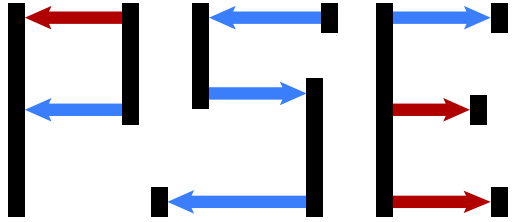
Maintenance of diversity in a parasite population capable of persistence and reinfection

joint work with Anton Wakolbinger

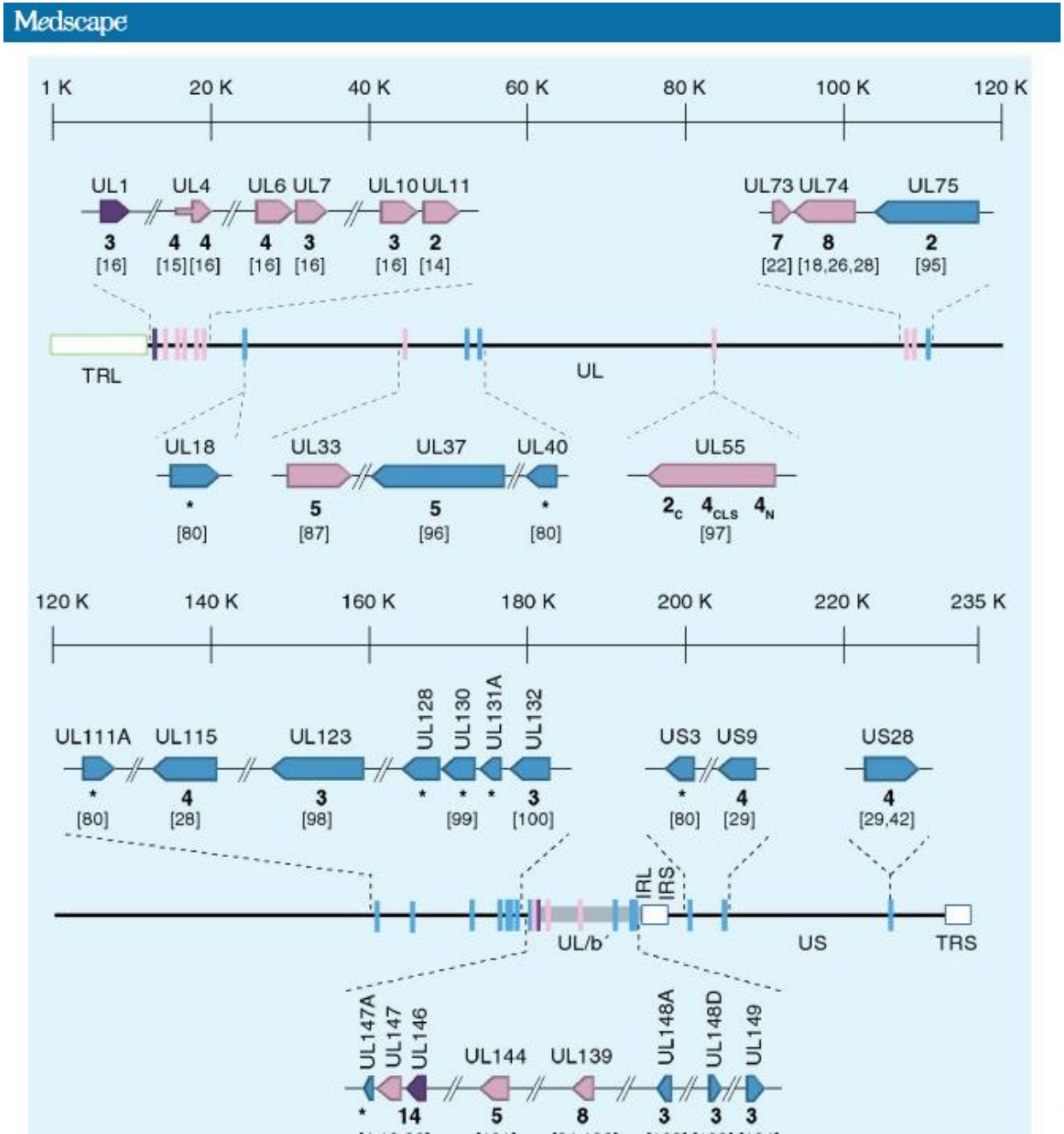
GOETHE UNIVERSITÄT FRANKFURT AM MAIN



Motivation: Human CytoMegaloVirus (HCMV)

- an (old) herpesvirus
- widespread in the human population (50%-90% infected worldwide)
- generally asymptomatic in the immunocompetent host
- dangerous for the immunocompromised hosts, like transplantation patients and fetuses

