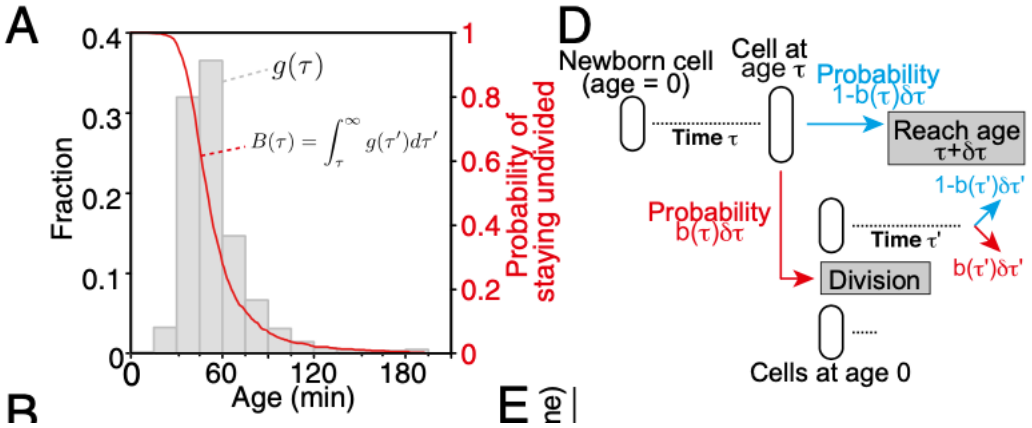
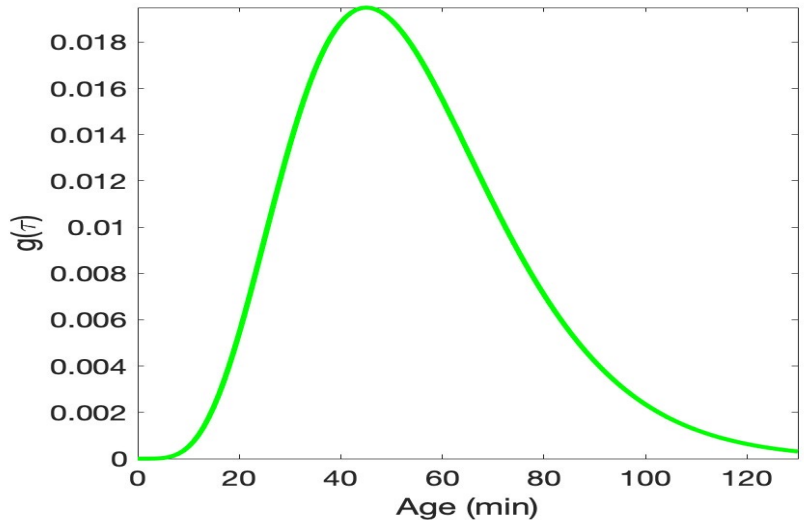


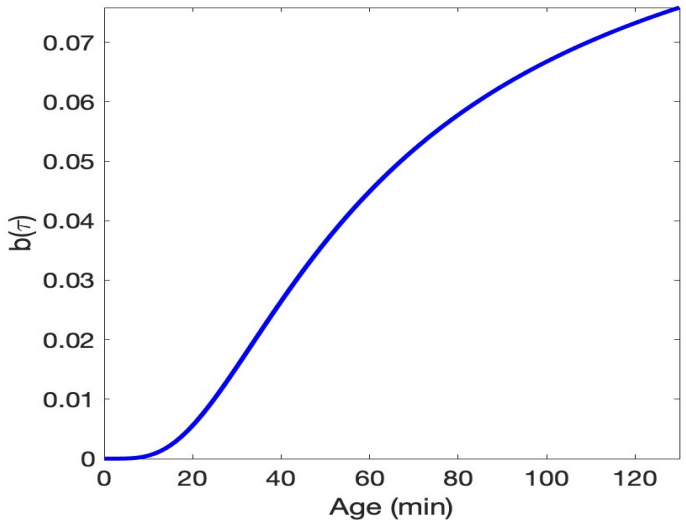
Fig. 6. Distribution of τ in Series Ba 4. Curve (i) Pearson Type III distribution fitted by maximum likelihood; (ii) Pearson Type V distribution fitted by maximum likelihood; (iii) Generalized Yule distribution (Equation 20) fitted by moments.

Hashimoto et al: Probabilistic model



$$b(\tau) = \frac{g(\tau)}{\int_{\tau}^{\infty} g(\tau')\, d\tau'}$$





Age-structured PDE model (homogeneous)

Age-structured population $N(t, \tau)$. Assume $d\tau/dt = 1$.

