Coexistence of metabolically co-operating replicators in a cellular automaton: the importance of space without mesoscopic structure

TAMÁS CZÁRÁN[1,2] AND EÖRS SZATHMÁRY[1,2]

[1]*Ecological Modelling Research Group, Dept of Plant Taxonomy and Ecology, Eötvös University, Ludovika tér 2, H-1083 Budapest, Hungary.*

[2]*Collegium Budapest (Institute for Advanced Study), Szentháromság u. 2, H-1014 Budapest, Hungary*

Fax: +36 1 457 7610, E-mail: szathmary@colbud.hu

Word count:

Page heading: *Spatial dynamical molecular coexistence*

Key words: molecular coexistence, cellular automata, mesoscopic structure, hypercycle, origin of life, metabolic systems

SUMMARY

We analyze the dynamics of metabolically coupled replicators on a surface by a cellular automaton model, implementing a reaction-diffusion system. Replicators contribute to metabolism synergistically, described by a multiplicative function. Despite this coupling, the corresponding ordinary differential equation system leads to competitive exclusion by the faster replicating species, which in turn results in extinction of the whole system. To the contrary, coexistence of replicators with different per capita fitnesses occurs in the spatial, discrete system, due to an advantage of rarity of less vigorously growing replicators. Synergism means obligatory complementation, and a rare species is more likely to be complemented by synergistic partners in a finite neighbourhood. Increasing neighbourhood size leads to system collapse because thus the limiting case of the homogeneous, well-stirred reactor arises. In contrast, increasing diffusion rate does not result in competitive exclusion, because a so-called trait-group model rather than a homogeneous system is approached. Harmful parasites are selected against in the model. The importance of such models is related to suggestions that life arose on mineral surfaces.

1. INTRODUCTION

The role of spatial population structure in promoting cooperation and mutualism has received much interest recently (e.g. Nowak & May 1992; Hammerstein & Hoekstra 1995; Killingback & Doebeli 1996). It has also been emphasized in the context of the origin of life, for essentially the same reason: it is a means to establish coexistence of potentially competing template replicators (see Maynard Smith & Szathmáry 1995 for review). There are three known, detailed model approaches: (1) structured deme type models (Wilson 1980; Michod 1983;

Szathmáry 1992); (2) replication-diffusion systems as modelled by cellular automata (Boerlijst & Hogeweg, 1991); and group selection of replicators encapsulated in compartments (Szathmáry 1986; Szathmáry & Demeter 1987; Maynard Smith & Szathmáry 1993).

The motivation of these studies originates with the seminal paper of Eigen (1971), arguing (i) that primitive genomes must have been segmented (consisting of physically unlinked genes); (ii) that these unlinked genes must have had the tendency to compete with one another; and, as a consequence, (iii) that some mechanism ensuring their coexistence was needed. He saw the hypercycle to fulfill this role: a system of cyclically interacting molecular mutualists. It turned out to be the case that the hypercycle in a spatially homogeneous setting is vulnerable to parasitism: a cheating replicator that does not give catalytic aid to any member of the cycle can kill it off, provided it receives more help from the cycle than the member that it competes with in the first place (Fig. 1). As it was realized, compartmentation is an efficient means to separate bad from good genes (Maynard Smith 1979; Eigen *et al.* 1981). The stochastic corrector model demonstrated that once we have compartments, their internal organization need not be hypercyclic: a so-called metabolic coupling of replicators is sufficient (Fig. 2), whereby genes contribute to the good of the compartment by catalyzing its metabolism at various points (Szathmáry & Demeter 1987). The stochastic corrector models assumes that there is an optimal template composition of compartments, which gives the highest protocell division rate (reviewed by Szathmáry & Maynard Smith 1995). Variation between compartments is generated by the stochastic effects in template reassortment upon cell division as well as in replication. Natural selection between the compartments is acting on this variation generated by stochasticity.

The structured deme type models essentially led to the same conclusion: interaction with only neighbours promotes coexistence. Michod (1983) showed

the resistance of the hypercycle against parasites in this setting. The viability of the hypercyclic (Fig. 1) and metabolic systems (Fig. 2) was also demonstrated in such a context (Szathmáry 1992).