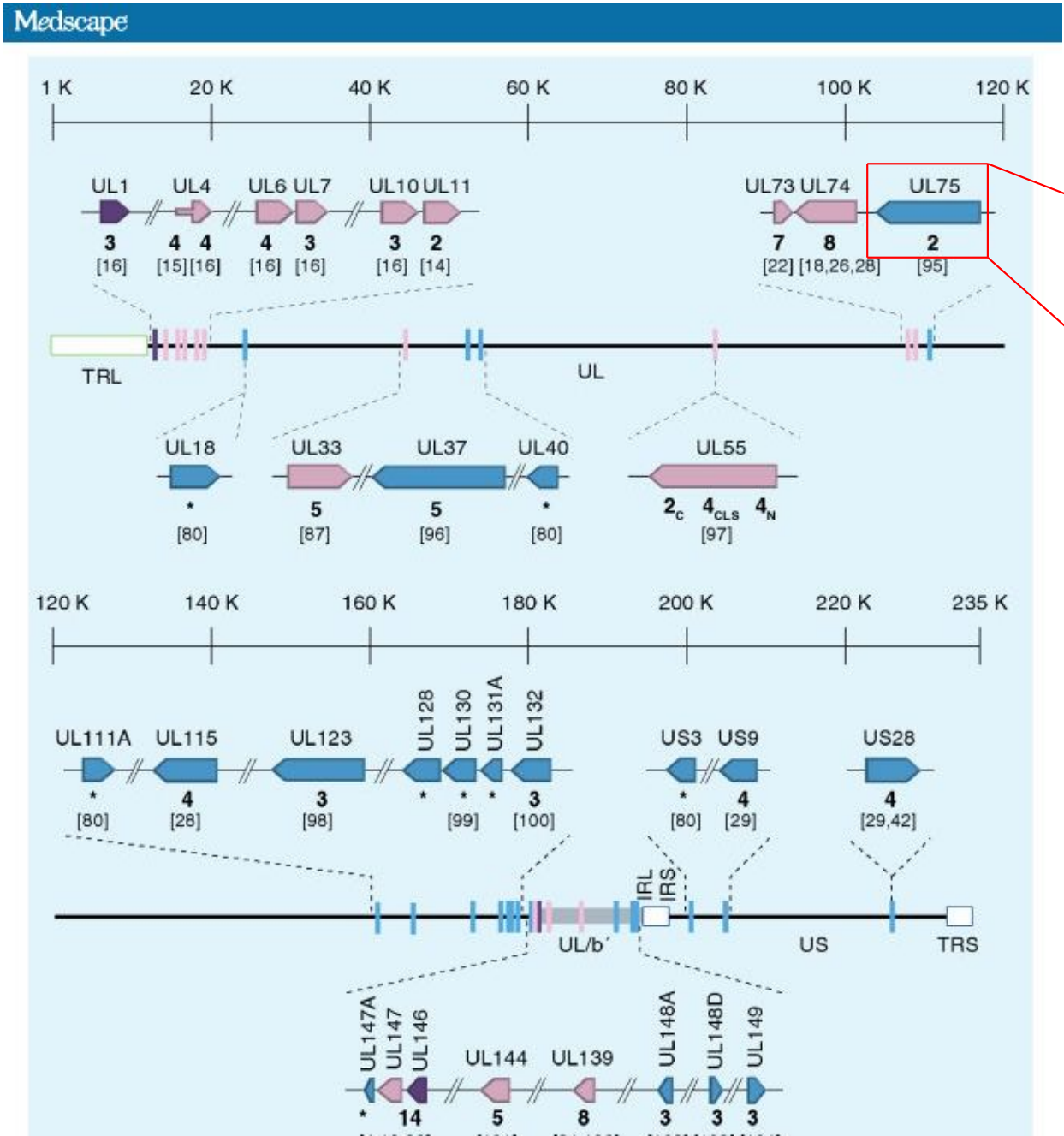
Diversity in coding regions

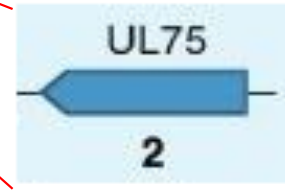


For many coding regions (loci) several types exist

Figure from Puchhammer-Stöckl and Görzer (2011)

Diversity in coding regions





For many coding regions (loci) several types exist

Figure from Puchhammer-Stöckl and Görzer (2011)

Fitness valley between genotypes

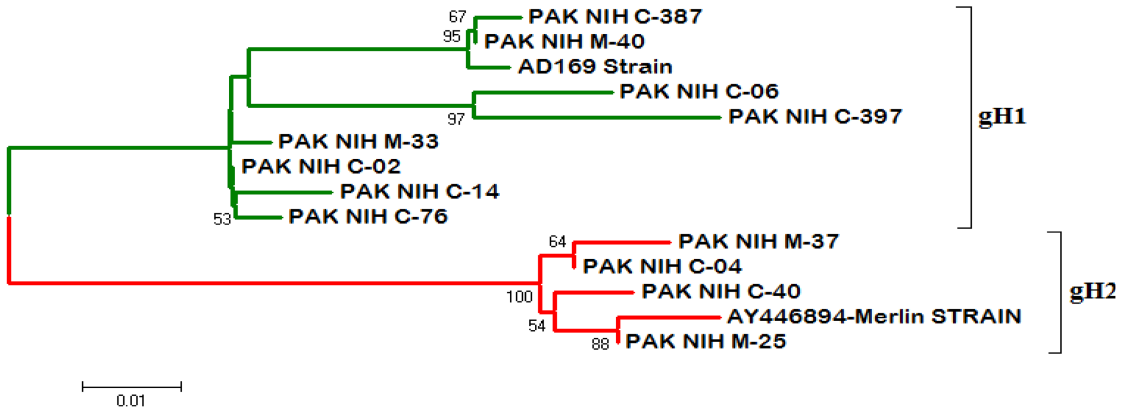


Figure from Mujtaba et al. (2016)

- Genotypes are phylogenetically distant
- e.g. UL 75 (also called gH) genotypes are distinguished by 1 deletion and 4 non-synonymous point mutations
- fitness landscape with a few, high peaks
- rare mutation between genotypes

How is this diversity maintained in the HCMV population?

Relevant characteristics of HCMV

- HCMV persists in its host lifelong

Figure downloaded from medigapcheap.com

Relevant characteristics of HCMV

- HCMV persists in its host lifelong

Figure downloaded from medigapcheap.com

Relevant characteristics of HCMV

- HCMV persists in its host lifelong
- a host can be reinfected by the virus

Figure downloaded from medigapcheap.com

Host types

TABLE I. Distribution of CMV gH Genotypes in Different Patient Groups

	Prevalence of gH genotypes in patients, n (%)		
Genotype	Newborns	Infants	Adults
gH1	17 (40.5)	31 (33.3)	14 (25.4)
gH2	19 (45.2)	40 (43.0)	21 (38.2)
gH1 + gH2	6 (14.3)	22 (23.7)	20 (36.4)
Total	42	93	55

n, number of patients.

Table from Paradowska et al. (2014)

- there exist hosts infected with a single type only
- there exist hosts infected with both types

Role of persistence and reinfection on parasite survival

Assumption: Diversity of the parasite population within hosts is advantageous for parasite survival

- De novo mutation? Too aggressive?!
- Persistence and reinfection?
  - Introduce the diversity of the surrounding parasite population into a single host step by step
  - Less aggressive?!

Modeling the evolution of a parasite population with persistence and reinfection

A parasite population is distributed over (infected) hosts.