$$\frac{\partial N}{\partial t} + \frac{\partial N}{\partial \tau} = -b(\tau)N$$

Birth process

$$N(t, 0) = 2 \int_0^\infty b(\tau)N \, \mathrm{d}\tau$$

Initial condition

$$N(0, \tau) = N_0(\tau)$$

Connection to ODE model

If $g(\tau) = \mu e^{-\mu t}$ then $b(\tau) = \mu$.
Age-structured PDE:

$$\frac{\partial N}{\partial t} + \frac{\partial N}{\partial \tau} = -\mu N$$

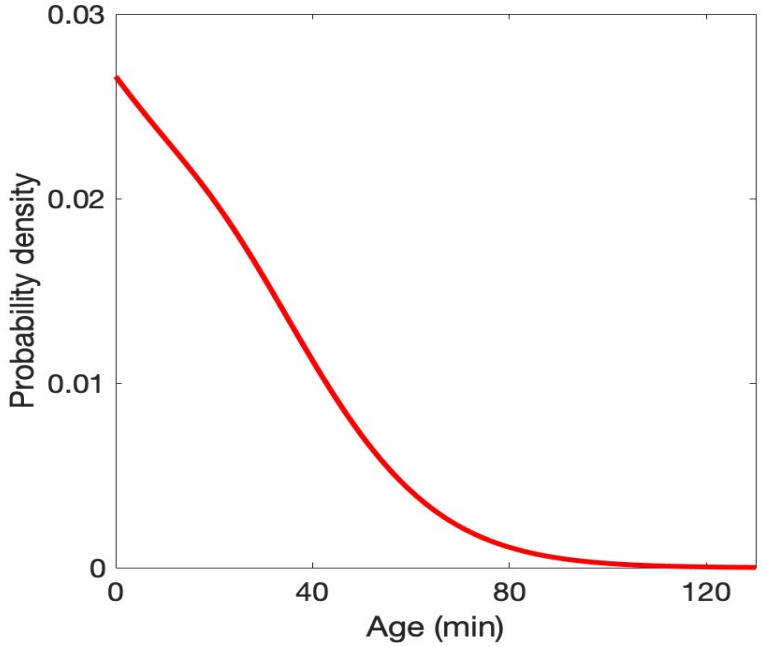
Integrate wrt τ and use boundary condition for birth. Denote population

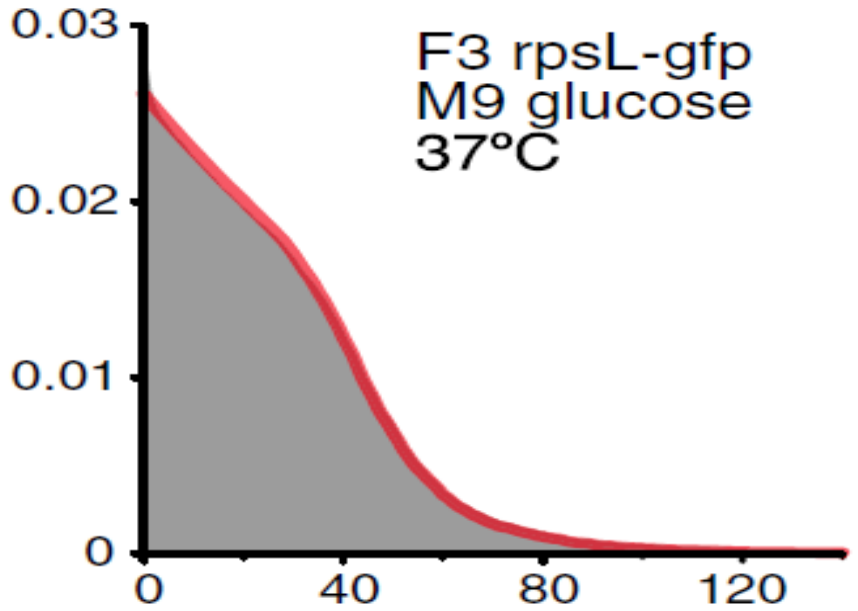
$$N^{tot}(t) = \int_0^\infty N(t, \tau)\, \mathrm{d}\tau$$

$$\frac{d}{dt} N^{tot}(t) = 2\mu N^{tot}(t) - \mu N^{tot}(t)$$

$$N^{tot}(t) = N^{tot}(0) e^{\mu t}$$

Age-structured model results





$$g(\tau) = \mathrm{Gamma}(6, 9)$$

Data from Hashimoto et al

Age-structured PDE model (homogeneous with stress)

Antibiotic (stress) negatively impacts cells at birth:

$$\frac{\partial N}{\partial t} + \frac{\partial N}{\partial \tau} = -b(\tau)N$$

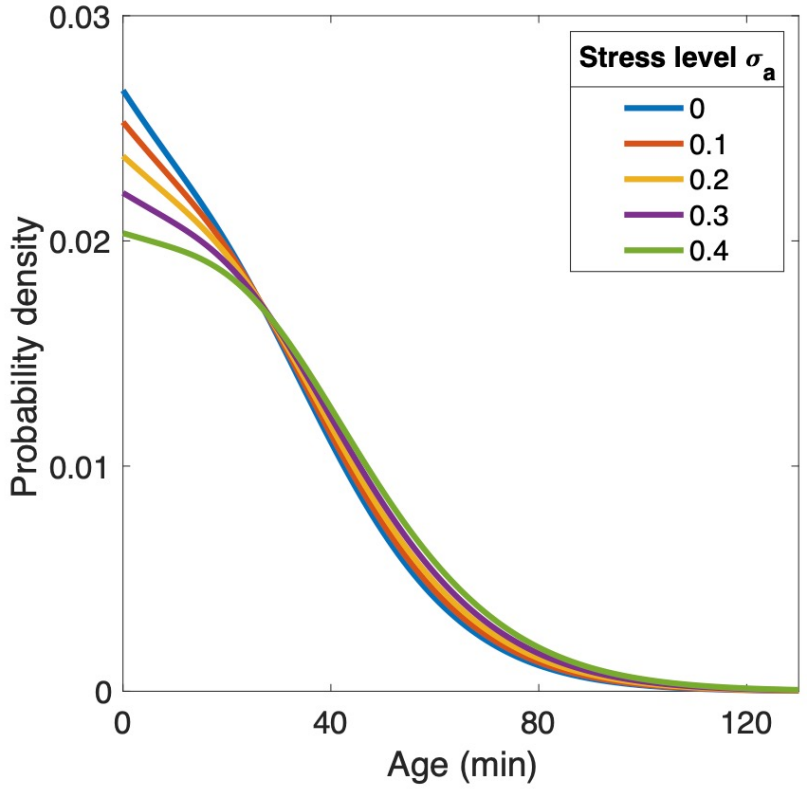
Birth process

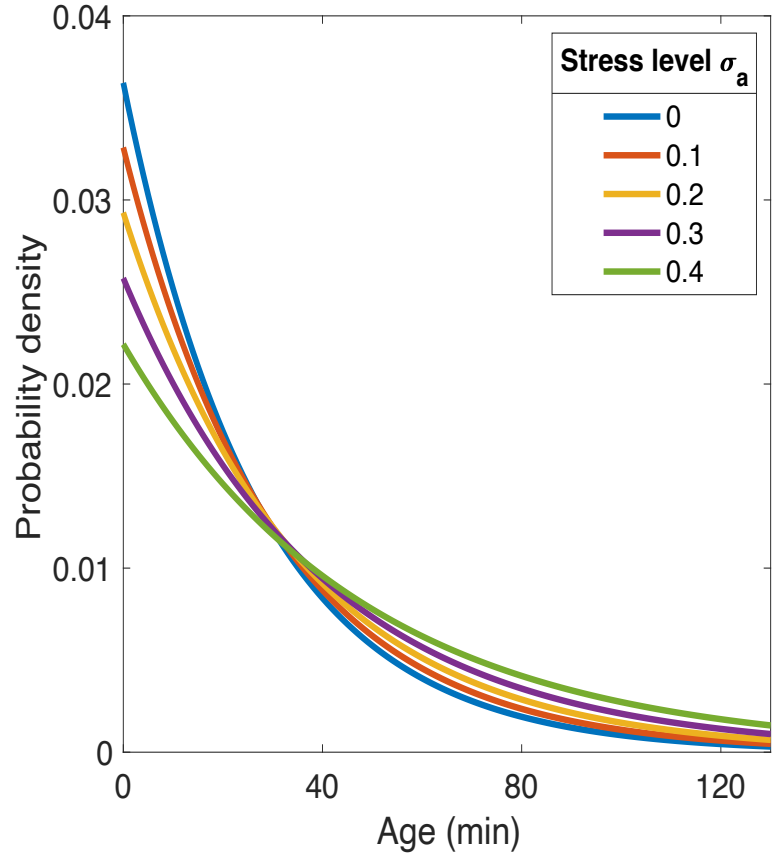
$$N(t,0) = 2(1-\sigma_a)\int_0^\infty b(\tau)N \, \mathrm{d}\tau$$

Initial condition

$$N(0,\tau) = N_0(\tau)$$

Selection in age as a function of stress





$$g(\tau) = \text{Gamma}(6, 9)$$

$$g(\tau) = \text{Exp}(1/9)$$

Age-structured (heterogeneous)

Two sub-populations $N_1(t,\tau)$ and $N_2(t,\tau)$

$$\frac{\partial N_1}{\partial t} + \frac{\partial N_1}{\partial \tau} = -b_1(\tau)N_1$$

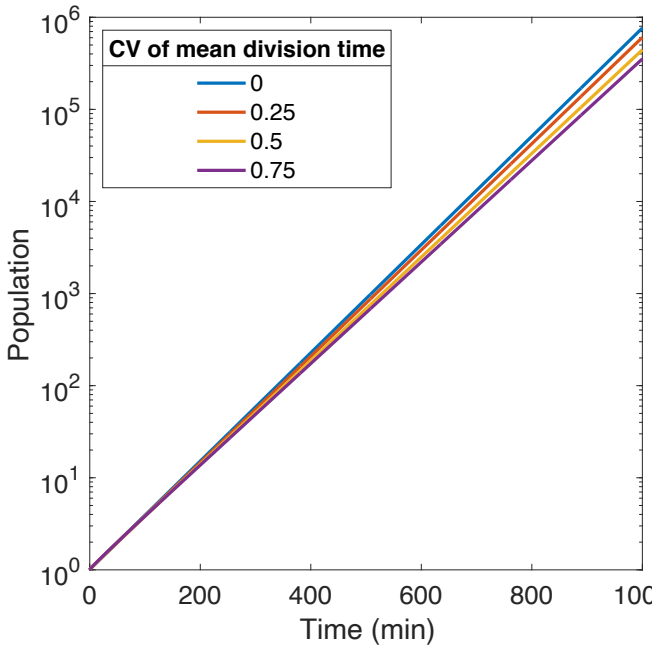
$$\frac{\partial N_2}{\partial t} + \frac{\partial N_2}{\partial \tau} = -b_2(\tau)N_2$$

