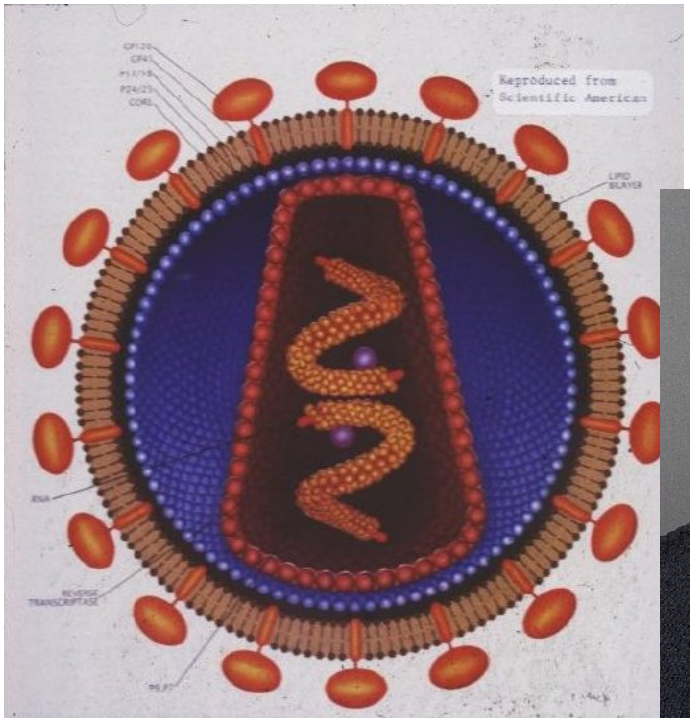
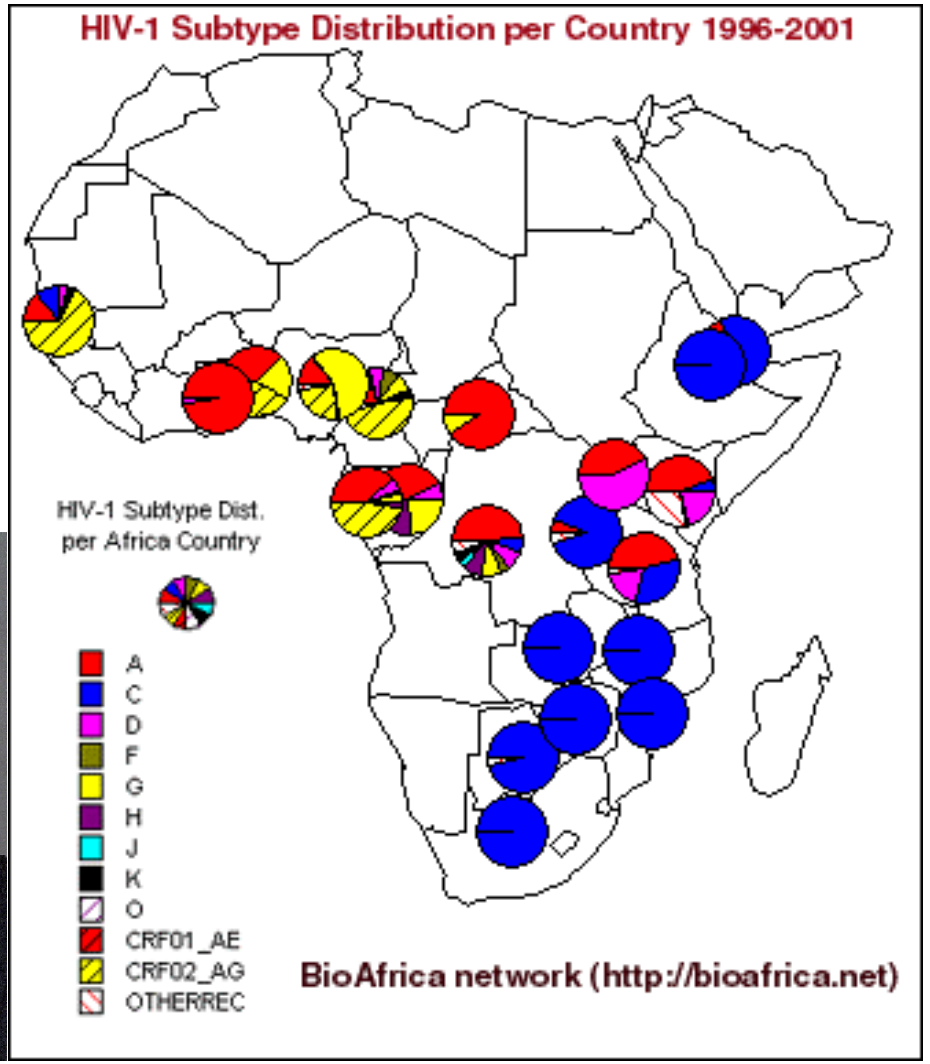
Modeling HIV Quasispecies Dynamics

Pietro Lio`
Computer Laboratory
University of Cambridge

- Artificial and Bioinformatics Group at Computer Laboratory (with J. Daugman, S. Holden and M. Jamnick).
- Mphil in Comp Biol (S. Tavare` and S. Eglen)
- CCBI (G. Micklem and S. Tavare`)
- Fitzwilliam College

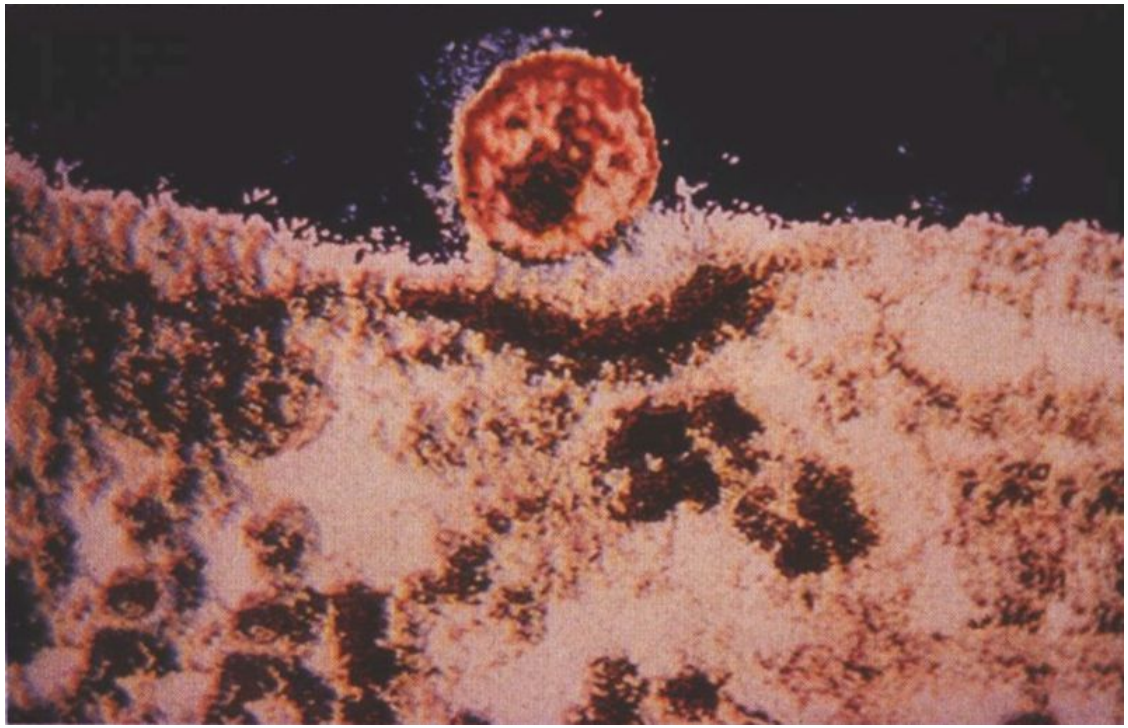
The coexistence and co-evolution of different strains of a virus within the same host are nowadays recorded for a growing number of viruses, most notably HIV,Marburg, Ebola and influenza, but the conditions for the formation and survival of new strains have not yet been understood.





How does one counsel two HIV-infected individuals who are in a monogamous relationship? Do they really need to still use condoms?

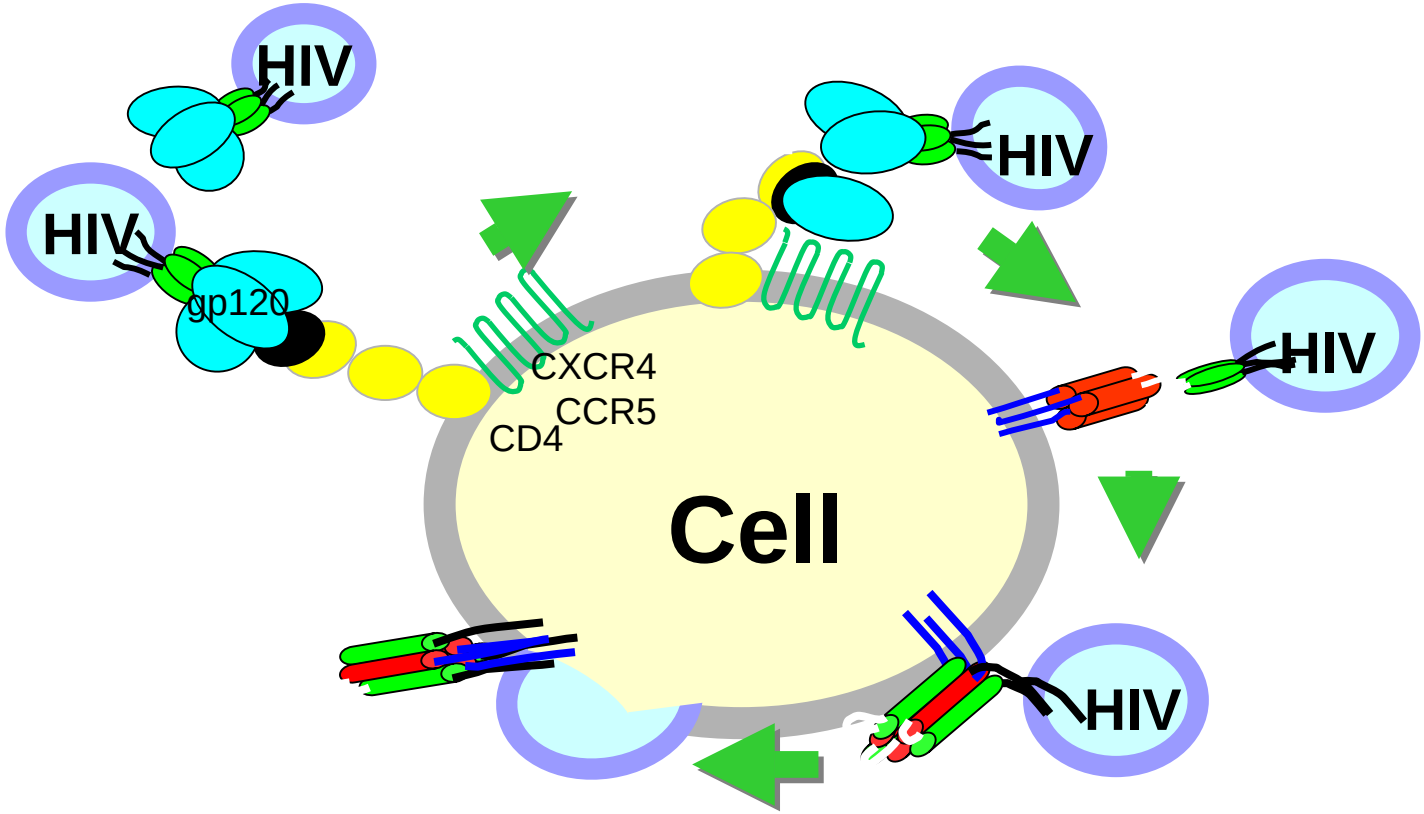
The answer to this complicated question depends in part on whether or not an HIV-infected patient can be infected with another HIV strain. It is now known that dual HIV infection, defined as the presence of two distinct HIV strains in a patient, can occur.