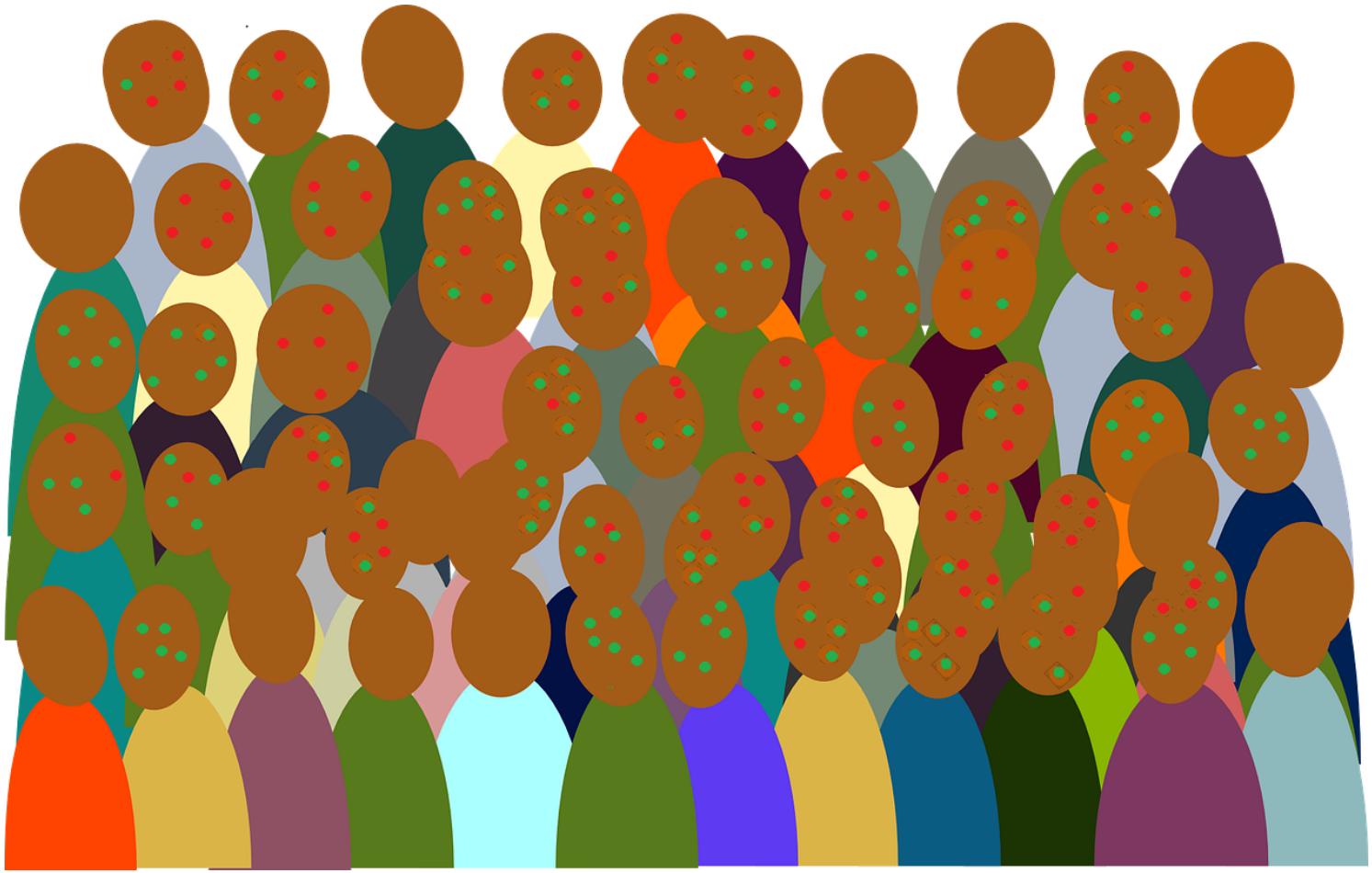


Figure downloaded from pixabay.com and modified

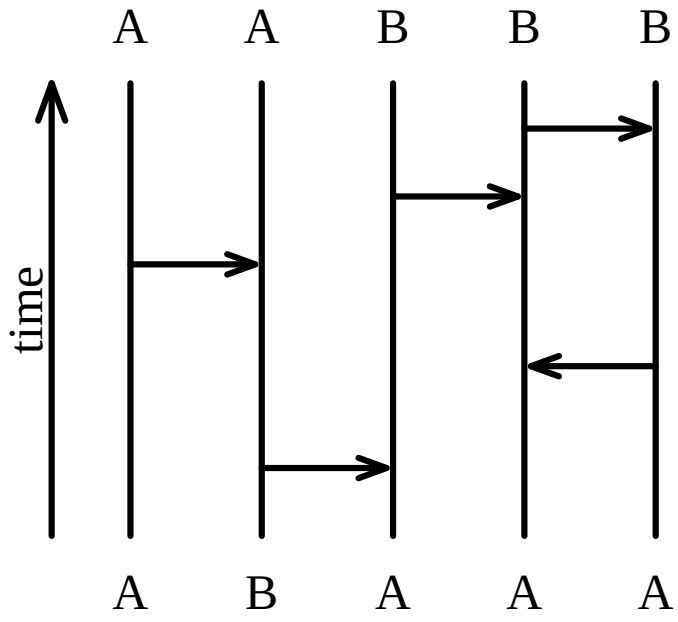
Modeling the evolution of a parasite population with persistence and reinfection

- M infected hosts
- within each (infected) host well-mixed (panmictic) parasite populations of constant size N
- two parasite types A and B

Factors driving the evolution

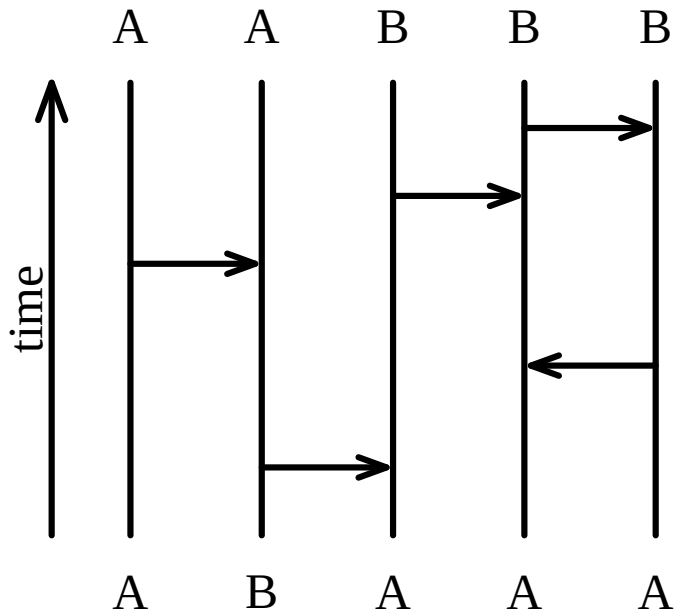
- reinfection
- host replacement
  (hosts die and new hosts are primary infected)
- balancing selection within hosts
  (maintains diversity in a host)

Reproduction of parasites



- population of constant size N
- two types A and B
- X_i^N - frequency of type A in host i
- Moran model: At reproduction a parasite splits into two and a randomly chosen parasite is removed
- no mutation

Balancing selection within hosts



- $s_N > 0$ - selection strength
- $\eta \in (0,1)$ - quasi-equilibrium frequency of type A
- parasites reproduce g_N-times faster than hosts
- in host i, parasites of type A reproduce at rate

$$g_N(1 + s_N(\eta - X_i^N))$$

- in host i, parasites of type B reproduce at rate

$$g_N(1 + s_N((1-\eta) - (1 - X_i^N)))$$

Adding the dynamics of hosts

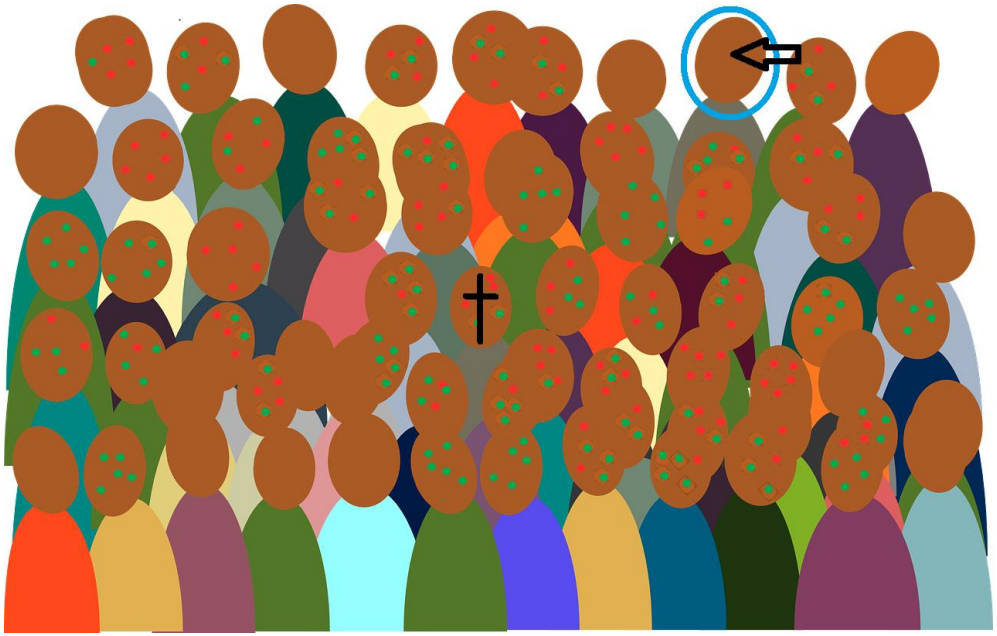


Figure downloaded from pixabay.com and modified

- Each host dies at rate 1 and is replaced by a so far uninfected host which is instantly primary infected $\rightarrow$ constant number M of infected hosts
- at infection only a single type is transmitted
- the type is chosen randomly from the infecting host