The cellular automaton approach applied so far to the problem of information integration (Boerlijst & Hogeweg 1991) is rather different from both the stochastic corrector and the structured deme framework. It is basically a discretised reaction-diffusion system: replication and diffusion of templates is imagined to take place on an adsorbing surface, without compartmentalisation. Resistance of a hypercycle against parasites is possible in such a reaction-diffusion system, provided the number of replicators exceeds four. The reason for this is that spiral waves emerge as spatial manifestations of the temporal limit cycle trajectory, itself the immediate consequence of the intransitive circle of mutualistic interaction assumed in the hypercycle model. Without the spirals, e.g. with fewer replicators, this particular system collapses if a parasite invades. Cronhjort and Blomberg (1995) have studied numerically the partial differential equation model of the same reaction-diffusion system, and found that the section of parameter space allowing for parasite resistance is smaller than it is in the cellular automaton of Boerlijst & Hogeweg (1991). The spirals are spontaneously emerging self-organised units of selection in the reaction-diffusion approach, which to a certain extent play the role of the compartments in the other two. Spiral waves as units of selection are of course much less definite than the compartments of the structured deme and the stochastic corrector models: whether a replicator molecule belongs to a certain spiral is not always easy to decide.

The fundamental difference of our cellular automaton model as compared to that of Boerlijst and Hogeweg (1991) is exactly in what the stochastic corrector differs from the hypercycle: the dynamical link among the replicator types is realised through a common metabolism, instead of the direct, intransitive

hypercyclic coupling. Moreover, unlike in the cellular automaton approximations to reaction-diffusion systems, the corpuscular appearence of the replicator macromolecules is not a methodological compromise in our model, but an essential feature of the object to be investigated: populations of macromolecules can be hardly imagined as fluids consisting of an infinite number of sizeless mass-points. We will show that corpuscularity can be, and in our model it is, of profound dynamical importance (cf. also Durrett & Levin 1994).

Using the cellular automaton model of the metabolic system, our aim was to show that (i) metabolic coupling can lead to coexistence of replicators in spite of an inherent competitive tendency; (ii) parasites cannot easily kill the whole system; (iii) complexity can increase by natural selection; (iv) varying the critical neighbourhood size and the diffusion rate, one can appoximate the behaviour of different other models.

2. THE MODEL

The model is a stochastic cellular automaton (Wolfram 1984; Czárán & Bartha 1992; Durrett .....) consisting of a 300 x 300 grid of sites with a toroidal topology (wrap-around margins) to avoid edge effects. Each site of the grid can contain at most one replicator molecule. At $t = 0$, half of the sites are occupied by n types of replicators; the types are equally abundant in the initial pattern. One generation consists of three essential component processes: replication, decay and diffusion. We discuss these in turn.

For a replication event of any template s into a neighbouring empty site to occur, s must be complemented by all other types present in its neighbourhood of a given size (Fig. 3). The claim of template s to replicate into an adjacent empty site is:

$$C(s) = k_s \left[\prod_{i=1}^{n} f(i) \right]^{1/n} \qquad (1)$$

where *f(i)* is the copy number of replicator type *i* within the neighbourhood of *s* and k_s is its specific replication rate. Notice that $C(s)$ is proportional to the geometric mean of the within-neighbourhood replicator frequencies. The chance of *s* to replicate into the empty site is:

$$P_s = \frac{C(s)}{C_e + \sum_l C(l)} \qquad (2)$$

where *l* are the four orthogonal nearest neighbours (the Neumann neighbourhood) of the empty site, and C_e is the claim of the empty site to remain empty. Thus the probability that the empty site remains empty is:

$$P_e = \frac{C_e}{C_e + \sum_l C(l)} \qquad (3)$$

Note that the spatially homogeneous dynamics (that is, a possible continuous mean-field approximation - cf. Durrett .....) for this system would be

$$\frac{dx_i}{dt} = x_i [k_i M(\mathbf{x}) - j(\mathbf{x})] \qquad (4)$$

where x_i is the concentration of replicator i, of which k_i is the growth rate. $M(\mathbf{x})$ expresses the effect of metabolism on the replication rate. It is a common multiplicative function of the replicator concentrations **x**, so that each replicator type needs the presence of all the others to be able to replicate, but the metabolic help received is aspecific. j is the outflow function (cf. Eigen 1971) acting as a density (concentration) dependent selection constraint. Although all replicators must be present for M to be positive, we know that this does not preclude competitive exclusion of all other types by the fittest (of largest k_i) replicator (Eigen & Schuster 1977), since M is the same in all equations of (4), and thus even very small positive concentrations of the competitively inferior types maintain the advantage of the dominant, in terms of the speed of replication. That is, we analyze the beneficial effect of spatial structure on coexistence on the basis of a worst-case assumption, similar to that of the stochastic corrector model (Szathmáry & Demeter 1987).