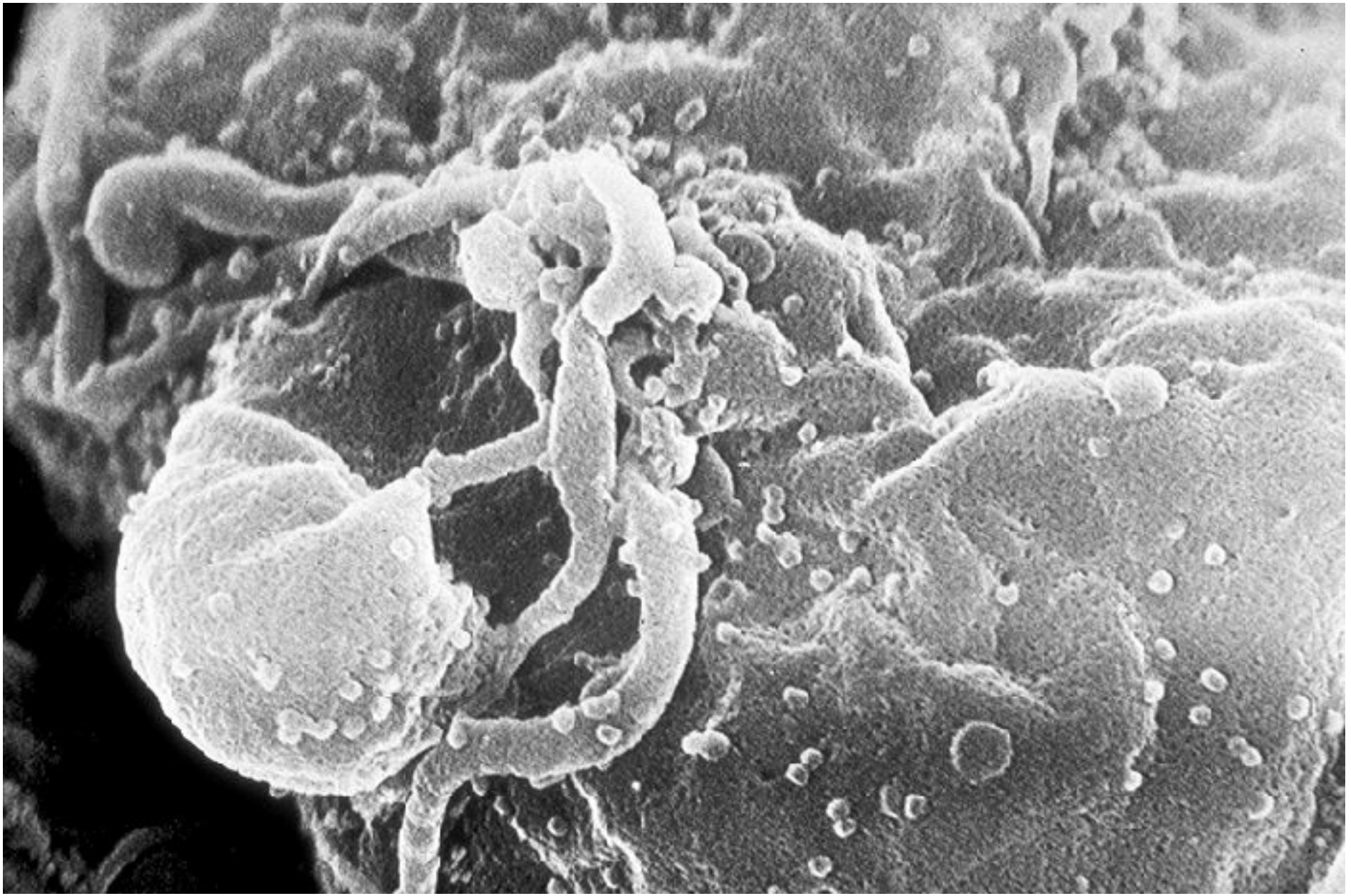
There are two potential mechanisms by which dual infection can occur: coinfection, in which a host is infected by two distinct HIV strains at or around the same time, and superinfection, in which there is sequential infection of a patient by two different viral strains.

Distinguishing between these possibilities has major implications for vaccine development.

HIV interaction with CD4+ cell



During the HIV infection several quasispecies of the virus arise, which are able to use different coreceptors, in particular the CCR5 and CXCR4 coreceptors (R5 and X4 phenotypes, respectively). The R5-X4 switch in coreceptor usage has been correlated with a faster progression of the disease to the AIDS phase.

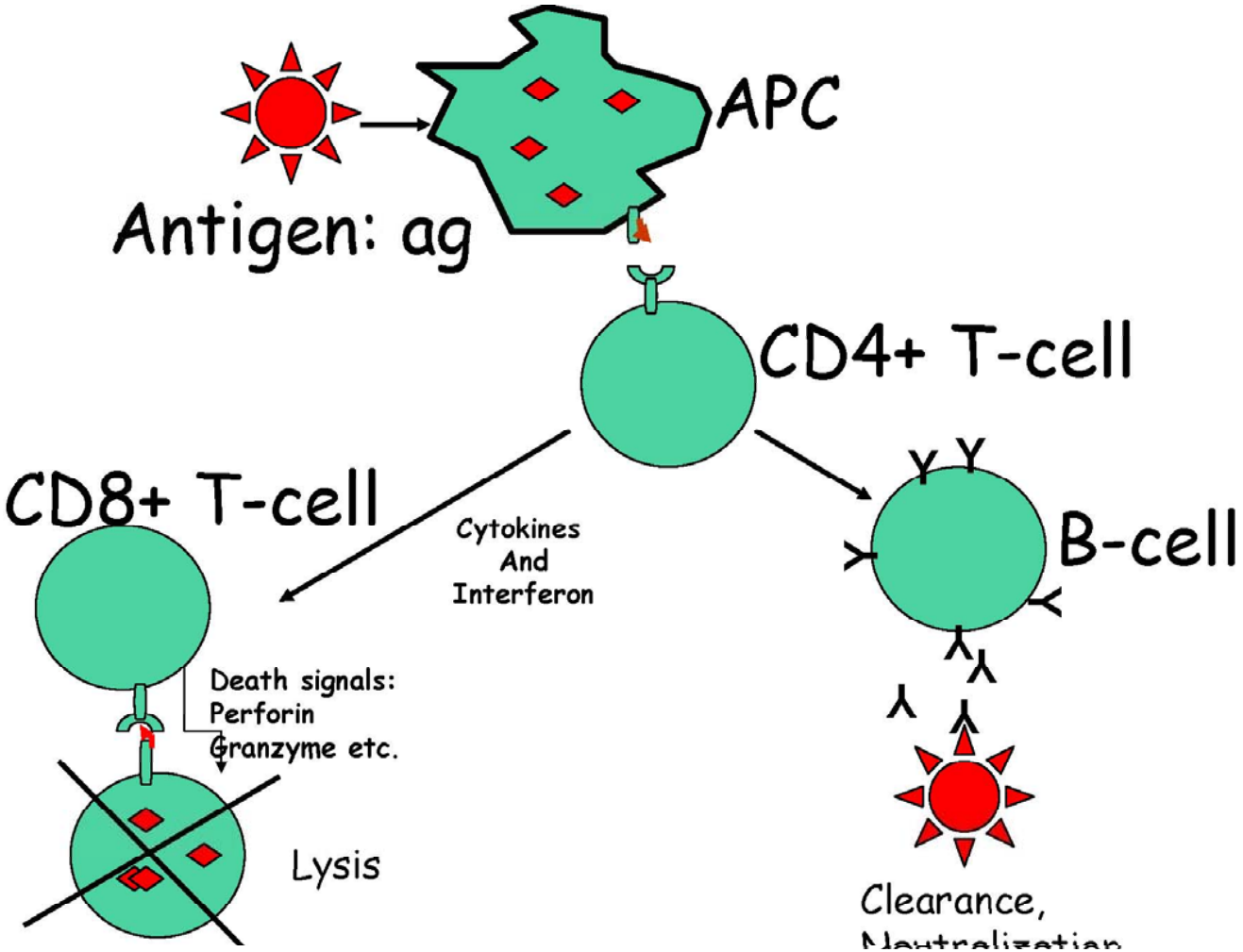


Modeling?

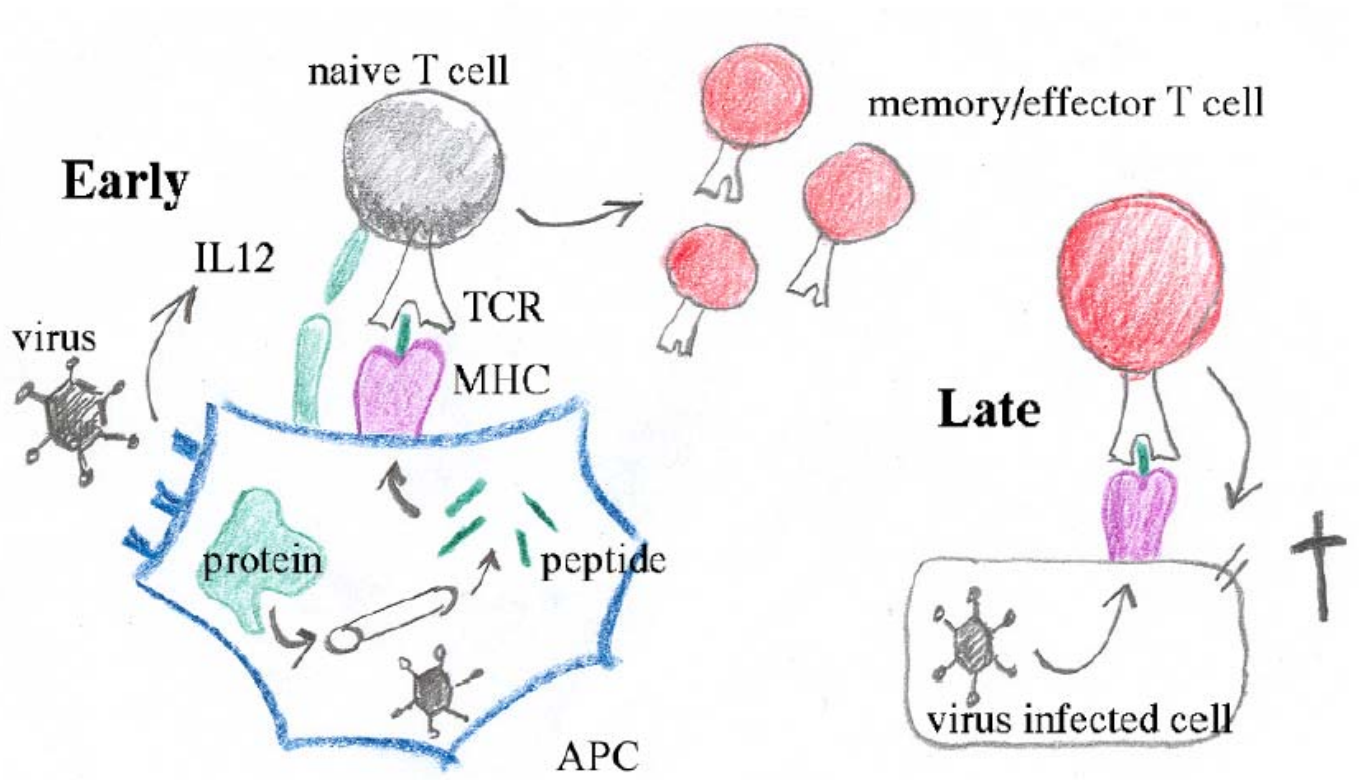
•**Discrete and stochastic -** Finest scale of representation for well stirred molecules. Exact description due to Gillespie and others

• **Continuous and stochastic -** The Langevin regime. Valid under certain conditions. Described by Stochastic Differential Equations (SDE).

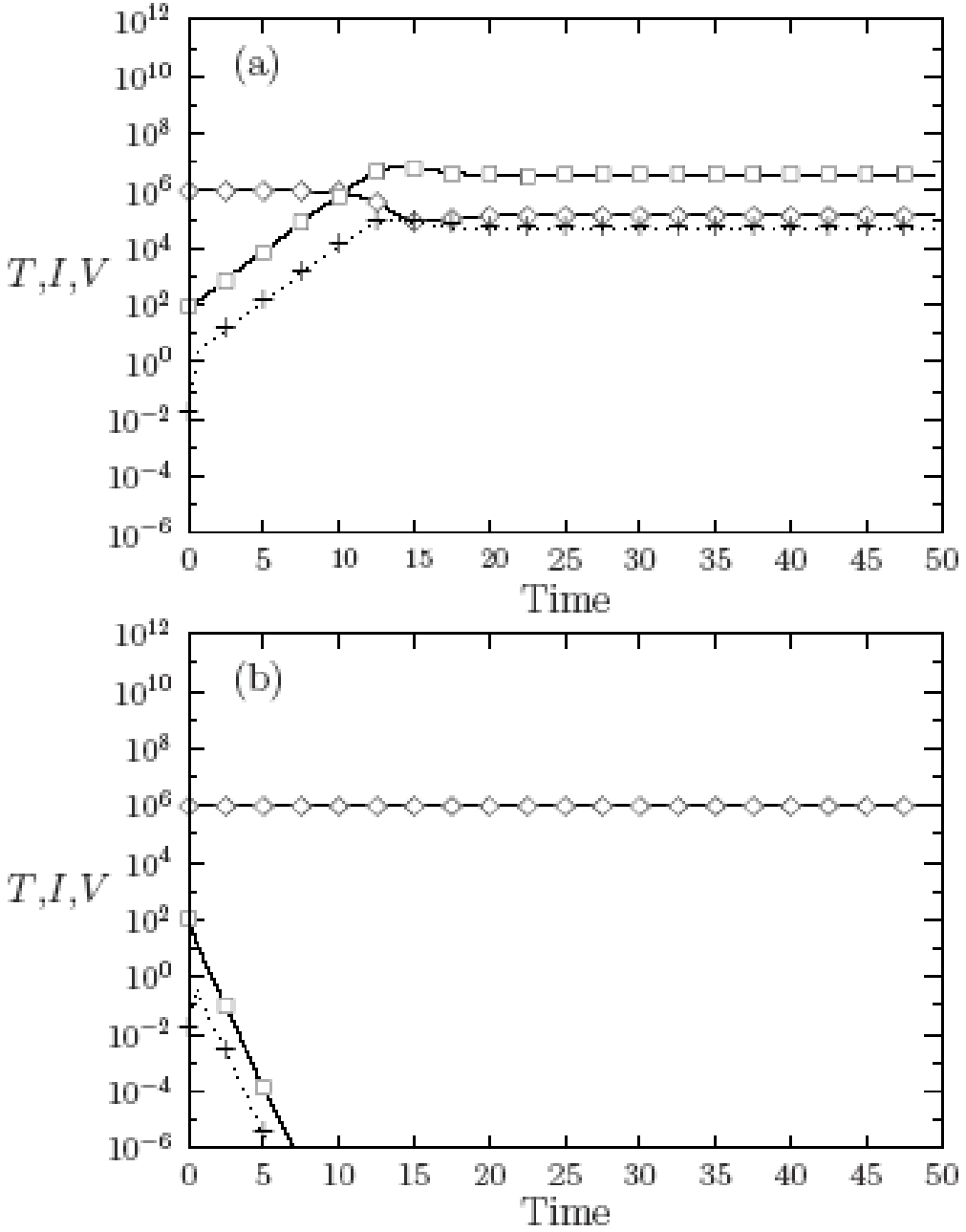
• **Continuous and deterministic -** The rate equations. Described by ordinary differential equations (ODE). Valid under further assumptions.