Adding the dynamics of hosts

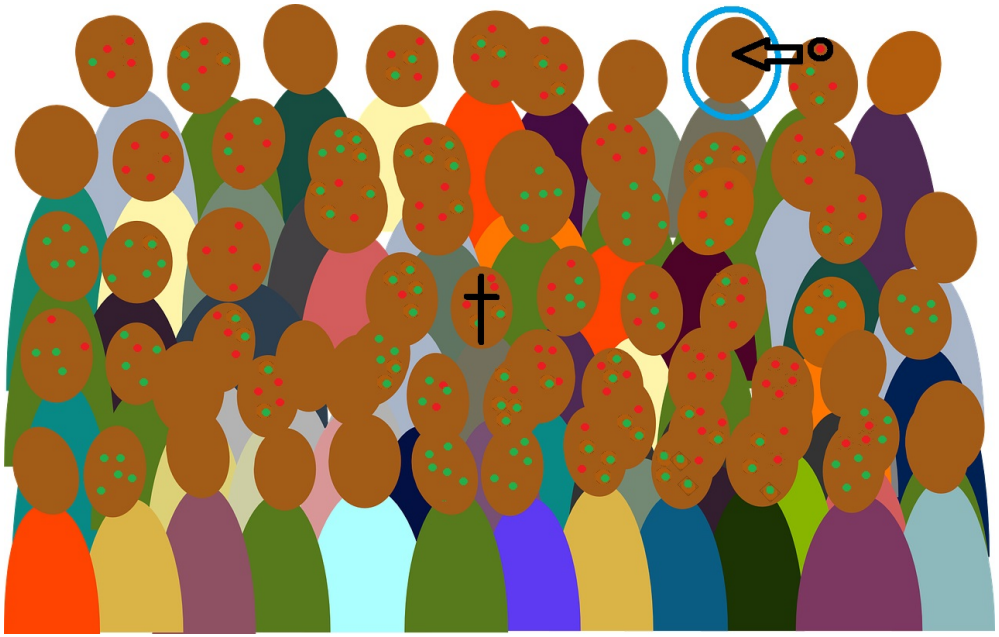


Figure downloaded from pixabay.com and modified

- Each host dies at rate 1 and is replaced by a so far uninfected host which is instantly primary infected $\rightarrow$ constant number M of infected hosts
- at infection only a single type is transmitted
- the type is chosen randomly from the infecting host

Adding the dynamics of hosts

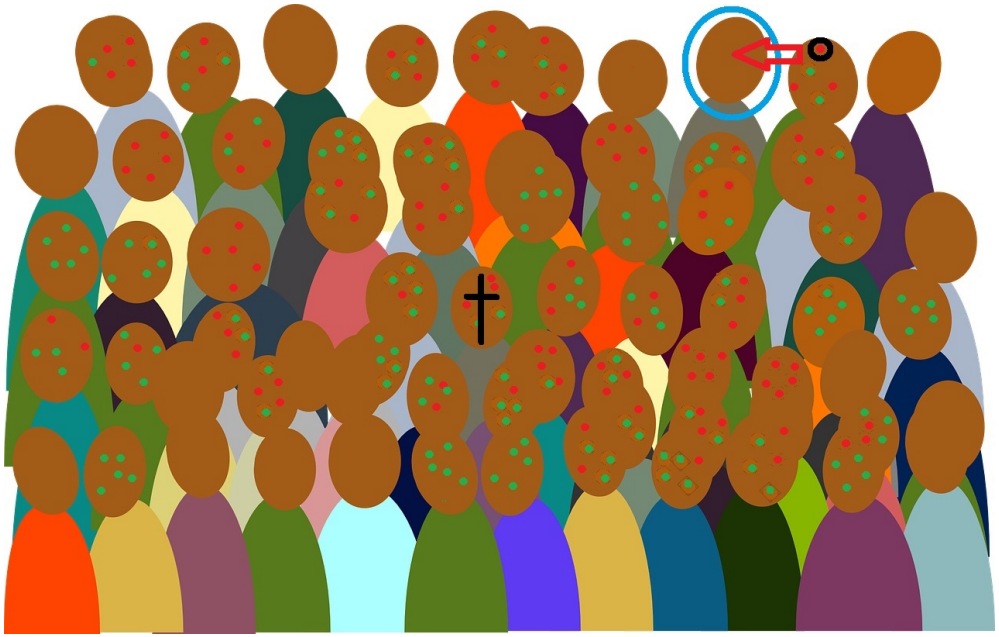


Figure downloaded from pixabay.com and modified

- Each host dies at rate 1 and is replaced by a so far uninfected host which is instantly primary infected $\rightarrow$ constant number M of infected hosts
- at infection only a single type is transmitted
- the type is chosen randomly from the infecting host

Adding the dynamics of hosts

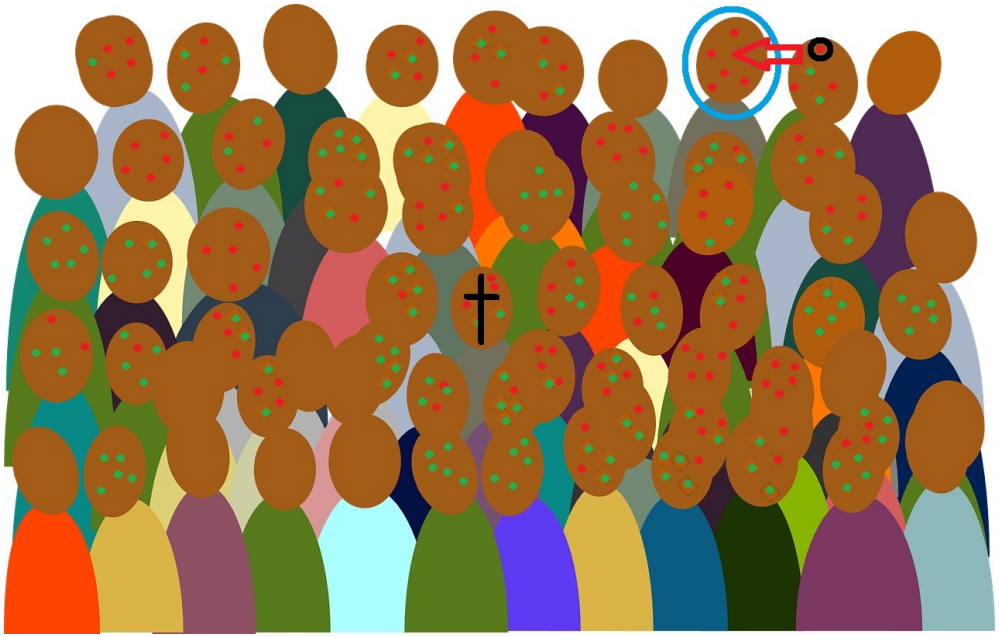


Figure downloaded from pixabay.com and modified

- Each host dies at rate 1 and is replaced by a so far uninfected host which is instantly primary infected $\rightarrow$ constant number M of infected hosts
- at infection only a single type is transmitted
- the type is chosen randomly from the infecting host