Decay is aspecific, defined as a constant probability for an occupied site to become empty in time $t + 1$, irrespective of what type of replicator it harbored in time t. The competitive (or chance) exclusion from the grid of any one replicator type results in the disappearance of all the other types as well, since they cannot replicate any more but decay continues.

Following Boerlijst and Hogeweg (1991), we modelled the diffusive movement of replicators by the algorithm of Toffoli and Margolus (1987), which preserves particle number and frequency distribution within the grid. One diffusion step is complete by executing the following three instructions: (i) divide the grid into 2 x 2 subgrids within a fixed frame; (ii) rotate each subgrid 90° clockwise or anticlockwise with equal probability; (iii) shift the grid frame one site diagonally with isotropic probability. Diffusion rate increases with the number of diffusion steps between two replication phases.

The update procedure of the cells in the replication phase can be either synchronous or asynchronous: since this is known to affect the arising spatial patterns as well as coexistence (Huberman & Glance 1993), we used both, and for the latter we chose a random update: one site was updated at a time, and that one was chosen at random. To ensure that each site was updated once on the average, the replication phase consisted of a number of updates equal to that of the sites in the grid.

The computer program implementing the cellular automaton was written in MS-FORTRAN 5.1, and run on an ALR Evolution Dual 6 (Dual Pentium Pro 133) machine. The program needs a lot of computing power, especially in terms of CPU time, therefore we could not run the program many times with all parameter sets. We produced replicate runs for a few points of the parameter space; the differences among replicate runs were negligible in all cases.

3. RESULTS

The most important, and also somewhat surprising result of the simulations is that the cellular automaton is capable of producing coexistence in a large part of its parameter space. No conspicuous mesoscopic pattern, similar to spiral waves, arises in any experiment, since a relatively homogeneous spatial distribution of the replicator types is necessary for many neighbourhoods to contain a metabolically sufficient set of macromolecules. Without this, there would not be enough replication events to compensate for the loss due to spontaneous decay. That is, a persistent system cannot show an aggregated pattern.

From the viewpoint of template coexistence, the most relevant parameters of the model are diffusion rate, neighbourhood size and system size (the latter being the number of different, metabolically necessary template types). We take a

closer look at the effects of these and also of the introduction of parasites below. Fig. 4. summarizes the ecology of the replicator system for different parameter sets.

Diffusion. Increased diffusion rate promotes coexistence. Boerlijst and Hogeweg (1991) were surprised to see that this did not lead to the reappearance of the spatially homogeneous dynamics in their model either: we give an explanation in the Discussion.

Neighbourhood size. For a fixed system size, there is an optimum neighbourhood size below which the chance that it contains a metabolically complete set of replicators is too small or even zero (metabolism "does not fit in") and above which replicators start to "feel" the overall population density rather than a strictly local one.

System size. For any fixed combination of neighbourhood size and diffusion rate, an increase in system size ultimately leads to the collapse of the system for the reason discussed above in relation to neighbourhood size reduction: neighbourhood size is the upper limit of system size. Decreasing system size makes coexistence more likely in any parameter setting, but it is to be noted that we do not consider the absolute efficiency of metabolism to be a function of system size in this model. This would be reasonable to assume, however, since more replicators might be more efficient in catalysing metabolism, giving more chance of survival for the larger system.

Parasites. By definition, parasite molecules are replicated by the others through metabolism, but they do not contribute to it. We found that a coexistent system cannot be killed by a parasite, even if its replication rate exceeds that of the fastest "altruistic" macromolecule type. If the parasite is very efficient, it can depress the concentration of the metabolically active replicators by simply occupying most of the surface of the substrate, but even then, local

neighbourhoods containing fewer copies of the parasite will be positively selected. This gives an advantage for the altruistic members of the system against the parasite: wherever there are too many cheaters, the system slows down in relative terms, while localities of comparatively low parasite concentrations speed up in replication. The result is a system stably coexisting with the parasite, as shown on Fig. 4.

4. DISCUSSION

The result that there is coexistence without any mesoscopic emergent pattern is robust and counter-intuitive. It is due to the inherent dicreteness (i.e., the corpuscular nature of the replicator molecule populations) and spatial explicitness of the model, which grasp essential features of the living world in general, and macromolecular replicator systems in particular. An inferior (that is, slowly replicating) molecule type does not die out since there is an advantage of rarity in the system: a rare template is more likely to be complemented by a metabolically sufficient set of replicators than a common one (cf. Fig. 3). This effect stems in the joint effect of corpuscularity and spatiality: neither is sufficient without the other.