$$\dot{T} = \left(\lambda + \gamma^{(T)} I\, T\right)(1 - T/K) - (\delta_T + \beta\, V)\, T,$$
$$\dot{I} = \beta\, V\, T - \left(\delta_I + \gamma^{(I)}\, T\right) I,$$
$$\dot{V} = \pi\, I - \left(c + \gamma^{(V)}\, T\right) V.$$



$$\dot{T} = \left(\lambda + \gamma^{(T)} I\, T\right)(1 - T/K) - (\delta_T + \beta\, V)\, T,$$
$$\dot{I} = \beta\, V\, T - \left(\delta_I + \gamma^{(I)}\, T\right) I,$$
$$\dot{V} = \pi\, I - \left(c + \gamma^{(V)}\, T\right) V.$$



Time evolution (in days) of the infection using single-species model.
Diamonds represent uninfected T cells, squares represent viruses and plus signs the infected T cells.

Plot (a) illustrates a scenario leading to a chronic infection.

In the presence of a further B cell response (plot (b)), the infection is defeated by the immune system.

Quasi species (multi strain) model

We assume that each viral strain is characterized by just one epitope i.

We make use of coupled differential equations, one for each viral quasispecies and T cell.

Although this mean-field approach disregards the effect of fluctuations and genetic drift in quasispecies abundances, it is useful for understanding the coarse-grain features of the behavior of the interplay between HIV and the immune system.

$$\dot{T}_i = \left(\lambda_i + \sum_k \gamma_{ik}^{(T)} I_k T_i\right)\left(1 - \frac{1}{K}\sum_i T_i\right) - \left(\delta_T + \sum_k \beta_k V_k\right) T_i,$$

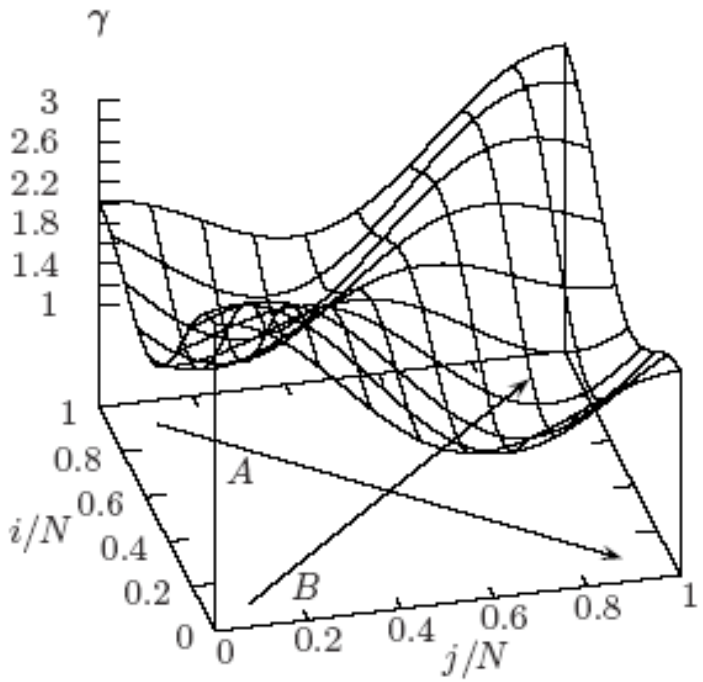
$$\dot{I}_k = \left(\sum_{k'} \mu_{kk'} \beta_{k'} V_{k'}\right)\left(\sum_i T_i\right) - \left(\delta_I + \sum_i \gamma_{ki}^{(I)} T_i\right) I_k,$$

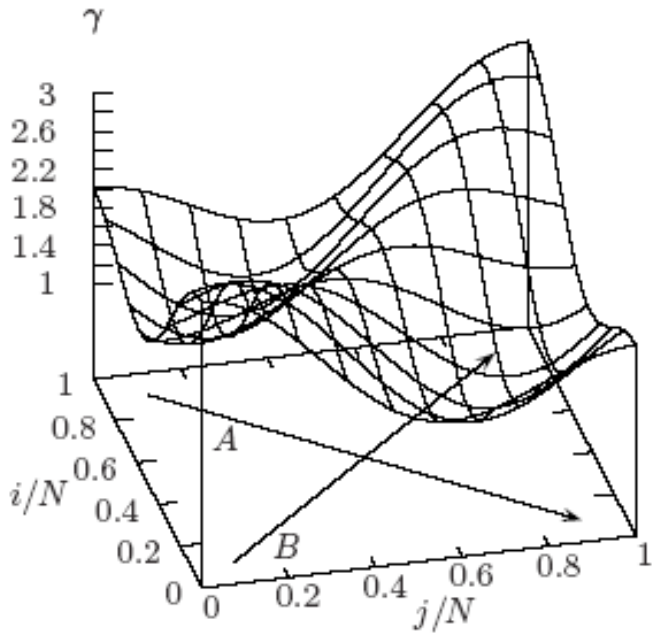
$$\dot{V}_k = \pi I_k - \left(c + \sum_i \gamma_{ki}^{(V)} T_i\right) V_k.$$

The model considers the following cell types: T-helper (CD4+) cells responding to virus strain i, (T_i); T cells (any strain) infected by virus strain k, (I_k); abundance of viral strain k, (V_k). This means that viral strain k are identified by just one epitope, which is then displayed on the surface of the T cell of class k, and that a T cell of class i can be activated at least by one CD4+ T cell carrying the epitope k, which is specific of the viral strain k. The indices i (k) range from 1 to N_i (N_k), and in the following we have used $N_i = N_k = N$.

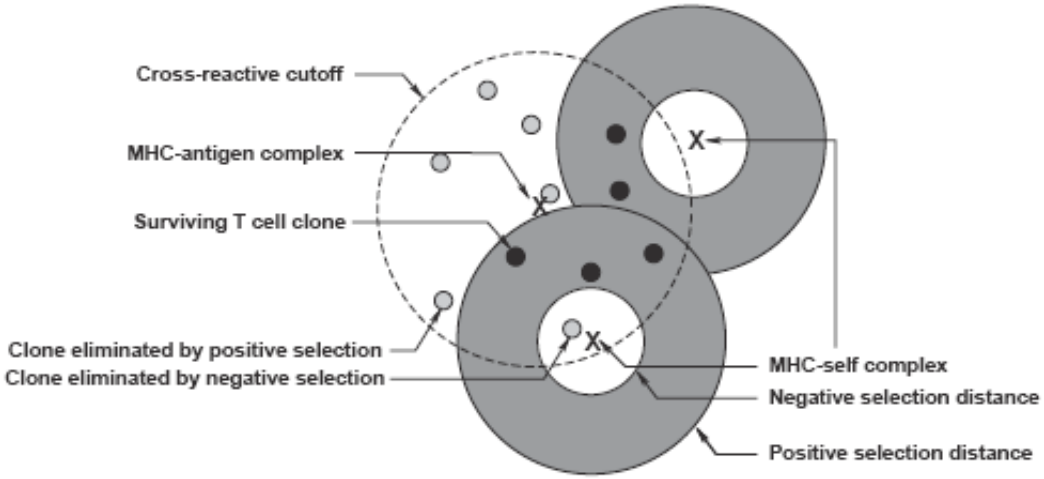
The gamma parameters incorporate the geometry and strength of cell-cell and virus-cell interactions

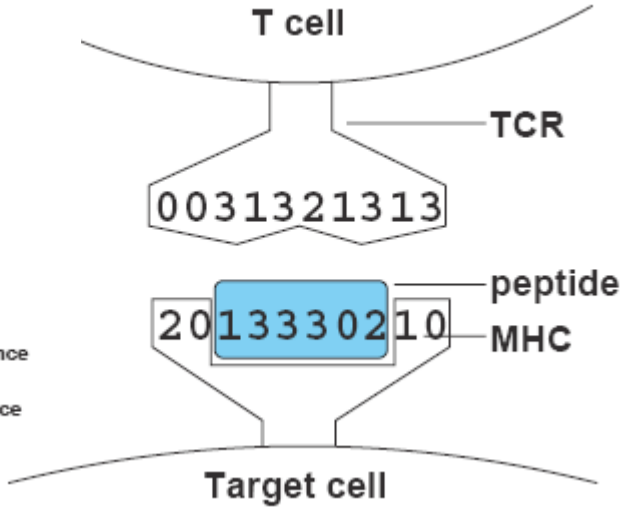
$$\gamma_{ij} = A\left\{\frac{1}{2}\left[1+\cos\left(\frac{2\pi}{N}(i-j)\right)\right]\right\}^{\varepsilon_A} + B\left\{\frac{1}{2}\left[1+\cos\left(\frac{\pi}{N}(i+j)\right)\right]\right\}^{\varepsilon_B}$$

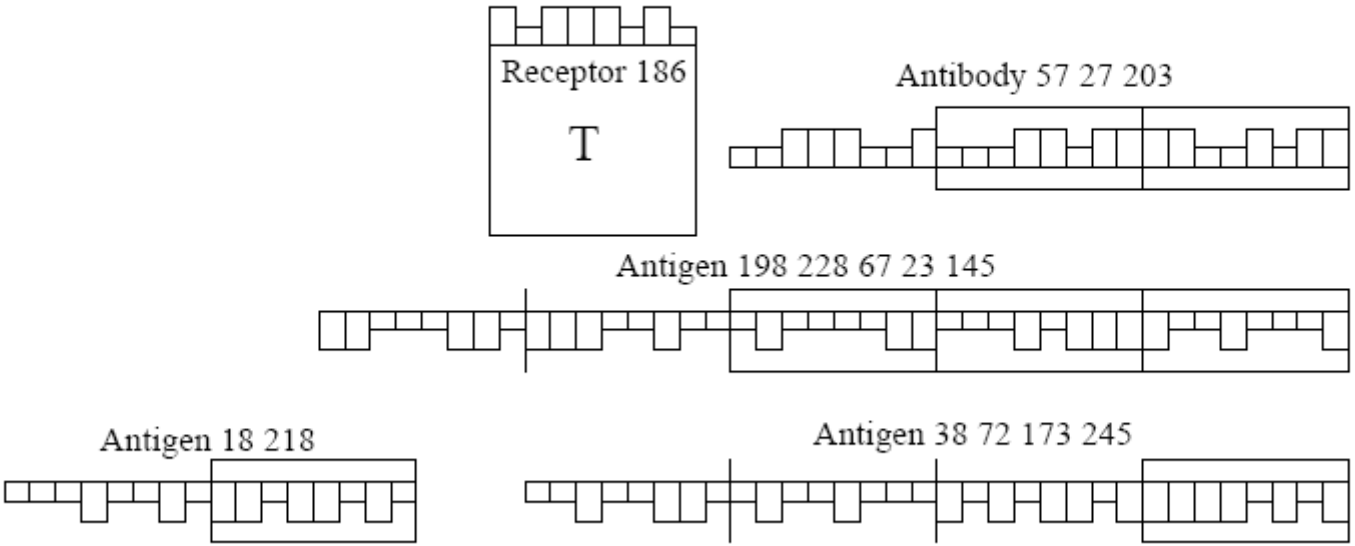




The parameters A and ε_A (B and ε_B) control the shape of the γ matrices in the direction transverse (parallel) to the main diagonal. The exponents ε_A and ε_B controls the smoothness of the variation along the corresponding direction. For ε small the variation of the function is very small, while for $\varepsilon \to \infty$ the corresponding function is δ-shape. So, a diagonal γ matrix is obtained by setting $B = 0$ and ε_A large.