Adding reinfection

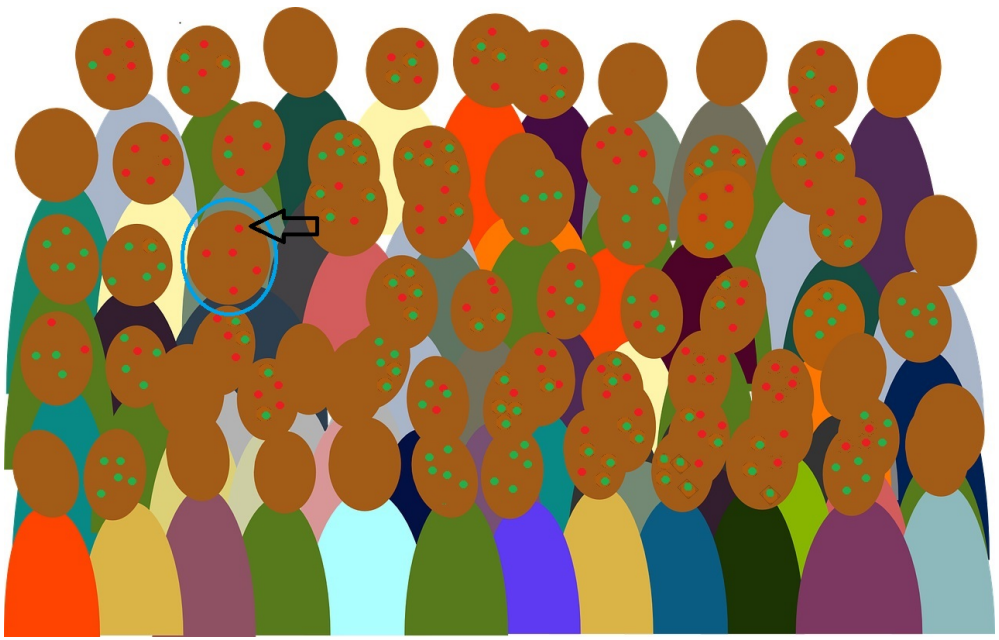


Figure downloaded from pixabay.com and modified

- Reinfection happens at rate r_N per host
- At reinfection the reinfecting host transmits a single parasite to the reinfected host

Adding reinfection

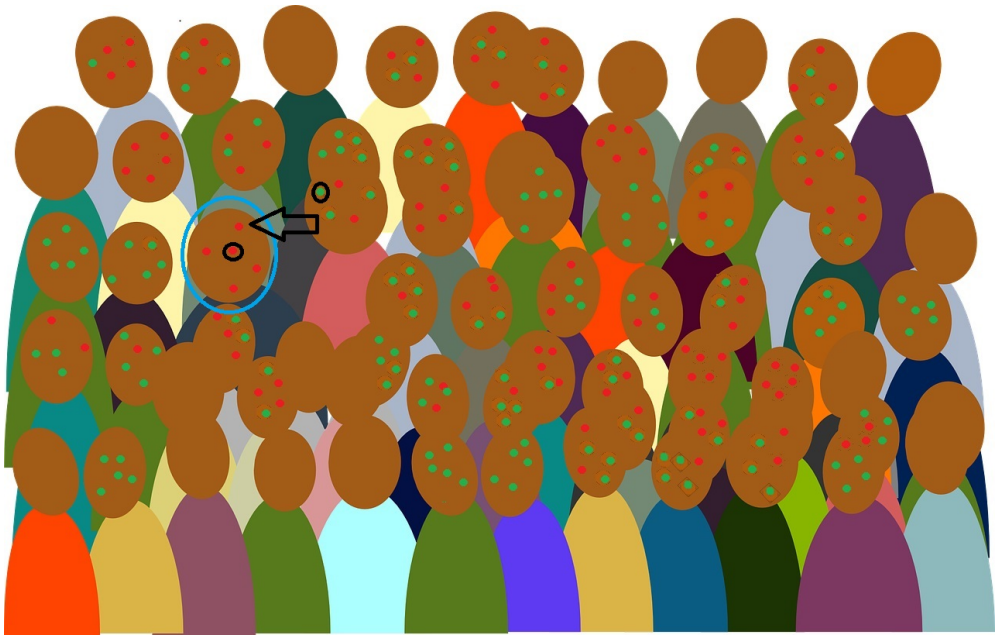


Figure downloaded from pixabay.com and modified

- Reinfection happens at rate r_N per host
- At reinfection the reinfecting host transmits a single parasite to the reinfected host
- the type of the transmitted parasite is chosen randomly from the reinfecting host
- in the reinfected host a single parasite is replaced by the transmitted parasite

Adding reinfection

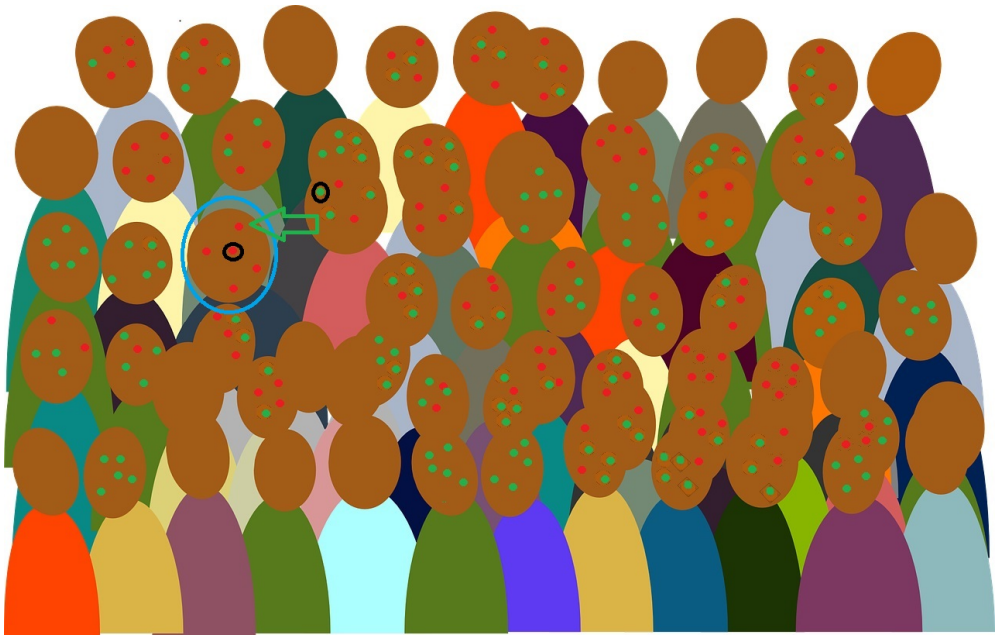


Figure downloaded from pixabay.com and modified

- Reinfection happens at rate r_N per host
- At reinfection the reinfecting host transmits a single parasite to the reinfected host
- the type of the transmitted parasite is chosen randomly from the reinfecting host
- in the reinfected host a single parasite is replaced by the transmitted parasite