What would happen in a spatial, but still continuous system could be best demonstrated by turning to the reaction-diffusion equivalent of (4). Obviously this would not help the system to become coexistent: the competitive dominance order of the replicators would be the same at every spatial location **u**, because the metabolic function $M(\mathbf{x}, \mathbf{u})$ is the same for all replicators at any location. The reaction-diffusion version of the model would collapse like (4), because the macromolecule type of largest replication rate k_i would dominate everywhere

everytime, thus driving the whole system to extinction (cf. Cronhjort & Blomberg 1995).

Allowing for corpuscularity alone without spatial inhomogeneity would not enhance coexistence either. This situation would best be modelled by our cellular automaton with a neighbourhood size equal to the size of the grid. In such a model, all possible neighbourhoods would be of the same composition, independently of spatial position. Increased neighbourhood size mimics, in contrast to increased diffusion rate, better mixing, whereby global rather than local densities determine metabolic efficiency. As demonstrated on Fig. 4, increasing neighbourhood size is detrimental for coexistence, which also means at the same time that abolishing spatiality would have the same effect.

Increased diffusion rate promotes coexistence because by it the system approaches the following dynamics: local replication -- random reassortment of groups -- local replication and so on. This is practically a so-called "trait group" model *sensu* Wilson (1980), for which template coexistence has been demonstrated analytically for the hypercyclic as well as the metabolic systems (Szathmáry 1992). Neighbourhood interaction represents a kind of temporary compartmentalisation, which helps maintaining the complete set of metabolically active replicators. Diffusion does not homogenise the compositions of the neighbourhoods below the resolution level of one site, therefore the advantage of rarity does not vanish with the intensity of diffusion getting larger. On the contrary, diffusion drives the two copies of a rare template type apart in space after replication, thus giving each a chance to replicate further in separate neighbourhoods.

Whether complexity can increase in such a system, is always a pertinent question. Since a parasite cannot kill the system, and the system cannot eliminate the parasite, the cheater, once appeared, can be around for a long time, and

mutate freely. Thus there is the possibility for a positive conversion to occur: the system can incorporate a mutant parasitic sequence into the metabolic machinery, and make it work for the common good. The beneficial effect might be facultative at first, giving only some (local) replication advantage to neighbourhoods containing the converted parasite, but later it can become an essential part of the metabolism. Once the interaction is obligate, we have a system one member larger, and metabolically more effective than the original. This scenario will be checked by future experiments. [Nem lehetne-e ezt megnezni?]

A general importance of surface dynamics seems more and more important for the origin of life in general: as Wächtershäuser (1988) pointed out, chemical evolution leading to more and more complicated networks, is likely to have taken place on the surface, especially on that of pyrite. Chemical dynamics on a surface occurs essentially in 2D. This has important thermodynamic and kinetic consequences. For example, an appropriate surface can act as a catalyst for the reactions in question. Water, liberated from the surface following condensation reactions leading to larger molecules, renders the reaction favourable, due to the increased entropy of the system as a whole. Surface dynamics of replicators with indefinite heredity is a natural outgrowth of this "primordial pizza" dynamics (cf. Maynard Smith & Szathmáry 1995). In this paper we have shown that spatiality and corpuscularity on a surface crucially change the outcome of selection, even for non-hypercyclic, but metabolically cooperative systems.

REFERENCES

Boerlijst, M. C. & Hogeweg, P. 1991 Spiral wave structure in pre-biotic evolution--hypercycles stable against parasites. *Physica* **D48,** 17-28.

Cronhjort M. B. & Blomberg, C. 1995 Hypercycles versus parasites in a two-dimensional partial differential equation model. *J. theor. Biol.* **169,** 31-49.

Czárán, T. & Barta, S. 1992 *Trends Evol. Evol.*

Durrett, ...

Durrett, R. & Levin, S. A. 1994 The importance of being discrete (and spatial). *Theor. pop. Biol.* **46,** 363-394.

Eigen, M. 1971 Self-organization of matter and the evolution of biological macromolecules. *Naturwiss.* **58,** 465-523.