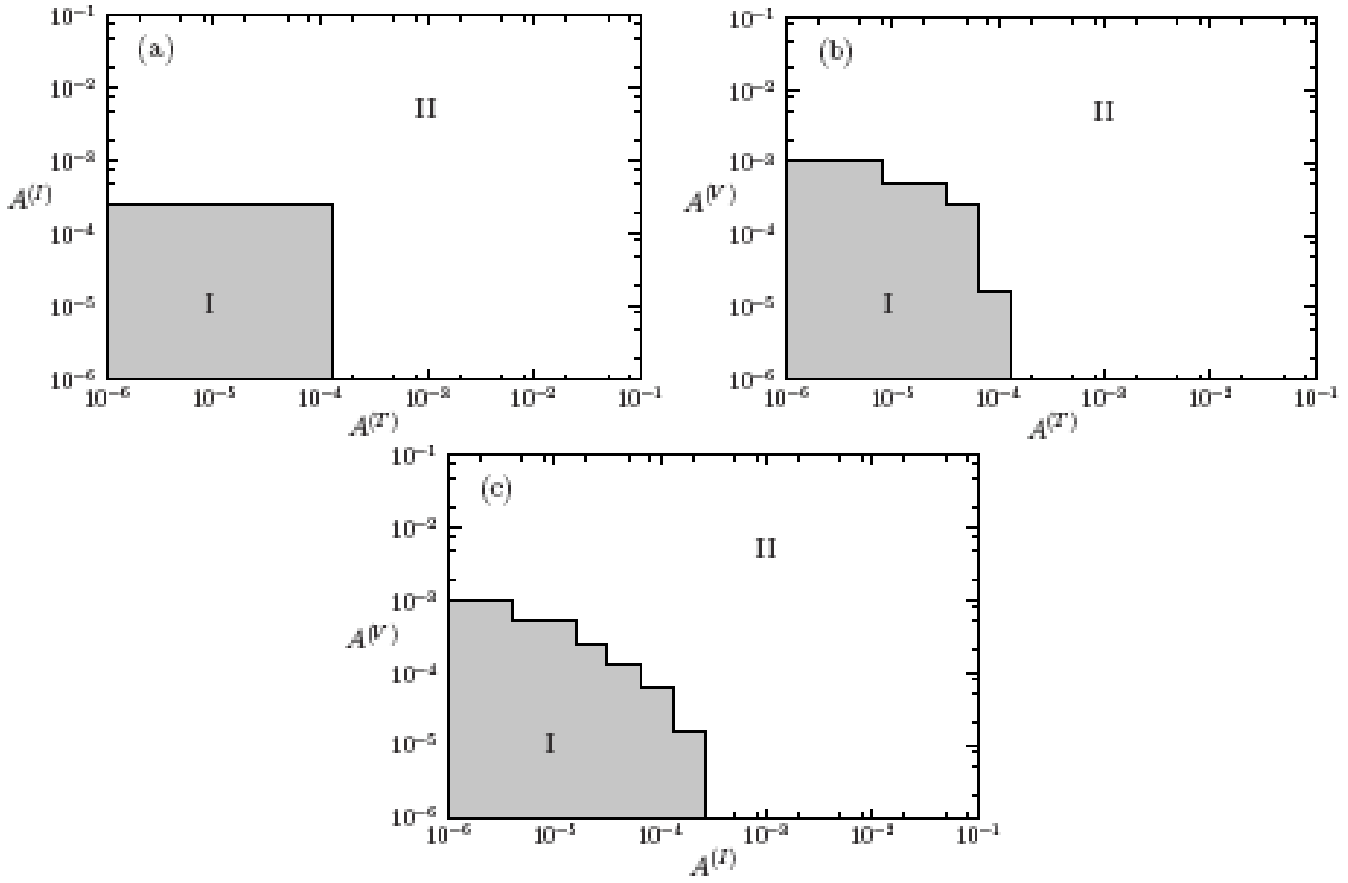

Cross-reactive cutoff
MHC-antigen complex
Surviving T cell clone
Clone eliminated by positive selection
Clone eliminated by negative selection
MHC-self complex
Negative selection distance
Positive selection distance


T cell
TCR
0031321313
peptide
20133302 10
MHC
Target cell

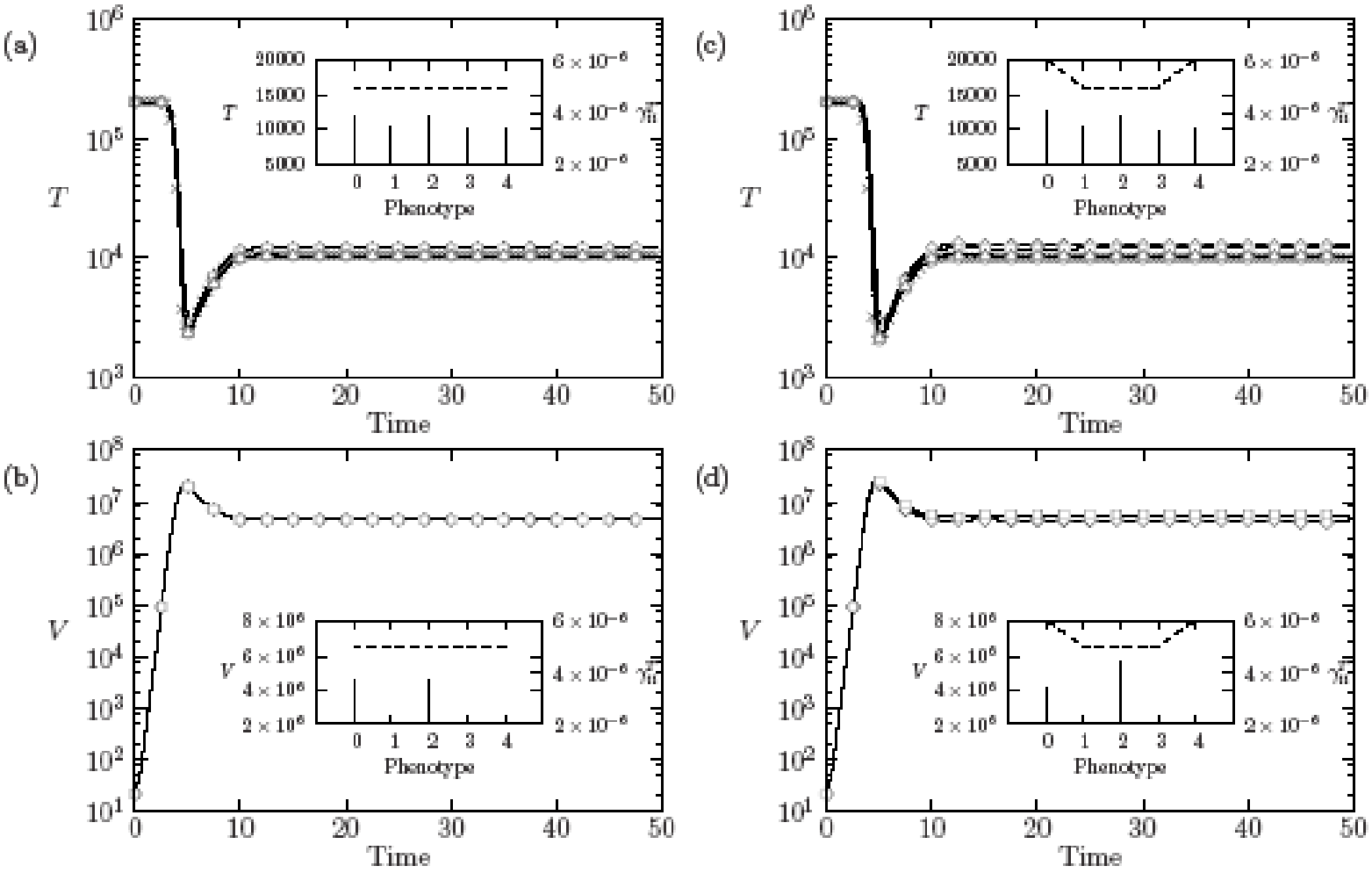

Receptor 186
T
Antibody 57 27 203
Antigen 198 228 67 23 145
Antigen 18 218
Antigen 38 72 173 245

A summary of model parameters

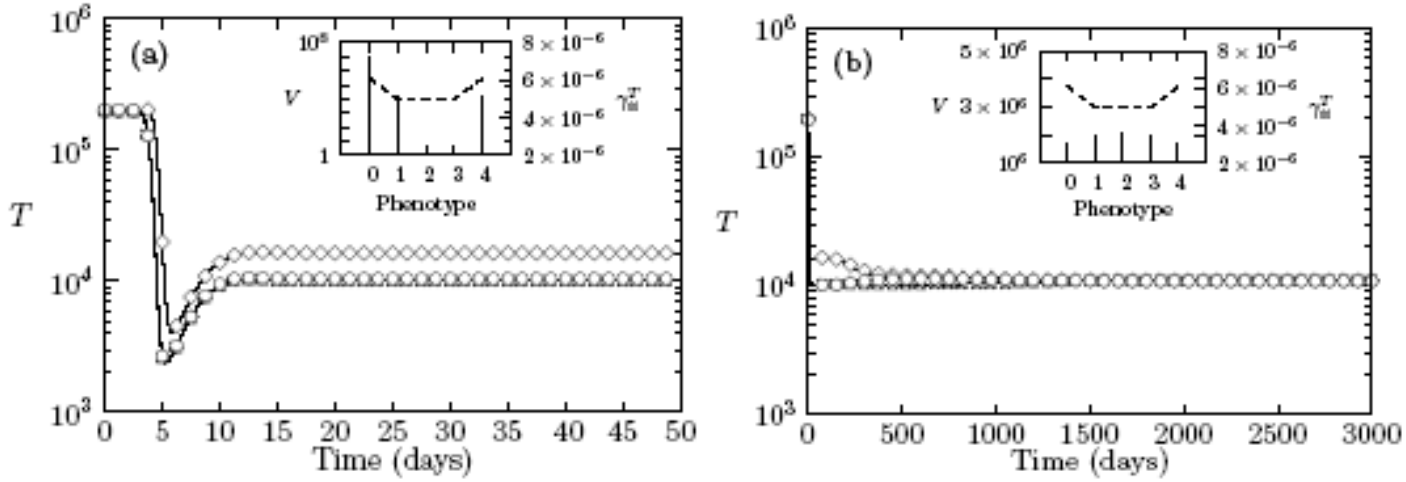
Attribute	Value	Reference
Time step (Δt)	10 min	
Naïve cell clone size	10 cells	Casrouge et al. (2000)
Maximum T cell recruitment rate (γ)	$1\,\text{day}^{-1}$	
Delay before a stimulated naïve cell becomes an effector (τ_n)	19 h	Oehen and Brduscha-Riem (1998); Gett and Ho Veiga-Fernandes et al. (2000); van Stipdonk et
Delay before a stimulated memory cell becomes an effector (τ_m)	1 h	Bachmann et al. (1999); Barber et al. (2003)
Naïve-derived active CTL death rate (δ_E)	$0.6\,\text{day}^{-1}$	Veiga-Fernandes et al. (2000)
Memory-derived active CTL death rate (δ_{E_m})	$0.4\,\text{day}^{-1}$	Veiga-Fernandes et al. (2000)
Time in B phase for CTL	5 h	van Stipdonk et al. (2001)
Average CTL cell cycle time	6 h	van Stipdonk et al. (2001)
Infected cell clearance rate (k^c)	$12\,\text{day}^{-1}$	Barchet et al. (2000)
Pre-selection repertoire size	2.5×10^8	
MHC string length	4 digits	Detours et al. (1999)
Peptide string length	6 digits	Detours et al. (1999)
Susceptible cell population (T)	10^6 cells	
Susceptible cell production rate (λ)	10^5 cells/day	
Susceptible cell death rate (δ_T)	$0.1\,\text{day}^{-1}$	
Virus infection rate (β)	2×10^{-7}	
Virus production rate (π)	$100\,\text{day}^{-1}$	
Virus clearance rate (c)	$2.3\,\text{day}^{-1}$	Lehmann-Grube (1988)
Infected cell death rate (δ_I)	$0.8\,\text{day}^{-1}$	