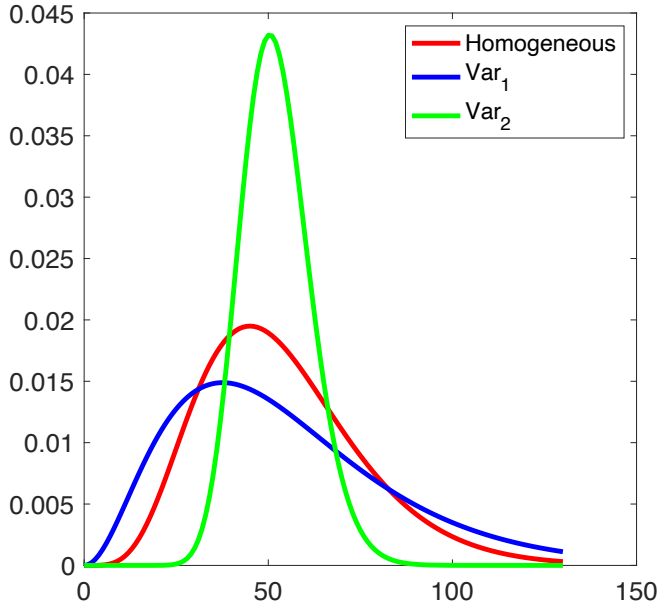
Birth process

$$N_1(t,0) = 2p_1 \int_0^\infty (b_1(\tau)N_1 + b_2(\tau)N_2)d\tau$$

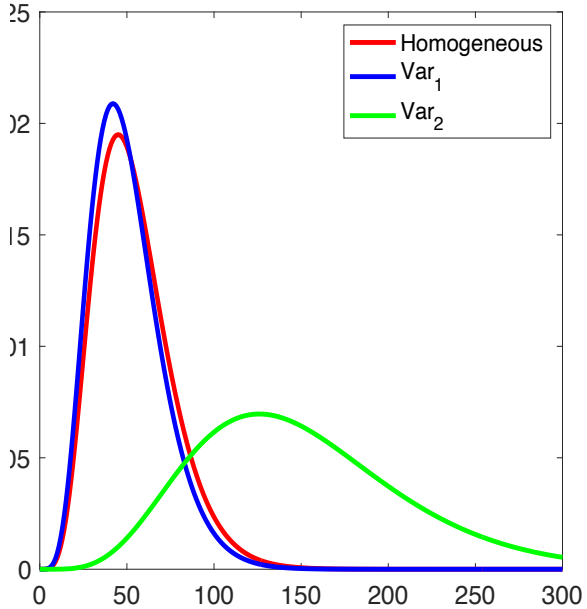
$$N_2(t,0) = 2p_2 \int_0^\infty (b_1(\tau)N_1 + b_2(\tau)N_2)d\tau$$

2 phenotype age-structured results



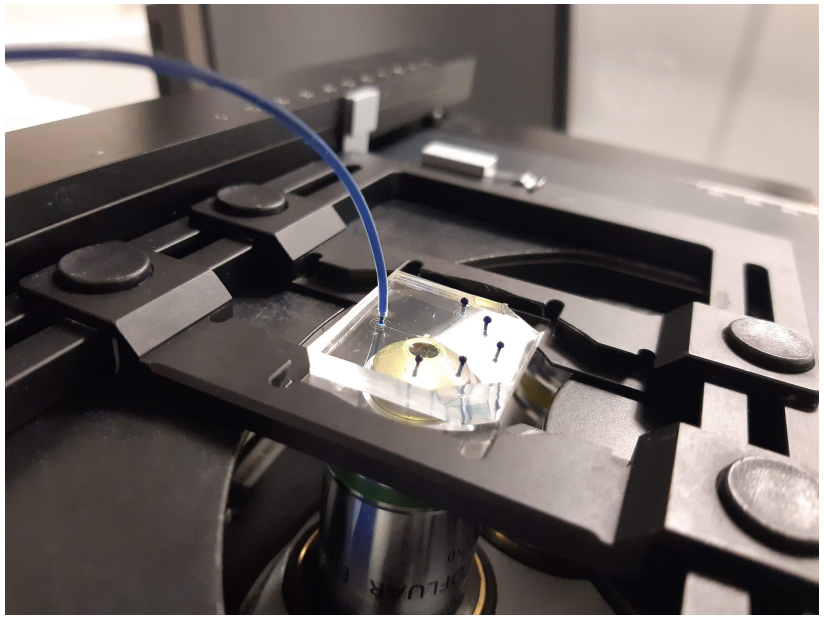


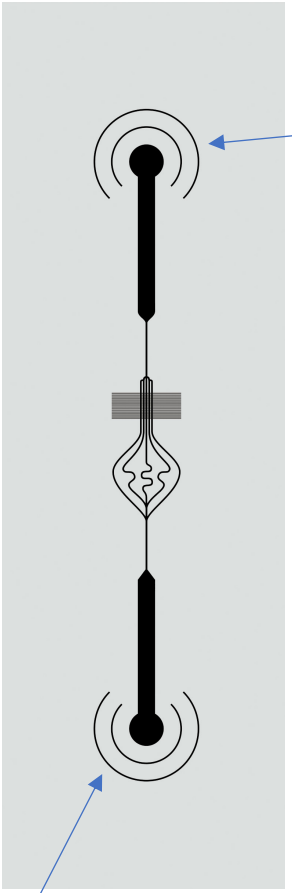
$CV = 0.25$



$CV = 0.75$

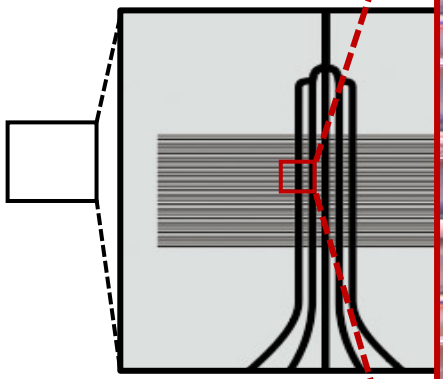
Microfluidics (Rozan, Michele, Paul)