Eigen, M. & Schuster, P. 1977 The hypercyce: a principle of natural self-organization. Part A: emergence of the hypercycle. *Naturwiss.* **64,** 541-565.

Eigen, M., Schuster, P., Gardiner, W. & Winkler-Oswatitsch, R. 1981 The origin of genetic information. *Sci. Amer.* **244(4),** 78-94.

Hammerstein, P. & Hoekstra, R. 1995. Mutualism on the move. *Nature* **376,** 121-122.

Huberman, B. A. & Glance, N. S. 1993 Evolutionary games and computer simulations. *Proc. natn. Acad. Sci. USA* **90,** 7716-7718.

Maynard Smith, J. 1979 Hypercycles and the origin of life. *Nature* **280,** 445-446.

Maynard Smith, J. & Szathmáry, E. 1993 The origin of chromosomes I. Selection for linkage. *J. theor Biol.* **164,** 437-446.

Maynard Smith, J. & Szathmáry, E. 1995 *The Major Transitions in Evolution.* Freeman, Oxford.

Michod, R. 1983 Population biology of the first replicators: on the origin of genotype, phenotype, and organism. *Amer. Zool.* **23,** 5-14.

Nowak, M. A. & May, R. M. 1992 Evolutionary games and spatial chaos. *Nature, Lond.* **359,** 826-829.

Szathmáry, E. 1986 The eukaryotic cell as an information integrator. *Endocytobiol. Cell Res.* **3,** 113-132.

Szathmáry, E. 1992 Natural selection and the dynamical coexistence of defective and complementing virus segments. *J. theor. Biol.* **157,** 383-406.

Szathmáry, E. & Demeter L. 1987 Group selection of early replicators and the origin of life. *J. theor Biol.* **128,** 463-486.

Szathmáry, E. & Maynard Smith, J. 1995 The major evolutionary transitions. *Nature, Lond.* **374,** 227-232.

Toffoli, & Margolus, ...

Wächtershäuser, G. 1988 Before enzymes and templates: theory of surface metabolism. *Microbiol. Rev.* **52,** 452-484.

Wilson, D.S. 1980 *The Natural Selection of Populations and Communities.* Benjamin Cummings, Menlo Park, California.

Wolfram, S. 1984 Cellular automata as models of complexity. *Nature* **311,** 419-424.

Figure 1. Cooperation of replicators (I_1,...,I_4) in the hypercycle. Circular arrows: replication; solid arrows: direct catalytic help in replication.

Figure 2. Cooperation of replicators (I_1,...,I_4) through a common metabolic system (M). Circular arrows: replication; dashed arrows: supply of monomers for replication; solid arrows: catalytic help for monomer production through metabolism.

Figure 3. Types of neighbourhoods applied in the simulations. (a) The replication neighbourhood (grey sites with x) of an empty (black) site. (b) Catalytic neighbourhoods of the replicator in the focal site (grey site with x) besides an empty (black) site. Grey sites: 3x3 neighbourhood; grey + white sites: 5x5 neighbourhood.

Figure 4. Time series of simulation results with different system sizes (*NR*), metabolic neighbourhood sizes (*NHS*) and diffusion (*DIFF*). Other parameters: C_e (vote weight of empty sites) is 2.0, p_d (decay probability of the replicators) is 0.2 in all cases; k_1 = 2.0, k_2 = 4.0 and k_3 = 6.0 if *NR* (system size) is 3; k_1 = 2.0, k_2 = 10/3, k_3 = 14/3 and k_4 = 6.0 if *NR* is 4.

Figure 5. The effect of increasing metabolic neighbourhood size on the persistence of a system of 3 replicator types and slow macromolecule diffusion. Parameters as in the middle graph of **Figure 3**, the only difference being *NHS* = 9x9.

Figure 6. The effect of a fast parasitic sequence appearing in a persistent 3-replicator system. Notice that the parasite is suppressed by the cooperative replicators, but it remains persistent. Parameters: *NHS* = 3x3, *DIFF* = 100, C_e = 2.0, p_d = 0.2, k_1 = 2.0, k_2 = 4.0, k_3 = 6.0 and k_p (the replication constant of the parasite) = 8.0.

Figure 7. A possible scenario for the evolution of system size. (a) a persistent 3-replicator system with parameters *NHS* = 5x5, *DIFF* = 100, C_e = 2.0, p_d = 0.2, k_1 = 2.0, k_2 = 4.0, k_3 = 6.0; (b) the same system with a parasite of k_p = 8.0 introduced; (c) the same system with the parasite converted to a cooperative (just as essential as the other 3) member of the system.