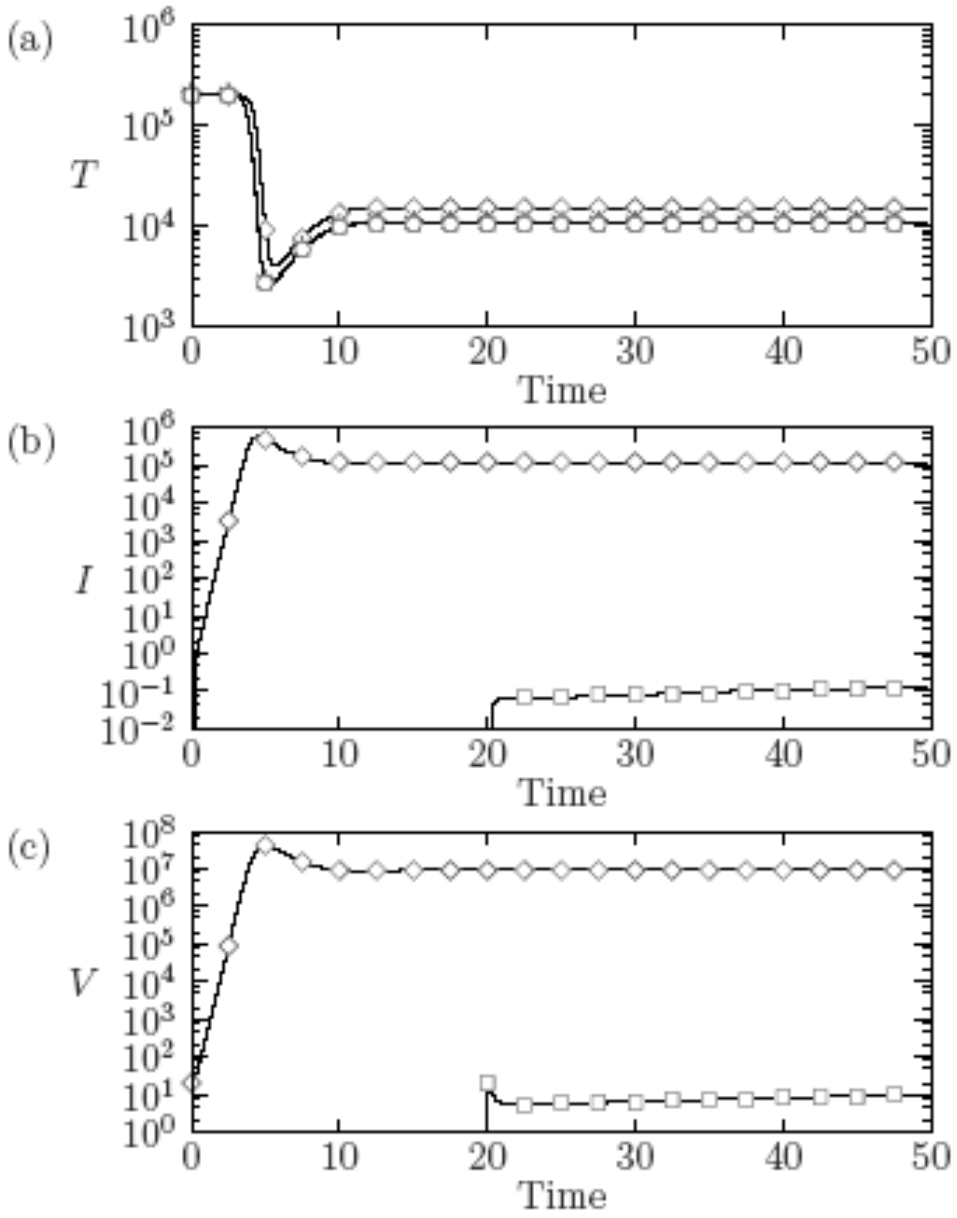


Phase space analysis of As parameters of the γ matrices. In plot (a) we fix $A^{(V)}$=10^{-6} and we show $A^{(I)}$ vs. $A^{(T)}$. In plot (b) we fix $A^{(I)}$=10^{-6} and we show $A^{(V)}$ vs. $A^{(T)}$. In plot (c) we fix $A^{(T)}$=10^{-6} and we show $A^{(V)}$ vs. $A^{(I)}$. The other parameters are: number of quasi species $N = 25$, $\varepsilon_A = 10^3$, $B = 0$ and $\mu = 0$. Region I corresponds to the coexistence of immune system and viruses (chronic infection) while region II corresponds to the defeat of viral infection.

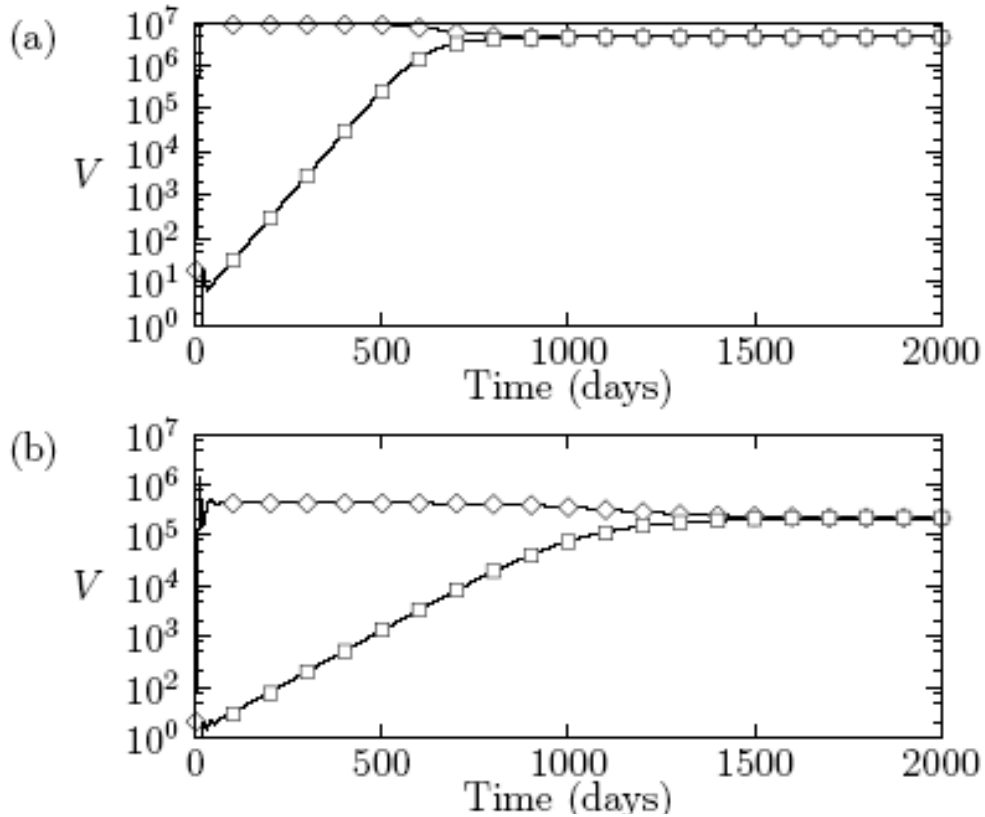


Time evolution of T (plots (a) and (c)) and V (plots (b) and (d)) populations after an inoculum with two strains (phenotypes 0 and 2), $\mu = 0$, $A = 5 \cdot 10^{-6}$, $\varepsilon_A = 10$, $N = 5$. Plots (a) and (b): $B = 0$; Plots (c) and (d): $B^{(T)} = 3 \cdot 10^{-6}$, $\varepsilon_B^{(T)} = 10$. In the inset the asymptotic distribution of T and V populations and the shape of the diagonal of $\gamma^{(T)}$ are shown.


(a)
10^6
10^5
10^4
10^3
T
0 5 10 15 20 25 30 35 40 45 50
Time (days)
10^5
V
1
0 1 2 3 4
Phenotype
8×10^{-6}
6×10^{-6}
4×10^{-6}
2×10^{-6}
γ_{ii}^T
(b)
10^6
10^5
10^4
10^3
T
0 500 1000 1500 2000 2500 3000
Time (days)
5×10^6
V 3×10^6
10^6
0 1 2 3 4
Phenotype
8×10^{-6}
6×10^{-6}
4×10^{-6}
2×10^{-6}
γ_{ii}^T



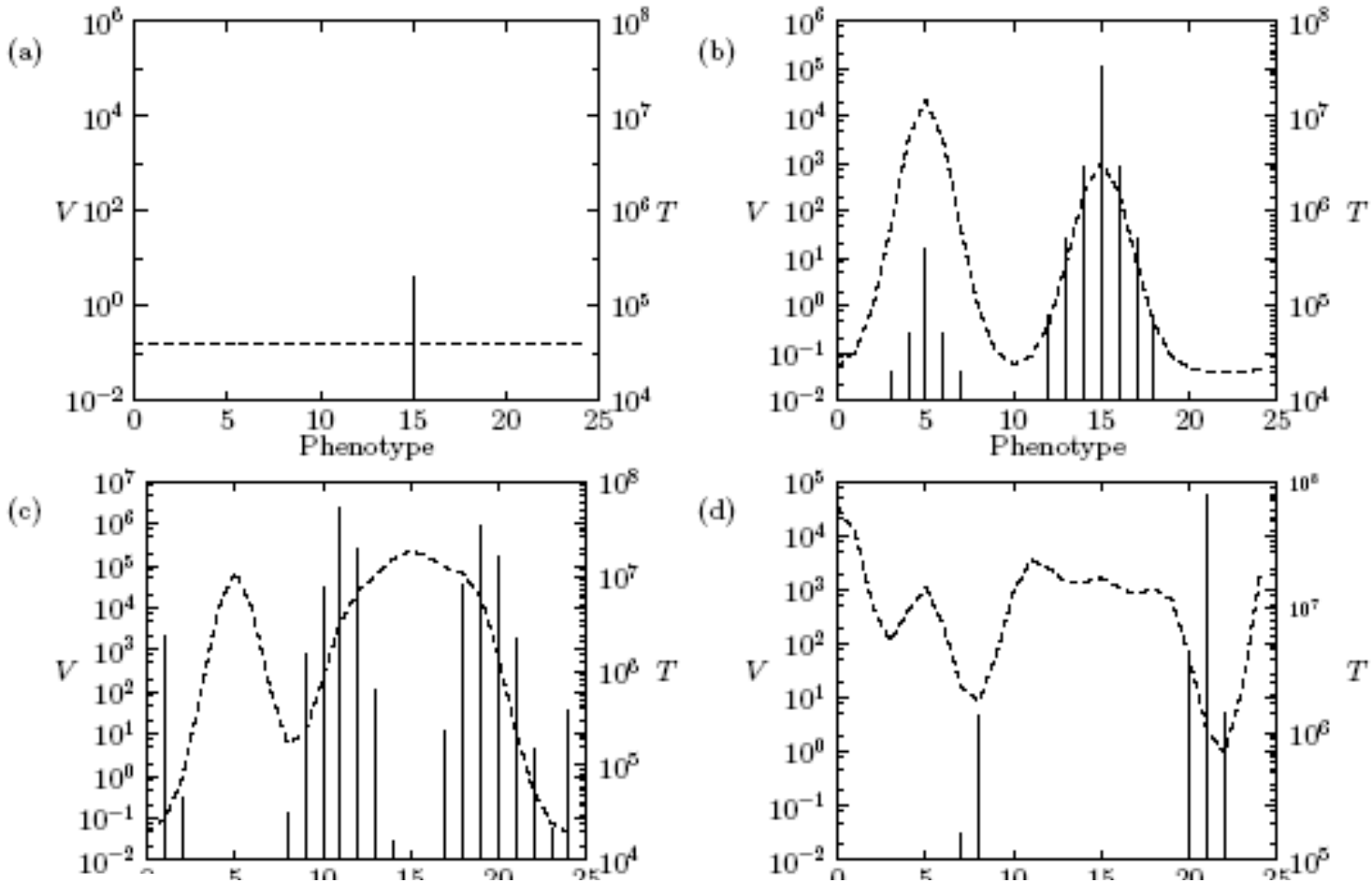
Short time behavior of uninfected T cells (a), infected T cells (b) and viruses (c) under superinfection.



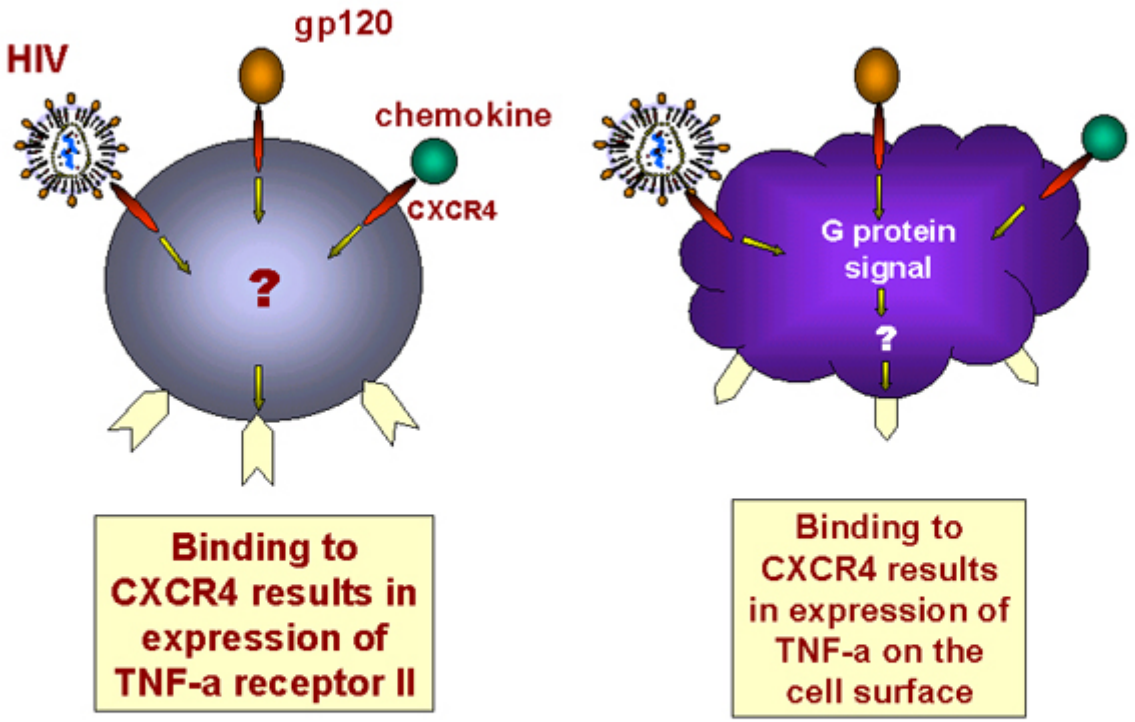
Viral counts, V, during a superinfection scenario. We set N=5 and no mutation is considered, thus μ=0. (a) A slow mounting of the second viral infection (□), having time scale of several months, is observed. In (b) a compromised immune system is considered. The time for the second strain to reach the same abundance of the first-infecting strain (◇) is greater than in (a).