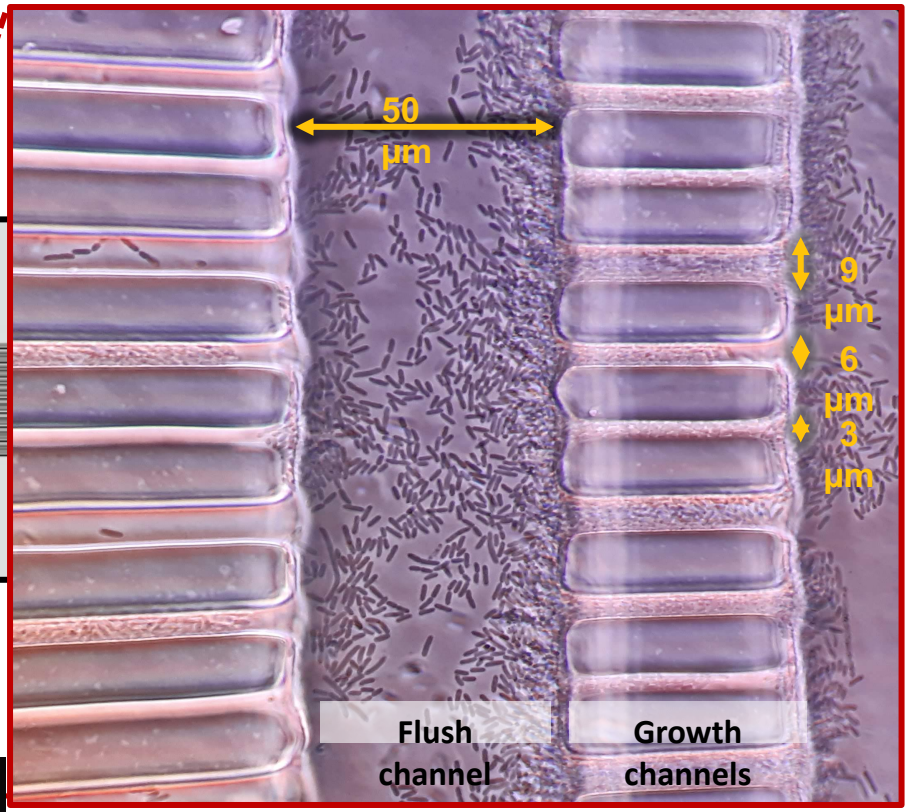


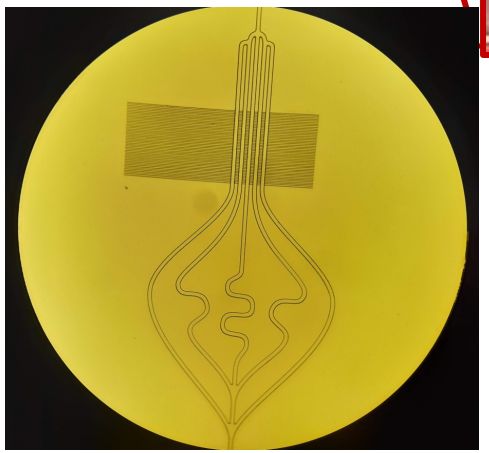


Tube input

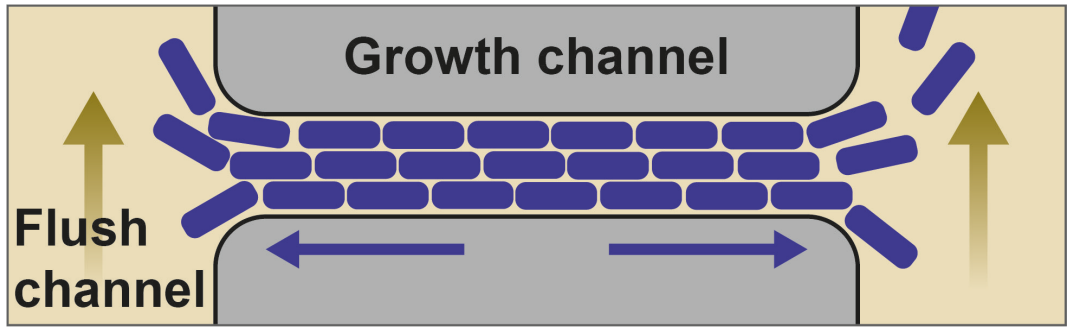
Tube output