Adding reinfection

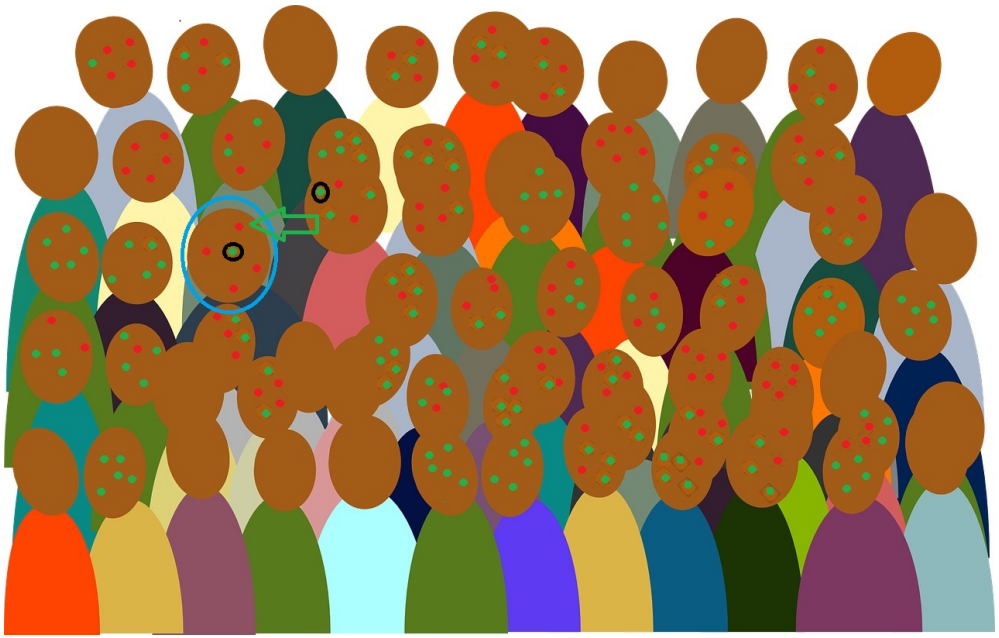


Figure downloaded from pixabay.com and modified

- Reinfection happens at rate r_N per host
- At reinfection the reinfecting host transmits a single parasite to the reinfected host
- the type of the transmitted parasite is chosen randomly from the reinfecting host
- in the reinfected host a single parasite is replaced by the transmitted parasite

The dynamics of the parasite frequencies in the hosts

$\mathbf{X}^{N,M} = (X_1^M, ..., X_M^M)$ - frequencies of type A in hosts $1, ..., M$

$\mathbf{X}^{N,M} = (X_1^M, ..., X_M^M)$ is a Markovian jump process with jumps from

$$\mathbf{x} = (x_1, ..., x_i, ..., x_M) \text{ to}$$

$$(x_1, ..., x_i + \tfrac{1}{N}, ..., x_M) \quad \text{at rate} \quad g_N(1 + s_N(\eta - x_i))Nx_i(1 - x_i) + r_N \tfrac{1}{M} \sum_{j=1}^{M} x_j(1 - x_i)$$

$$(x_1, ..., x_i - \tfrac{1}{N}, ..., x_M) \quad \text{at rate} \quad g_N(1 + s_N(x_i - \eta))Nx_i(1 - x_i) + r_N \tfrac{1}{M} \sum_{j=1}^{M} (1 - x_j)x_i$$

$$(x_1, ..., 1, ..., x_M) \quad \text{at rate} \quad \bar{\mathbf{x}} = \tfrac{1}{M} \sum_{j=1}^{M} x_j$$

$$(x_1, ..., 0, ..., x_M) \quad \text{at rate} \quad 1 - \bar{\mathbf{x}}$$

Asymptotic behaviour as $N \to \infty$?

Aim:

Find a parameter regime

- with a long maintenance of both parasites types
- with hosts carrying a single parasite type as well as hosts carrying both types (as observed in samples of HCMV hosts)

Ingredients

- Host replacement: Creates monomorphic hosts
- Reinfection: Creates polymorphic hosts

Choose parameters r_N, s_N, g_N such that

- effective reinfection events and host replacements appear on the same time scale

Conditions C

- (moderate selection)

$$s_N = N^{-\epsilon}$$

for $0 < \epsilon < \frac{1}{5}$.
In particular, as $N \to \infty$

$$s_N \to 0,$$

but

$$Ns_N \to \infty.$$

- (frequent reinfection)

$$r_N s_N \xrightarrow{N\to\infty} r$$

for some $r > 0$.

- (fast viral reproduction)

$$N^{5\epsilon} \ll g_N \ll \exp(N^{1/5}).$$

Theorem

Let $M \in \mathbb{N}$. Assume Conditions C are valid and

$$\mathbf{X}^{N,M}(0) \xrightarrow{\text{weakly}} \rho_0$$

as $N \to \infty$, for some distribution ρ_0 concentrated on $(\{0\} \cup [\delta, 1-\delta] \cup \{1\})^M$ for some $0 < \delta < 1$.

Then $\mathbf{X}^{N,M}$ converges on each time interval $(0, t]$ in distribution (in the Skorohod topology) to the process $\mathbf{Y}^M$, with $\mathbf{Y}^M(0)$ being the image of the ρ_0 under the map $0 \mapsto 0, 1 \mapsto 1, [\delta, 1-\delta] \ni x \mapsto \eta$.

The process $\mathbf{Y}^M$ is of the following form:

The process $\mathbf{Y}^M = (Y_1^M, ..., Y_M^M)$ starting in $\mathbf{y}^0$ is a pure jump process on $\{0, \eta, 1\}^M$ with

jumps from $y = (y_1, ..., 1, ..., y_M)$ to

$(y_1, ..., 0, ..., y_M)$	at rate	$1 - \bar{\mathbf{y}}$	(host replacement)
$(y_1, ..., \eta, ..., y_M)$	at rate	$2r(1-\eta)(1-\bar{\mathbf{y}})$	(effective reinfection)

jumps from $y = (y_1, ..., 0, ..., y_M)$ to

$(y_1, ..., 1, ..., y_M)$	at rate	$\bar{\mathbf{y}}$	(host replacement)
$(y_1, ..., \eta, ..., y_M)$	at rate	$2r\eta\bar{\mathbf{y}}$	(effective reinfection)

jumps from $y = (y_1, ..., \eta, ..., y_M)$ to

$(y_1, ..., 1, ..., y_M)$	at rate	$\bar{\mathbf{y}}$	(host replacement)
$(y_1, ..., 0, ..., y_M)$	at rate	$1 - \bar{\mathbf{y}}$	(host replacement).