We report the following finding:

after superinfection or coinfection, quasispecies dynamics has time scales of several months and becomes even slower at low number of CD4+ T cells.

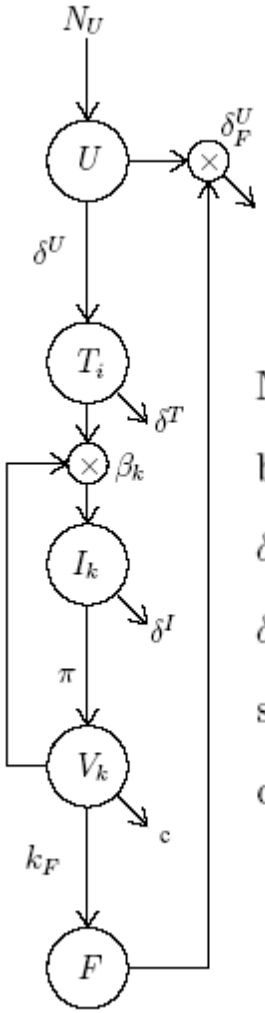


Speciation of virus quasispecies and uninfected T cells dynamics after competitive superinfection at four different times: $t = 0$ (a), $t = 4.5$ (b), $t = 5.25$ (c) and $t = 5.75$ (d). Virus strain 15 is present at time $t = 0$, while strain 5 is inoculated at time $t = 1$. Mutation rate $\mu = 10^{-4}$ and non-uniform interaction strength The dashed line represents the abundances of T cells targeting each viral phenotype, represented as vertical stems.

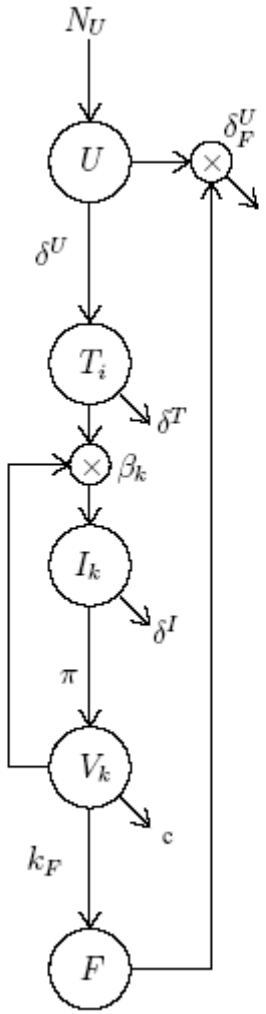


Apoptosis of T cells

Schematic description of the model for the switching from R5 to X4 viral phenotype.



Naïve T-cells, U, are generated at constant rate N_U and removed at rate δ^U. They give birth to differentiated, uninfected T-cells, T. These in turn are removed at constant rate δ^T and become infected as they interact with the virus. Infected T-cells, I, die at rate δ^I and contribute to the budding of viral particles, V, that are cleared out at rate c. As soon as the X4 phenotype arise, the production of the TNF starts, proportional to the X4 concentration and contribute to the clearance of naïve T-cells, via the δ_F^U parameter.



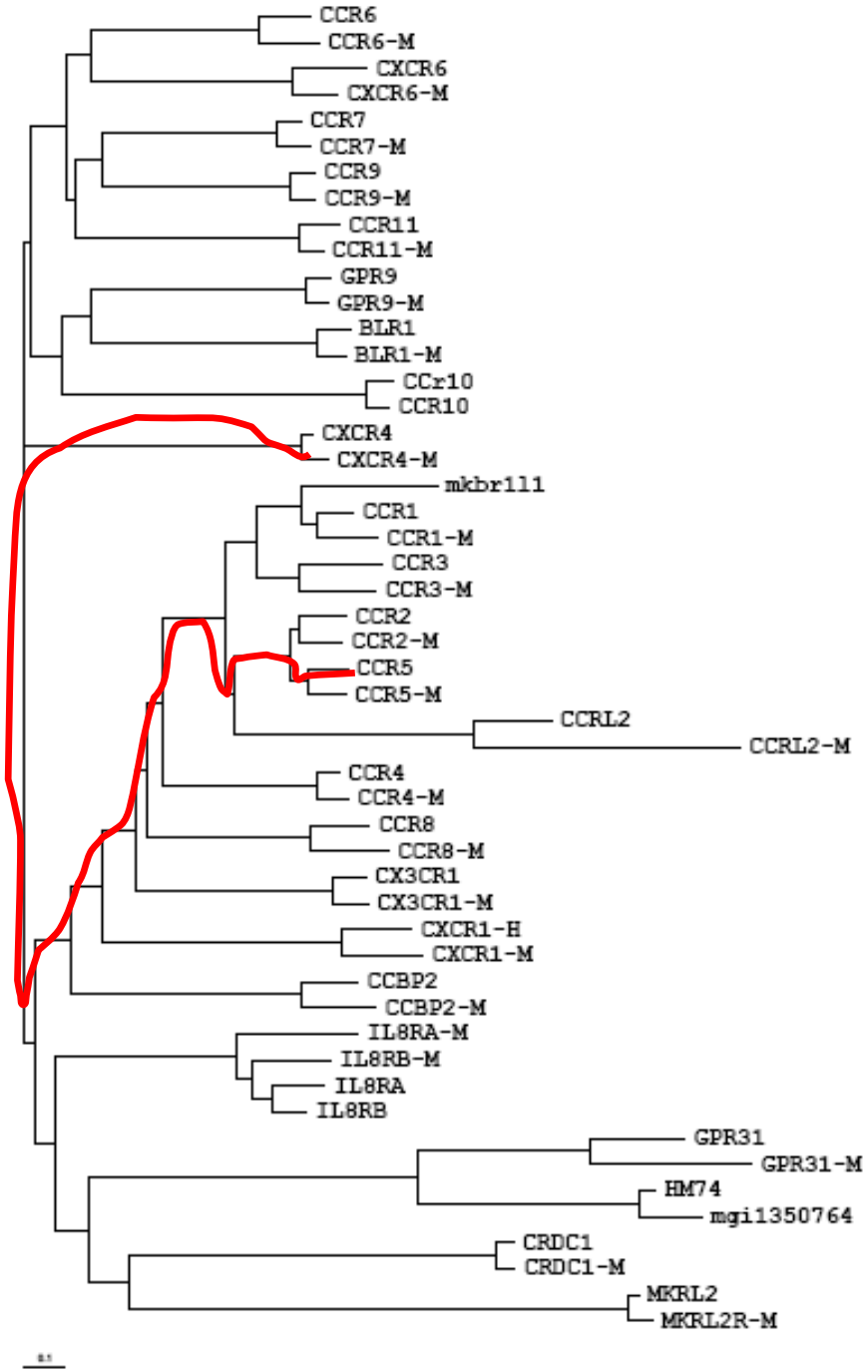
$$\frac{\mathrm{d}U}{\mathrm{d}t} = N_U - \delta^U U - \delta_F^U U F$$

$$\frac{\mathrm{d}T_i}{\mathrm{d}t} = \delta^U U - \left(\sum_k \beta_k V_k\right) T_i - \delta^T T_i$$

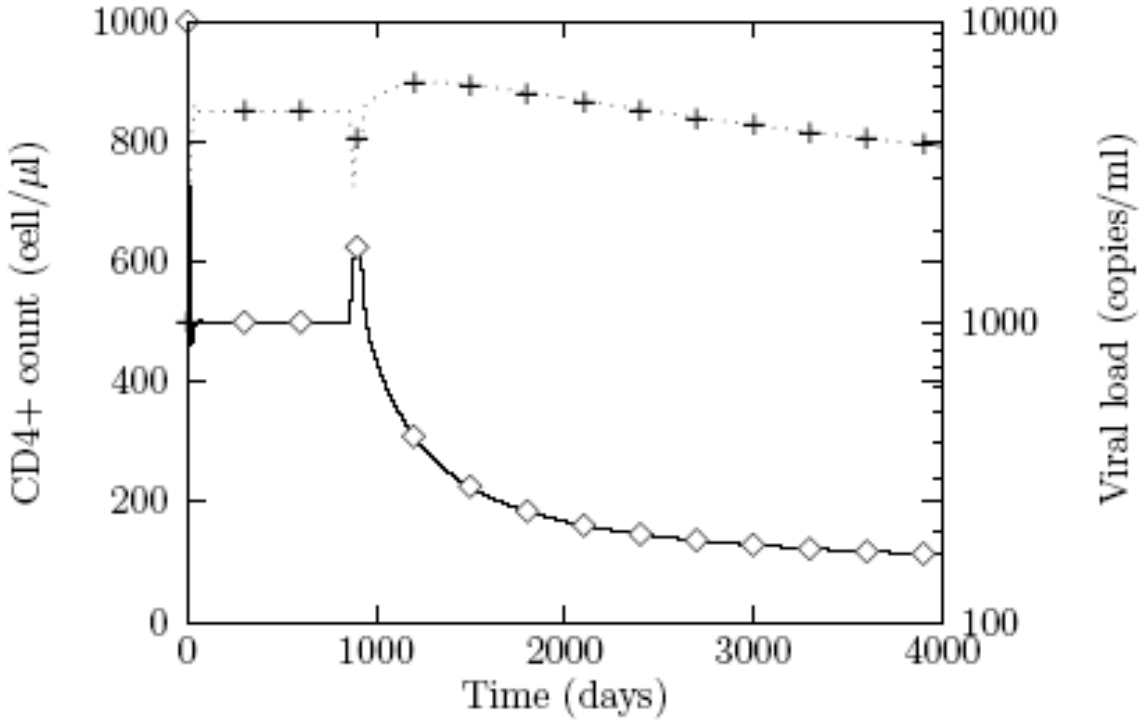
$$\frac{\mathrm{d}I_k}{\mathrm{d}t} = \left(\sum_{k'} \mu_{kk'} \beta_{k'} V_{k'}\right) \left(\sum_i T_i\right) - \delta^I I$$

$$\frac{\mathrm{d}V_k}{\mathrm{d}t} = \pi I_k - c V_k$$

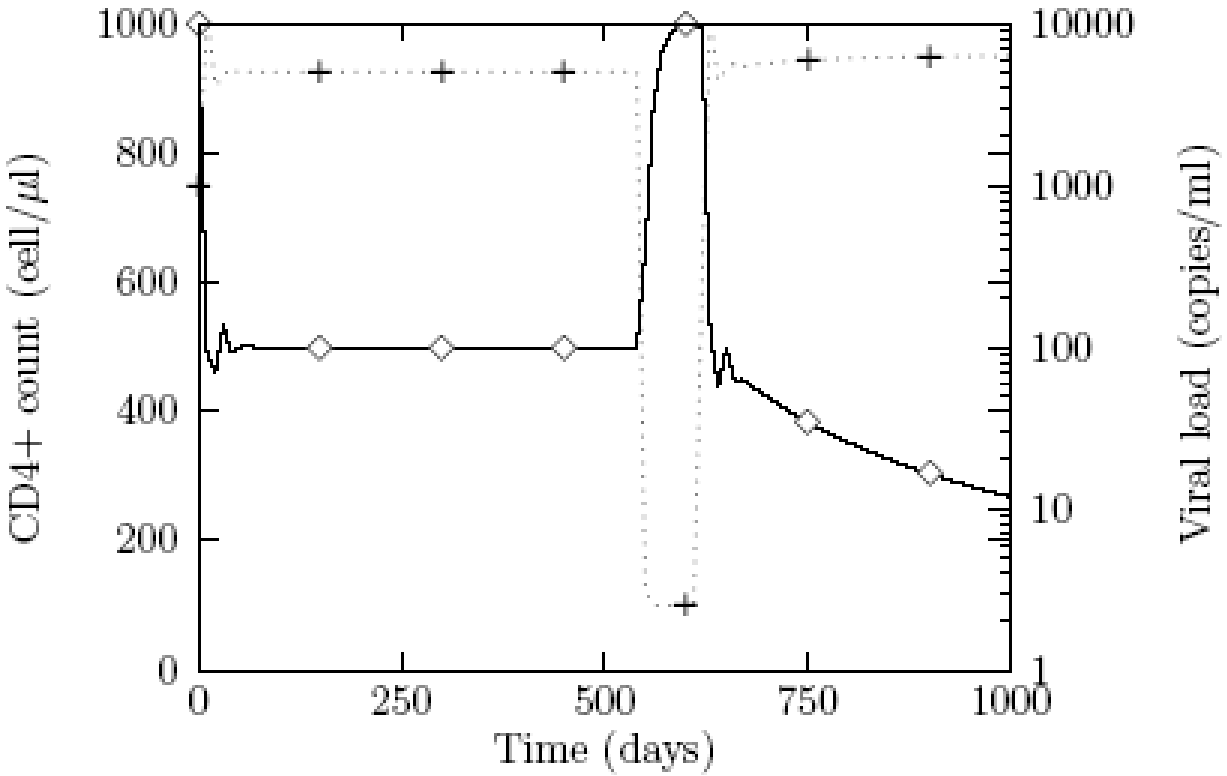
$$\frac{\mathrm{d}F}{\mathrm{d}t} = k_F \sum_{k \in X4} V_k$$