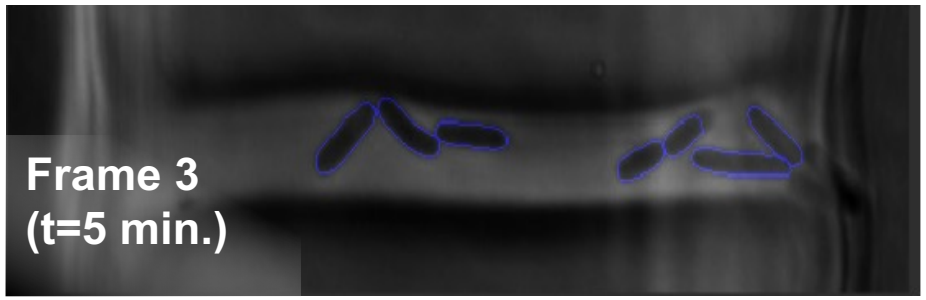


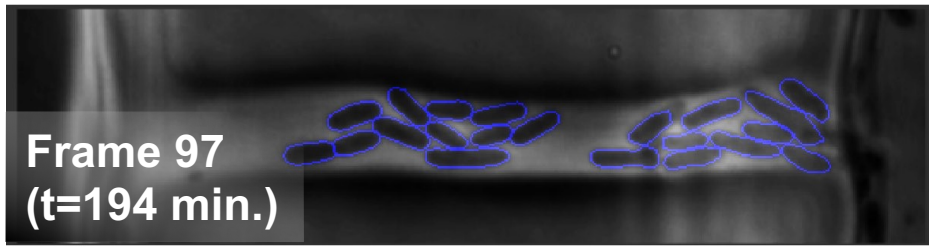


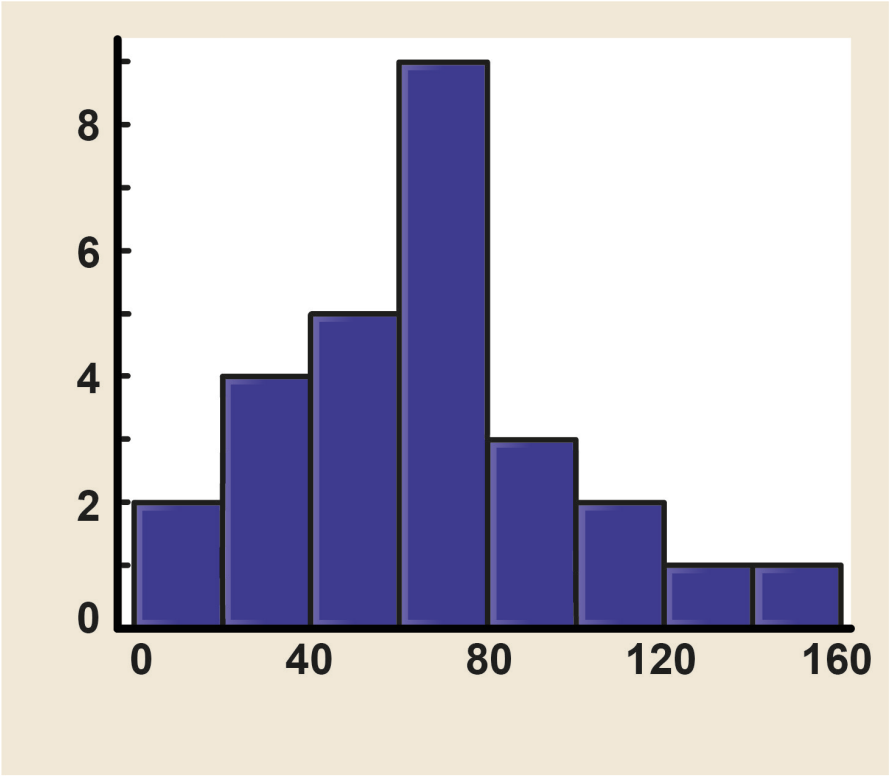


Data analysis





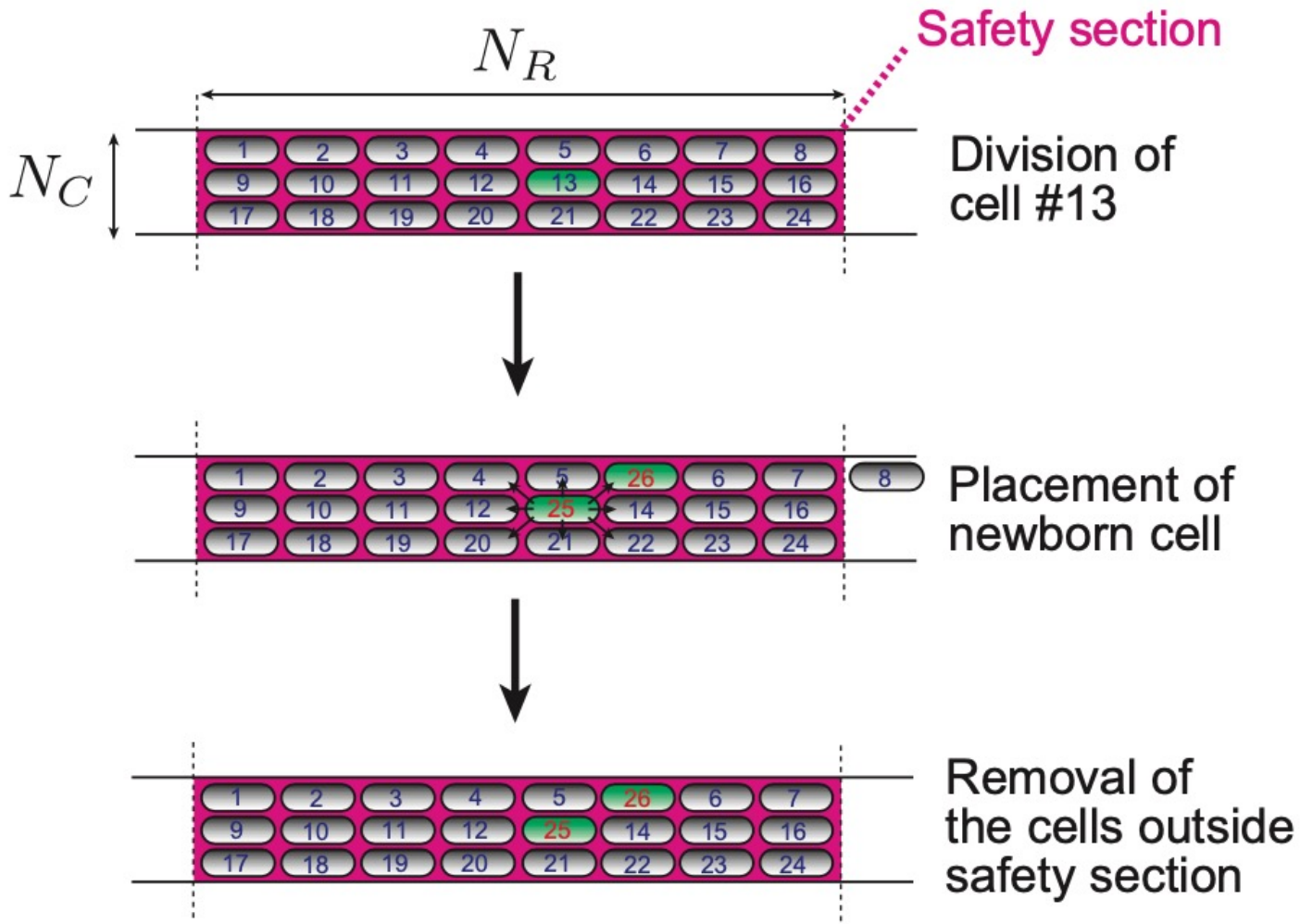