Sketch of the proof

X_i^N, $i = 1, ..., M$, **are concentrated on the states 0, 1 und** $(\eta - s_N, \eta + s_N)$

- A reinfection is effective if in the reinfected host the neighbourhood $(\eta - s_N^{3/2}, \eta + s_N^{3/2})$ of the frequency η is reached.
- Parasite frequencies within hosts are only correlated by reinfection and host replacement events

In a single host ...

without host replacement

- Ineffective excursions are short, i.e. $\ll N^{2\epsilon}/g_N$
- Ineffective excursions are "rare" on the host time scale (they do not overlap)
- Time to balance is short:
  After an effective reinfection the neighbourhood $(\eta - s_N^{3/2}, \eta + s_N^{3/2})$ is reached after time $\ll N^{4\epsilon}/g_N$
- Then the frequency remains within $(\eta - s_N, \eta + s_N)$ for a long time $\gg \exp(N^{3/10})/g_N$

In a single host ...

without host replacement

- Ineffective excursions are short, i.e. $\ll N^{2\epsilon}/g_N$
- Ineffective excursions are "rare" on the host time scale (they do not overlap)
- Time to balance is short:
  After an effective reinfection the neighbourhood $(\eta - s_N^{3/2}, \eta + s_N^{3/2})$ is reached after time $\ll N^{4\epsilon}/g_N$
- Then the frequency remains within $(\eta - s_N, \eta + s_N)$ for a long time $\gg \exp(N^{3/10})/g_N$

Since $g_N \gg N^{5\epsilon}$ as $N \to \infty$

- Ineffective excursions disappear
- Effective reinfection events result in a jump to frequency η

In a single host ...

without host replacement

- Ineffective excursions are short, i.e. $\ll N^{2\epsilon}/g_N$
- Ineffective excursions are "rare" on the host time scale (they do not overlap)
- Time to balance is short:
  After an effective reinfection the neighbourhood $(\eta - s_N^{3/2}, \eta + s_N^{3/2})$ is reached after time $\ll N^{4\epsilon}/g_N$
- Then the frequency remains within $(\eta - s_N, \eta + s_N)$ for a long time $\gg \exp(N^{3/10})/g_N$

Since $g_N \ll \exp(N^{1/5})$ as $N \to \infty$

- $(\eta - s_N, \eta + s_N)$ is left only due to a host replacement event and not due to random fluctuations

"Correct" rates

- Reinfections in single hosts do not overlap
- "Probability to balance"

$$2\eta s_N + o(s_N),$$

if a pure B-type host was infected with type A, and

$$2(1-\eta)s_N + o(s_N),$$

if a pure A-type host with reinfected with type B.

$\Rightarrow$ Effective reinfection rate

$$2\eta s_N r_N \to 2\eta r \text{ and } 2(1-\eta)r$$

The process $\mathbf{Y}^M = (Y_1^M, ..., Y_M^M)$ starting in $\mathbf{y}^0$ is a pure jump process on $\{0, \eta, 1\}^M$ with

jumps from $y = (y_1, ..., 1, ..., y_M)$ to

$(y_1, ..., 0, ..., y_M)$	at rate	$1 - \bar{\mathbf{y}}$	(host replacement)
$(y_1, ..., \eta, ..., y_M)$	at rate	$2r(1-\eta)(1-\bar{\mathbf{y}})$	(effective reinfection)

jumps from $y = (y_1, ..., 0, ..., y_M)$ to

$(y_1, ..., 1, ..., y_M)$	at rate	$\bar{\mathbf{y}}$	(host replacement)
$(y_1, ..., \eta, ..., y_M)$	at rate	$2r\eta\bar{\mathbf{y}}$	(effective reinfection)

jumps from $y = (y_1, ..., \eta, ..., y_M)$ to

$(y_1, ..., 1, ..., y_M)$	at rate	$\bar{\mathbf{y}}$	(host replacement)
$(y_1, ..., 0, ..., y_M)$	at rate	$1 - \bar{\mathbf{y}}$	(host replacement).

Coupling

Define an approximation $\hat{\mathbf{X}}^{N,M}$ of $\mathbf{X}^{N,M}$.
$\hat{\mathbf{X}}^{N,M}$ has the same dynamics as $\mathbf{X}^{N,M}$, with modified reinfections:

- Ignore a reinfection event
  if the hitted host is not in state 0 or 1.
- Otherwise, if the reinfected host is in state 0 (resp. 1) and the transmitted type is 1 (resp. 0), toss a coin which shows up head with probability $2s_N\eta$ (resp. with prob. $2s_N(1-\eta)$).
- If the coin shows tail, ignore the reinfection.
- If the coin shows head, start a transition from $1/N$ (resp. $1-1/N$) to $\eta - s_N^{3/2}$ (resp. $\eta + s_N^{3/2}$).

Coupling

Then

- the finite dimensional distributions (fdd) of $\hat{\mathbf{X}}^{N,M}$ converge to those of $\mathbf{Y}^{\mathbf{M}}$
- $\hat{\mathbf{X}}^{N,M}$ and $\mathbf{X}^{N,M}$ have the same limiting fdd

Hence, the convergence follows by showing that $\mathbf{X}^{N,M}$ is tight.

Mean field limit for $M \to \infty$

Consider $\mathbf{Z}^M = (Z_t^0, Z_t^\eta, Z_t^1)_{t\geq 0}$ with

$$Z_t^\ell := \frac{|\{i \in \{1, ..., M\} | Y_i^M(t) = \ell\}|}{M},$$

the frequency of type-ℓ hosts in the population at time t, for $\ell \in \{0, \eta, 1\}$.
Assume $\lim_{M\to\infty} \mathbf{Z}_0^M = \mathbf{v}^0 \in S^3$.
For $M \to \infty$ the process $\mathbf{Z}^M$ converges to the deterministic dynamical system $\mathbf{v} = (v_t^0, v_t^\eta, v_t^1)_{t\geq 0}$ with

$$\begin{aligned}
\dot{v}^0 &= (1-\eta)v^\eta - 2r\eta v^0(v^1 + \eta v^\eta) \\
\dot{v}^\eta &= -v^\eta + 2r(\eta^2 v^0 v^\eta + (1-\eta)^2 v^1 v^\eta + v^0 v^1) \\
\dot{v}^1 &= \eta v^\eta - 2r(1-\eta)v^1(v^0 + (1-\eta)v^\eta).
\end{aligned}$$

started in $\mathbf{v}^0$.

Propagation of chaos: Typical host type frequency process

Let $k \in \mathbb{N}$. Denote by $V_i = \lim_{M\to\infty} Y_i^M$ for $i \in \{1, ..., k\}$. Then V_i are independent copies of the jump process V which jumps at time s from any state to state

$$\begin{array}{lll} 0 & \text{at rate} & v_s^0 + (1-\eta)v_s^\eta \\ 1 & \text{at rate} & v_s^1 + \eta v_s^\eta, \end{array}$$

from state 0 to state η at rate $2r\eta(v_s^1 + \eta v_s^\eta)$, and
from state 1 to state η at rate $2r(1-\eta)(v_s^0 + (1-\eta)v_s^\eta)$.

Remark:
Also in the case $M = M_N$ and $M_N \xrightarrow{N\to\infty} \infty$ this propagation of chaos holds.