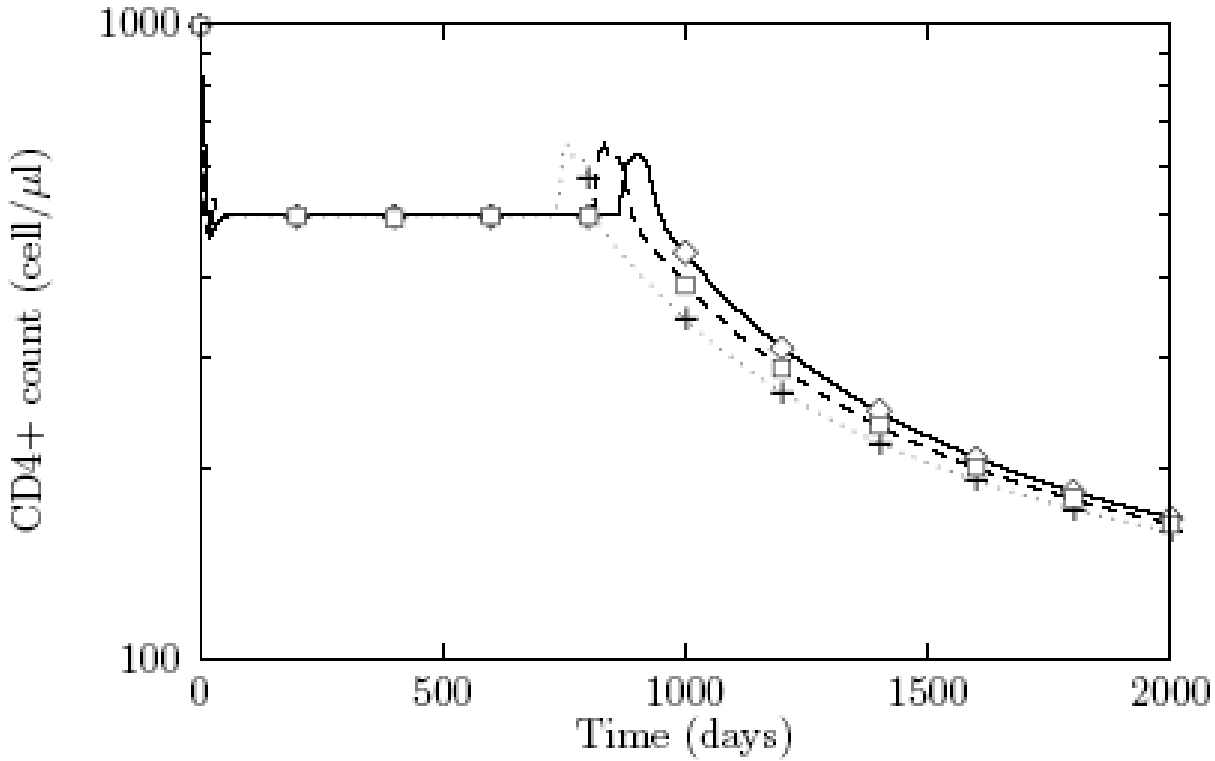
Phylogenetic inference of chemokine receptors suggests that viral mutational pathway may generate R5 variants able to interact with chemokine receptors different from CXCR4.



Time evolution of the concentrations of uninfected T-cells ($\diamond$) and viruses (+), during R5 to X4 switch, occurring at time $t \approx 900$. The time of appearance of the X4 strains depends on the mutation rate and on the phenotypic distance between R5 and X4 viruses. After the appearance of the X4 phenotype a continuous slow decline in CD4+ T-cells level leads to AIDS phase (CD4 counts below 200cells/ml). We set $\mu \approx 0.001$ and $d_P \approx 5$.



The efficacy of HAART therapy may be disrupted by a sudden interruption in drugs treatment. If time has passed for mutations to populate the R5 strains closer to the X4 phenotypes, an earlier appearance of X4 strains may occur. Uninfected T-cells (◇) and viruses (+).

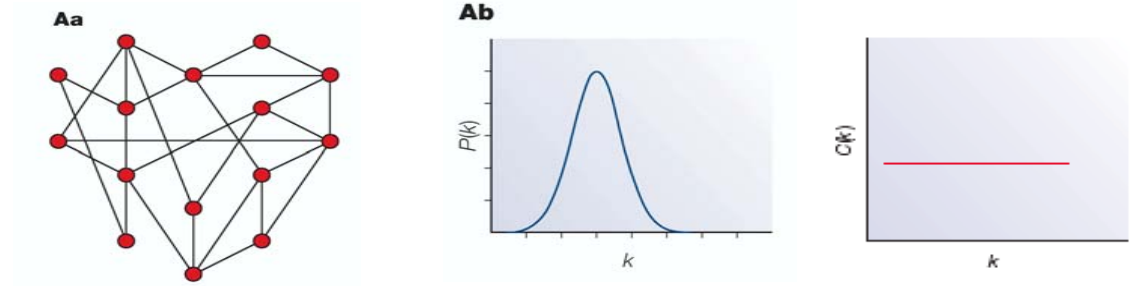


CD4+ T-cells concentration during HIV-1 super-infection by a R5 viral strain. Different signs represent: evolution without superinfection, (◇); superinfection occurring at time t=100 and 400, (+) and (□), respectively. For a superinfection event occurring after the R5 to X4 switching the dynamics is qualitatively the same as for a single infection, (◇). If the second delayed infection occurs before the R5 to X4 switching, the time of appearance of X4 viruses may be shorter, when the superinfecting strain is closer to the X4 phenotypes, (+,□).

We report the following finding:

The decays of the number of CD4+ and CD8+ T cells during AIDS late stage can be described taking into account the X4 related Tumor Necrosis Factor dynamics (TRIAL).

A. Random Networks [Erdos and Rényi (1959, 1960)]



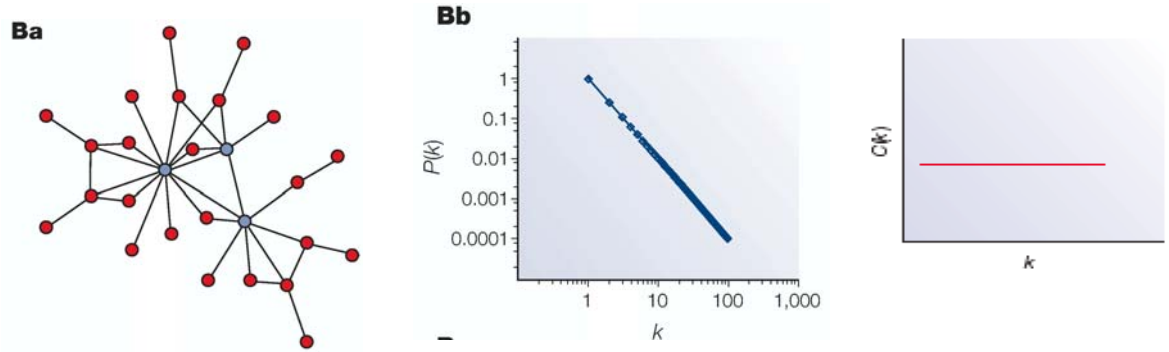
$$P(k) = \frac{e^{-\bar{k}}\bar{k}^{k}}{k!}$$

Mean path length ~ ln(k)

Phase transition:

Connected if: $p \geq \ln(k)/k$

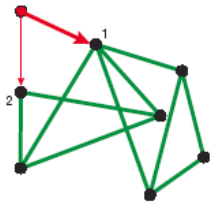
B. Scale Free [Price,1965 & Barabasi,1999]



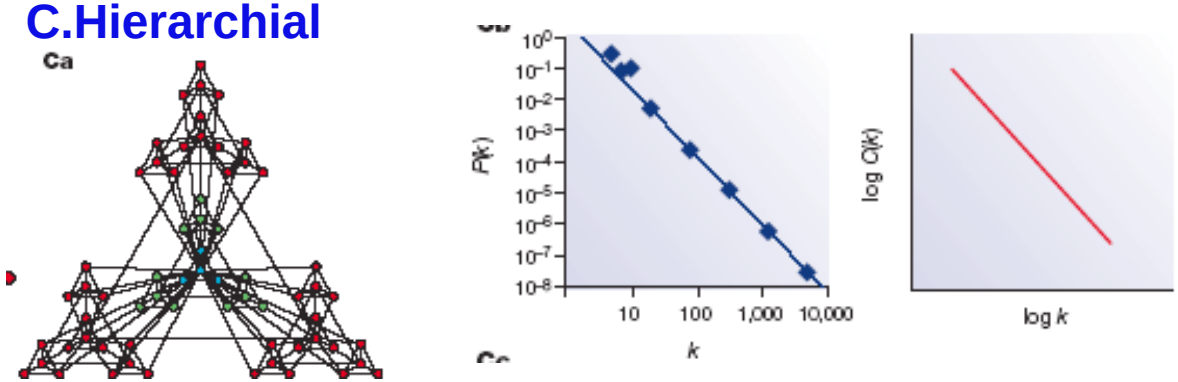
$$P(k) \sim k^{-\gamma},\ k >> 1,\ \ 2 < \gamma$$

Mean path length ~ lnln(k)

Preferential attachment. Add proportionally to connectedness



C.Hierarchial



Copy smaller graphs and let them keep their connections.

We start from the infection probability given S "independent" infected neighbors, that is:

$$P_I^S = 1 - (1 - p)^S$$

where p represents the probability of infection per contact. The expression of p is given by:

$$p = p_0 - p_P.$$

The first term, p_0, accounts for the intrinsic infection probability of the virus and the second,

$$p_P = H + J,$$

represents the perception of an individual of the risk of becoming infected. This term is composed by two terms, one representing the external infuences on the behavior (e.g. TV, radio etc), H, and the second mimicking the strategy of behaving of the single individual, depending mainly on the status of the neighbors, J.

We represent the spatial structure of the population with a graph. Each node represents a person, whose status is 1 if the person is infected, 0 otherwise. We describe the network of contacts by the M matrix, whose elements, M_{ij} , describe the link between the i-th and the j-th node. With this schematization, the equation becomes:

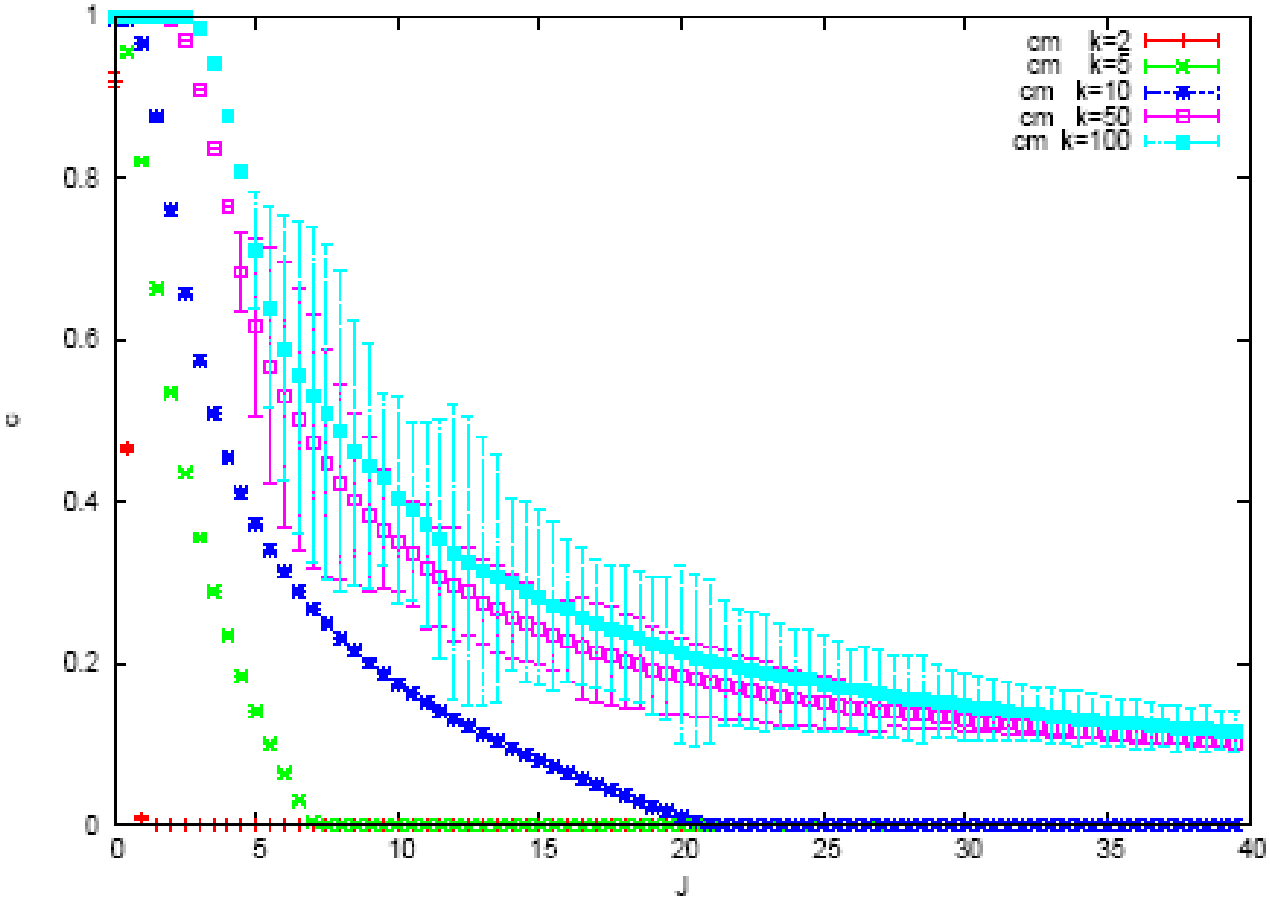
$$P_I = 1 - \Pi_j (1-p)^{M_{ij}\sigma_j}$$