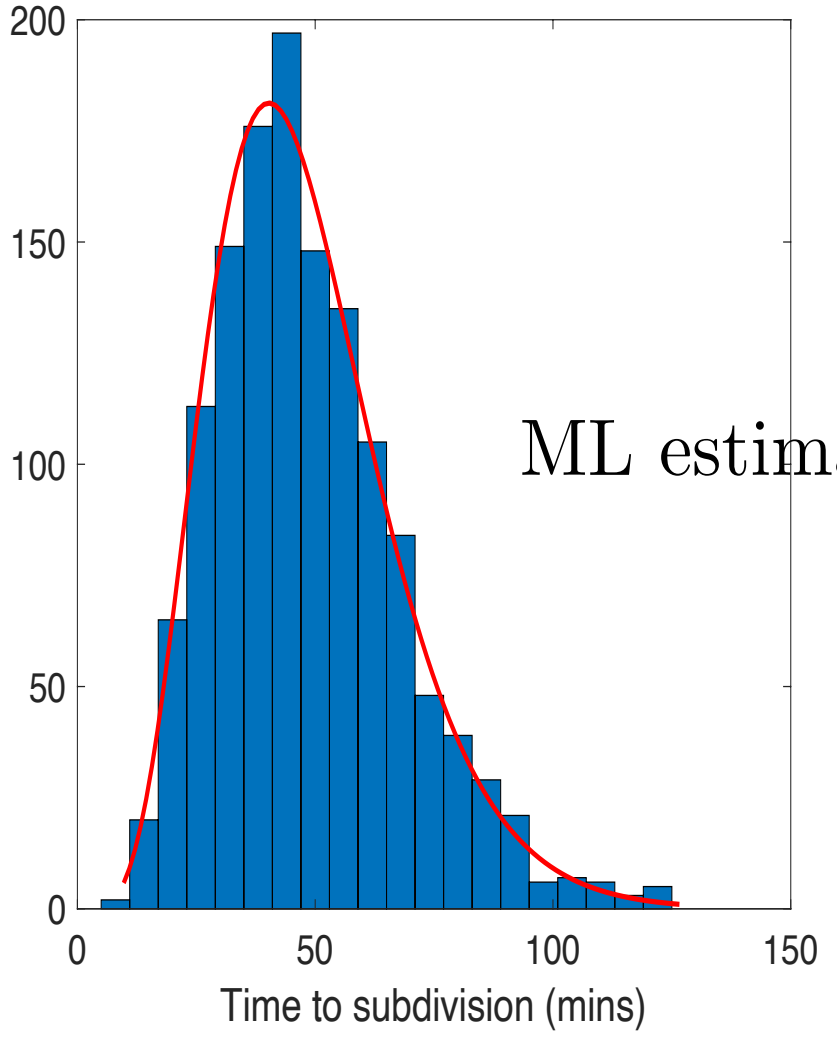
Time to subdivision (mins)

SuperSegger – Omnipose
MATLAB & Python based machine learning segmentation software