Fixed points and stability of the dynamical system

The fixed points of the dynamical system $\mathbf{v}$ are

$$(1,0,0)\ ,(0,0,1),$$

and $\mathbf{u} = (u^0, u^\eta, u^1)$ with

$$u^0 = \frac{2r\eta(1-\eta)^2 - (2\eta-1)}{2r\eta^2 + 4r^2\eta^3(1-\eta)} \qquad u^1 = \frac{2r(1-\eta)\eta^2 + 2\eta - 1}{2r(1-\eta)^2 + 4r^2\eta(1-\eta)^3}$$

$$u^\eta = \frac{4r^2\eta^3(1-\eta)^3 - (2\eta-1)^2(2r\eta(1-\eta)+1)}{2r\eta^2(1-\eta)^2(1+2r\eta(1-\eta)}.$$

If $r > \max\{\frac{2\eta-1}{2\eta(1-\eta)^2}, \frac{1-2\eta}{2\eta^2(1-\eta)}\}$, then
$\mathbf{u} \in \operatorname{int} S^3$ and $\mathbf{u}$ is on $S^3 \backslash \{(0,0,1),(1,0,0)\}$ a globally stable equilibrium; (0,0,1) and (1,0,0) are saddle points.

Maintenance of a polymorphic state

- For finite N and finite $M = M_N$ eventually one type gets lost due to random fluctuations
- Add a (small) mutation rate μ_N at which parasites mutate (on the host time scale, population wide mutation rate). Eventually this turns a monomorphic parasite population to a polymorphic one.

How long does it take until a polymorphic state is reached from a monomorphic one and vice versa?

Switching between monomorphic and polymorphic states

- Let

$$T_{\text{mono}} = \inf\{t > 0 | \bar{\mathbf{X}}^{N,M}(t) \in \{0, 1\}\}$$

- For $\delta > 0$ let

$$T^{\delta}_{\text{poly}} = \inf\{t > 0 | \bar{\mathbf{X}}^{N,M}(t) > \delta \text{ and } 1 - \bar{\mathbf{X}}^{N,M}(t) > \delta\}$$

with $\bar{\mathbf{X}}^{N,M}(t) = \frac{\sum_{i=1}^{M} X_i^N(t)}{M}$.

Switching between monomorphic and polymorphic states

Theorem

Assume Conditions (C), $\mu_N \ll r_N$ and $r > \max\{\frac{\eta}{2(1-\eta)^2}, \frac{1-\eta}{2\eta^2}\}$.
Then for any $\gamma > 0$:

- If $\mathbf{X}^{N,M}$ is started in a monomorphic state, for any $\delta > 0$

$$T^{\delta}_{\text{poly}} = O(\frac{1}{\mu_N s_N}) + O(M_N^{\gamma}).$$

- If $\mathbf{X}^{N,M}$ is started in a polymorphic state,

$$T_{\text{mono}} \gg \exp(M_N^{1-\gamma})$$

Switching between monomorphic and polymorphic states

Theorem

Assume Conditions (C), $\mu_N \ll r_N$ and $r > \max\{\frac{\eta}{2(1-\eta)^2}, \frac{1-\eta}{2\eta^2}\}$. Then for any $\gamma > 0$:

- If $\mathbf{X}^{N,M}$ is started in a monomorphic state, for any $\delta > 0$

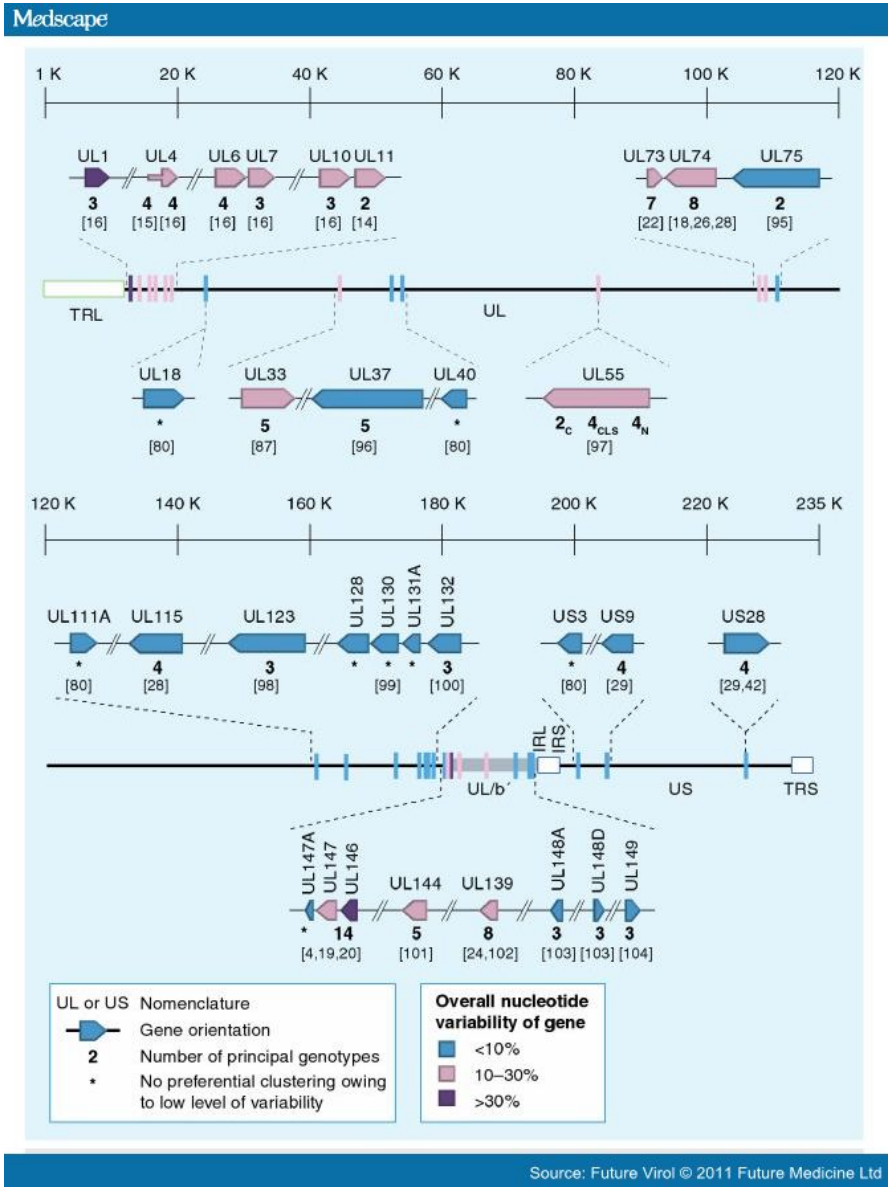
$$T^{\delta}_{\text{poly}} = O(\frac{1}{\mu_N s_N}) + O(M_N^{\gamma}).$$

- If $\mathbf{X}^{N,M}$ is started in a polymorphic state,

$$T_{\text{mono}} \gg \exp(M_N^{1-\gamma})$$

If $\frac{1}{\exp(M_N^{1-\gamma})s_N} \ll \mu_N \ll r_N$, then $T^{\delta}_{\text{poly}} \ll T_{\text{mono}}$. In this case both types coexist most of the time in the parasite population.

Perspective



Persistence and reinfection are effective mechanisms

- to introduce diversity into parasite populations of single hosts
- maintain diversity in the parasite population also in the case of small mutation rates

Figure from Puchhammer-Stöckl and Görzer (2011)

Thank you!

Literature:

P., Wakolbinger: Maintenance of diversity in a hierarchical host-parasite model with balancing selection and reinfection, submitted, Preprint: arXiv:1802.02429.