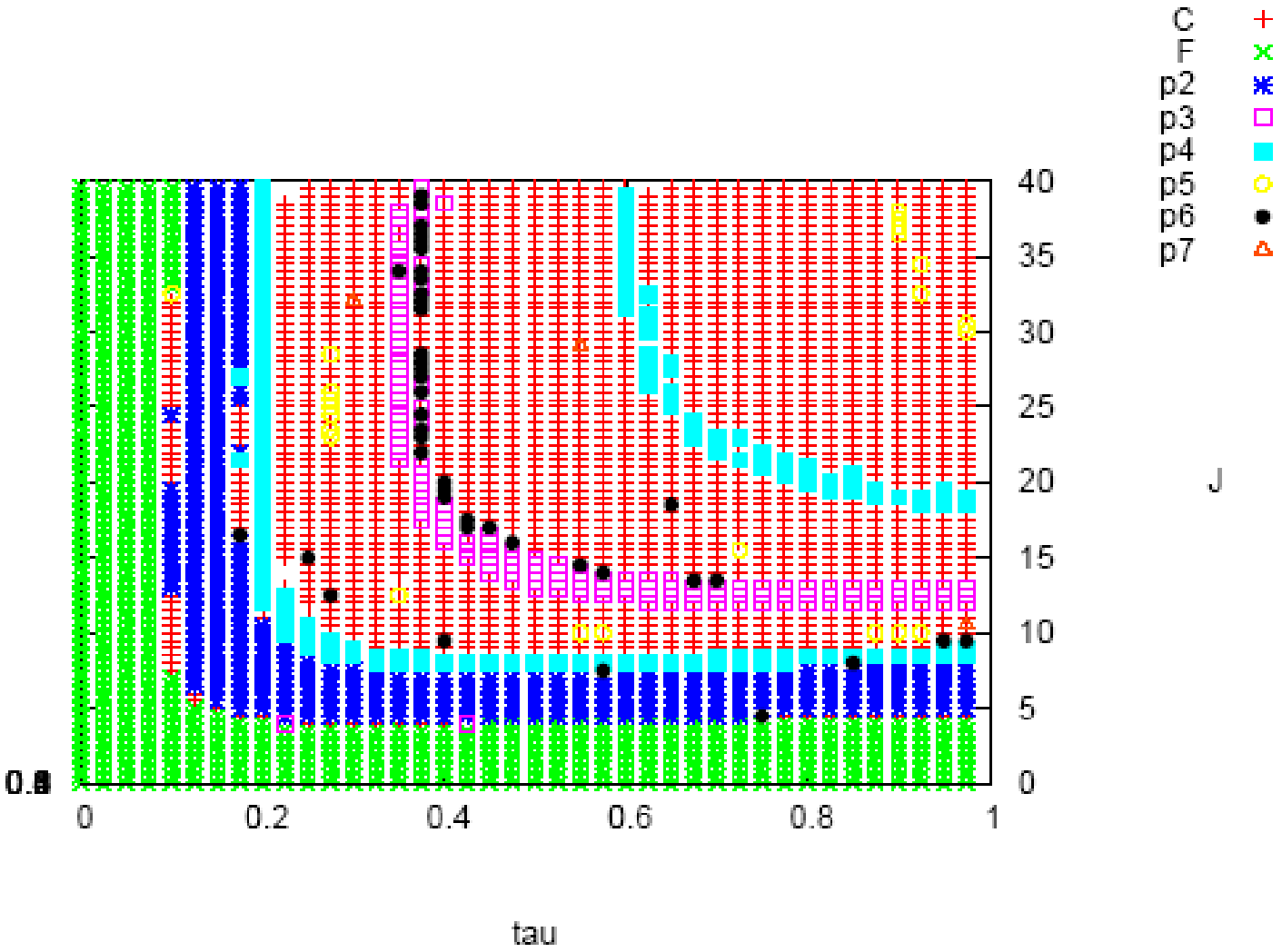
with

$$p = p_0 - \alpha M_{ij}\sigma_j$$

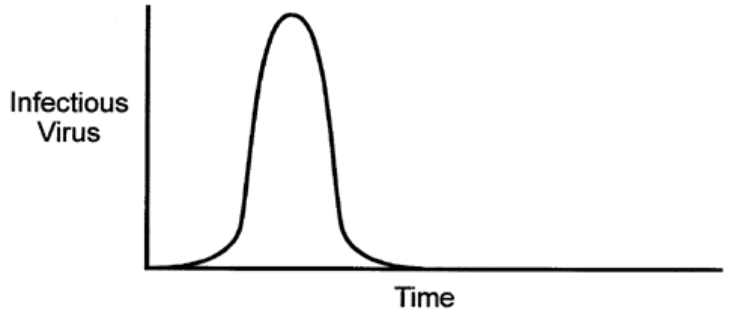
Mean field of percentage of infected

$$c^{t+1} = \sum_{k=1}^{N} (Nk) c^k (1-c)^{N-k} [1 - (1 - p_0 + \alpha k)^k]$$


cm k=2
cm k=5
cm k=10
cm k=50
cm k=100
c
J
0
0.2
0.4
0.6
0.8
1
0
5
10
15
20
25
30
35
40

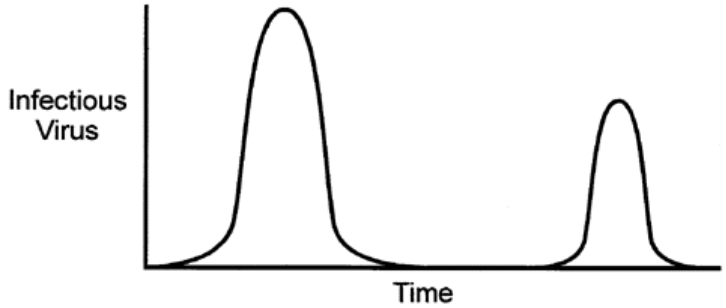

C
F
p2
p3
p4
p5
p6
p7
J
0
5
10
15
20
25
30
35
40
0.8
0
0.2
0.4
0.6
0.8
1
tau

ACUTE



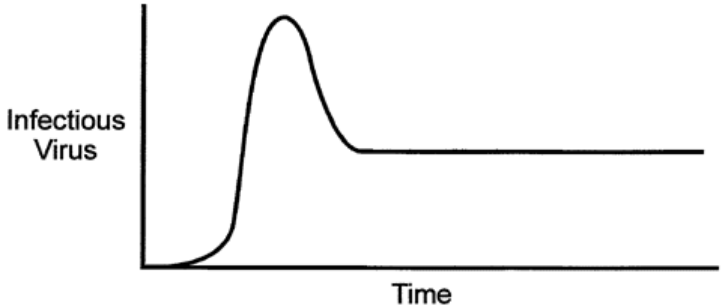
Rhinovirus
Influenza
Yellow Fever

LATENT PERSISTENT



Herpes simplex
Varicella-zoster
Measles-SSPE

CHRONIC PERSISTENT

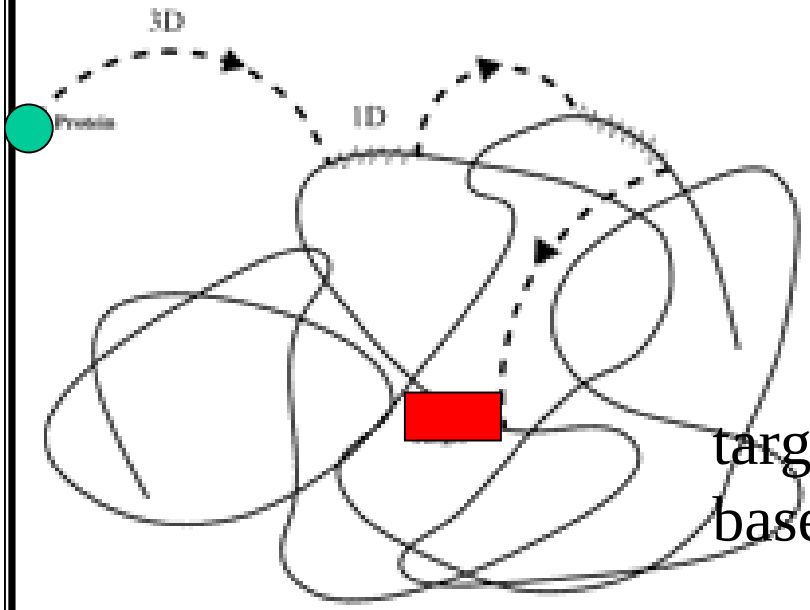


Hepatitis B
LCMV in Newborn Mice

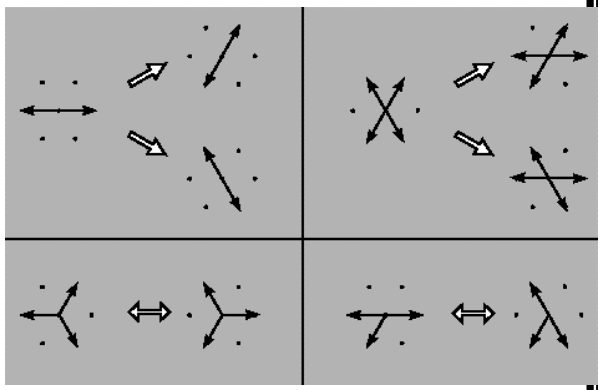
FIG. 2. General patterns of viral infection.

REACTIVE LATTICE GAS MODEL of METABOLISM

How a protein finds its target site on the DNA is still unclear. Here, I use a lattice gas (LG) model (figure below on the right shows some details of simulating molecular collisions in a LG) to describe the kinetics of target site localization of a protein on DNA. The model explicitly combines one-dimensional motion along the DNA and three-dimensional excursions in the solution (below left).



target site: (5-6 DNA bases in 3 billions).



CA and PDE of LIVER REGENERATION

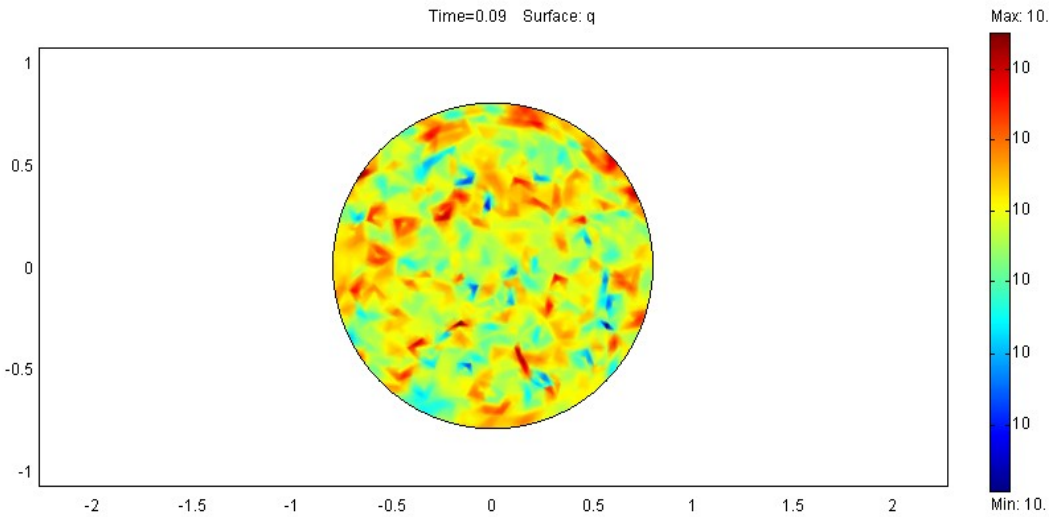
Models of early stages of vascular network formation (angiogenesis) and liver regeneration.

$$\frac{\partial n}{\partial t} + \nabla(nv) = \sigma n(\tanh(o - o_0))$$

$$\frac{\partial v}{\partial t} + v\nabla v = \beta \nabla c$$

$$\frac{\partial c}{\partial t} = D\Delta c + \alpha - \tau^{-1} c -$$

$$\frac{\partial o}{\partial t} = d\Delta o - \mu no + \gamma n(\tanh(o - o_1$$



where c is an aggregation factor (that acts effectively as a chemotactical agent), n is the number of cells in the continuum approximation, v is the velocity field , o is the oxygen and Neumann boundary conditions The figures above show the lobule patterns and the initial stages of vascular cell aggregation. I will consider medical applications.

Comp and System Biology at CL

Pietro Lio’

Richard van der Wath
Building stochastic genetic networks of intra -and inter-cellular communication (e.g. embryo development, organ regeneration and tumour growth)

Elizabeth van der Wath
Embedding Gene expression data analysis and Phylogenetic applications into the EGEE grid infrastructure.

Collaborations

Franco Bagnoli

Luca Sguanci