Simulation of dynamic cytometer



Homogeneous distribution



ML estimates

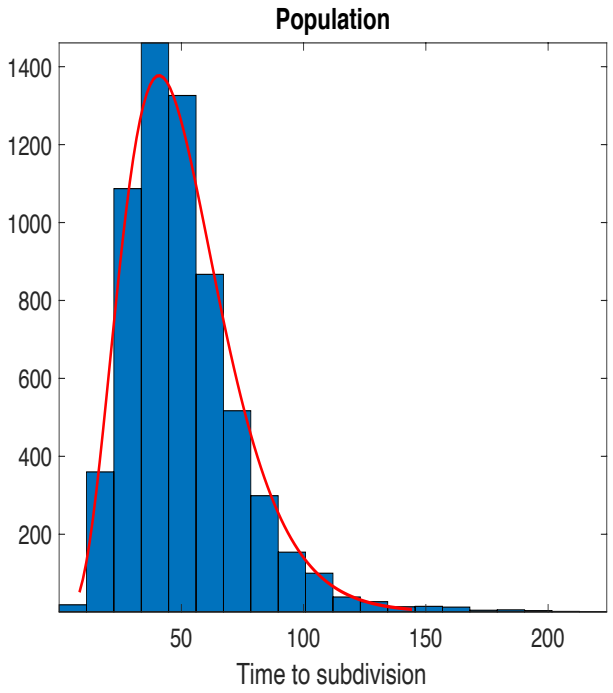
$$N_R = 10 \qquad N_C = 1$$

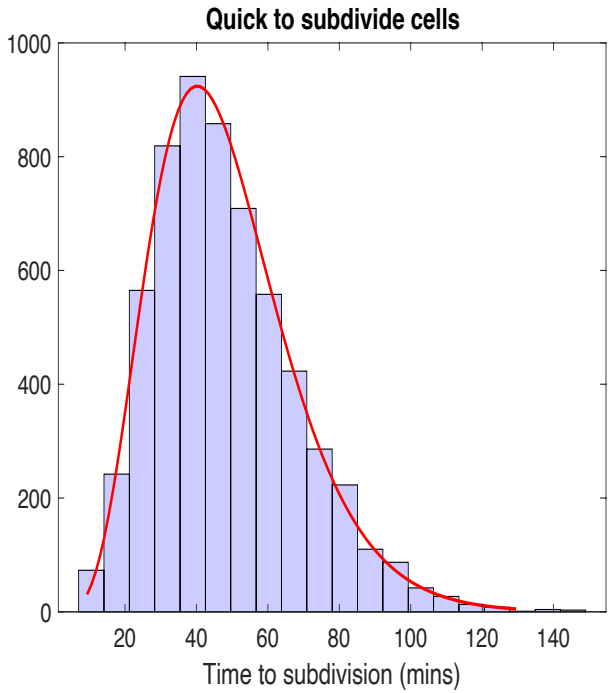
$$g(\tau) = \text{Gamma}(k, \theta) = \text{Gamma}(6, 9)$$

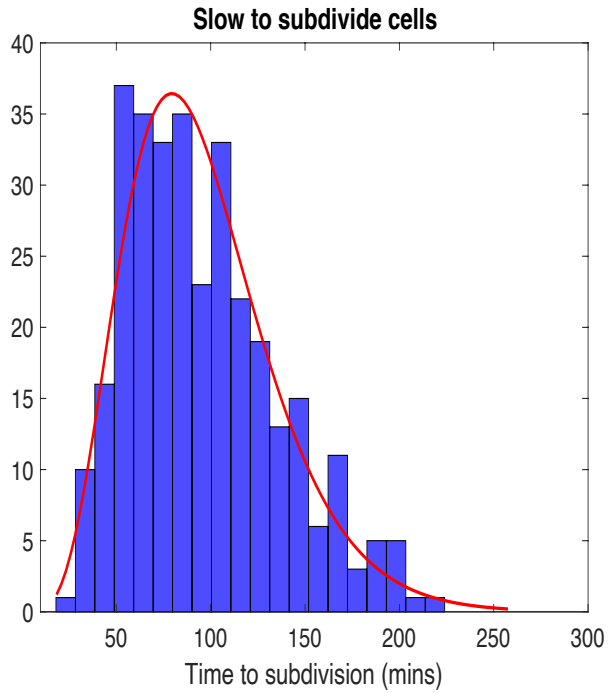
$$k = 6.3, \qquad CI = [5.8, 6.7]$$

$$\theta = 7.7, \qquad CI = [7.2, 8.3]$$

2-Phenotypes







Probability of slow cell birth

$p = 0.1$

$\mathrm{Gamma}(6, 9)$

$\mathrm{Gamma}(6, 20)$

Inference of heterogeneity

Assume cell time to subdivision x is generated from mixture of K models

$$p(x;\eta) = \sum_{l=1}^{K} w_l \, p(x;\theta_l)$$

$\theta_1, \ldots, \theta_K$ pdf parameters

$w_1, \ldots, w_K,$ non-negative weights with $\sum_{l=1}^{K} w_l = 1$

Model parameters

$$\eta = (\theta_1, w_1, \ldots, \theta_K, w_K)$$

Snob: finite mixture model (Matlab)

Assume two sub-populations:

Class 1: 85.8% (90%), k=6.6 (6), θ=7.2 (9)
Class 2: 14.2% (10%), k=4.1 (6), θ=18.0 (20)

Assume one population: k=5.1, θ=10

AIC and BIC (model selection)

$$\mathrm{AIC} = 2n_p - 2\ln(\hat{L})$$

$$\mathrm{BIC} = n_p \ln(N) - 2\ln(\hat{L})$$

$$\mathrm{AIC}_1 = 28164 \qquad \mathrm{AIC}_2 = 28092$$

$$\mathrm{BIC}_1 = 28171 \qquad \mathrm{BIC}_2 = 28108$$

Strong evidence data generated from of 2-phenotype model

Conclusions

- Range of modeling frameworks for heterogeneous population growth.
- Simple ODE model suggest mechanism for cell persistence.
- Homogeneous age-structured PDE model agrees well with Hashimoto.
- Included stress at birth. Selection effect depending on age.
- Extended to 2-phenotype age-structured. Reduced growth v homogeneous.
- Simple probabilistic model of dynamic cytometer.
- Synthetic data generated fitted by ML and finite mixture models
- Promising ability to do model selection AIC and BIC