Parameter	Description	Values taken
C_e	Vote weight of empty cells for remaining empty	2.00
p_d	Decay probability (per generation) of replicators	0.20
k_i	Replication constant of replicator type i	2.00 - 8.00
NR	Number of replicator types (system size)	3 - 9
NHS	Metabolic neighbourhood size	3x3 - 9x9
DIFF	Number of diffusion steps per generation	1 - 300

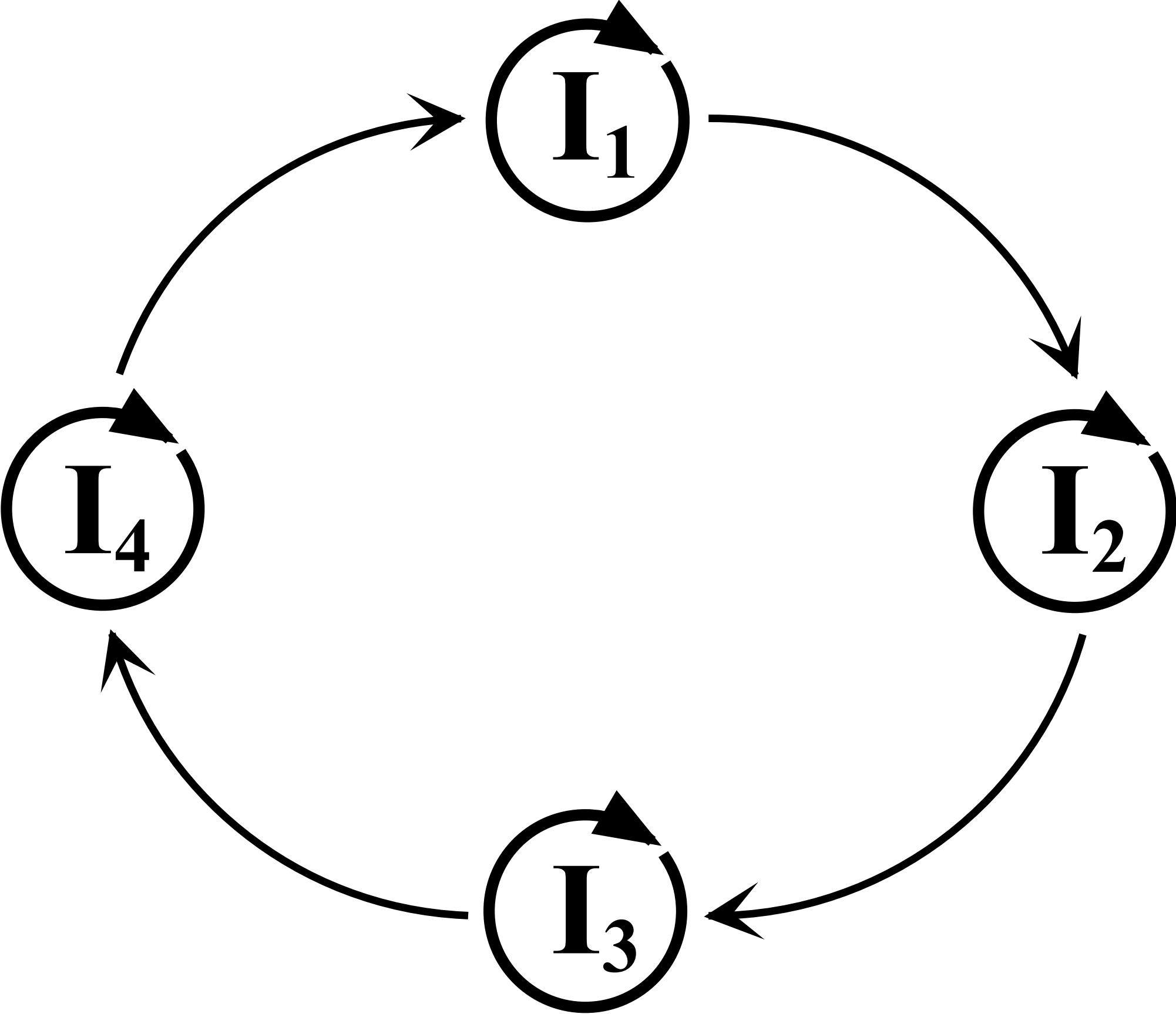
Table I. Parameters of the model

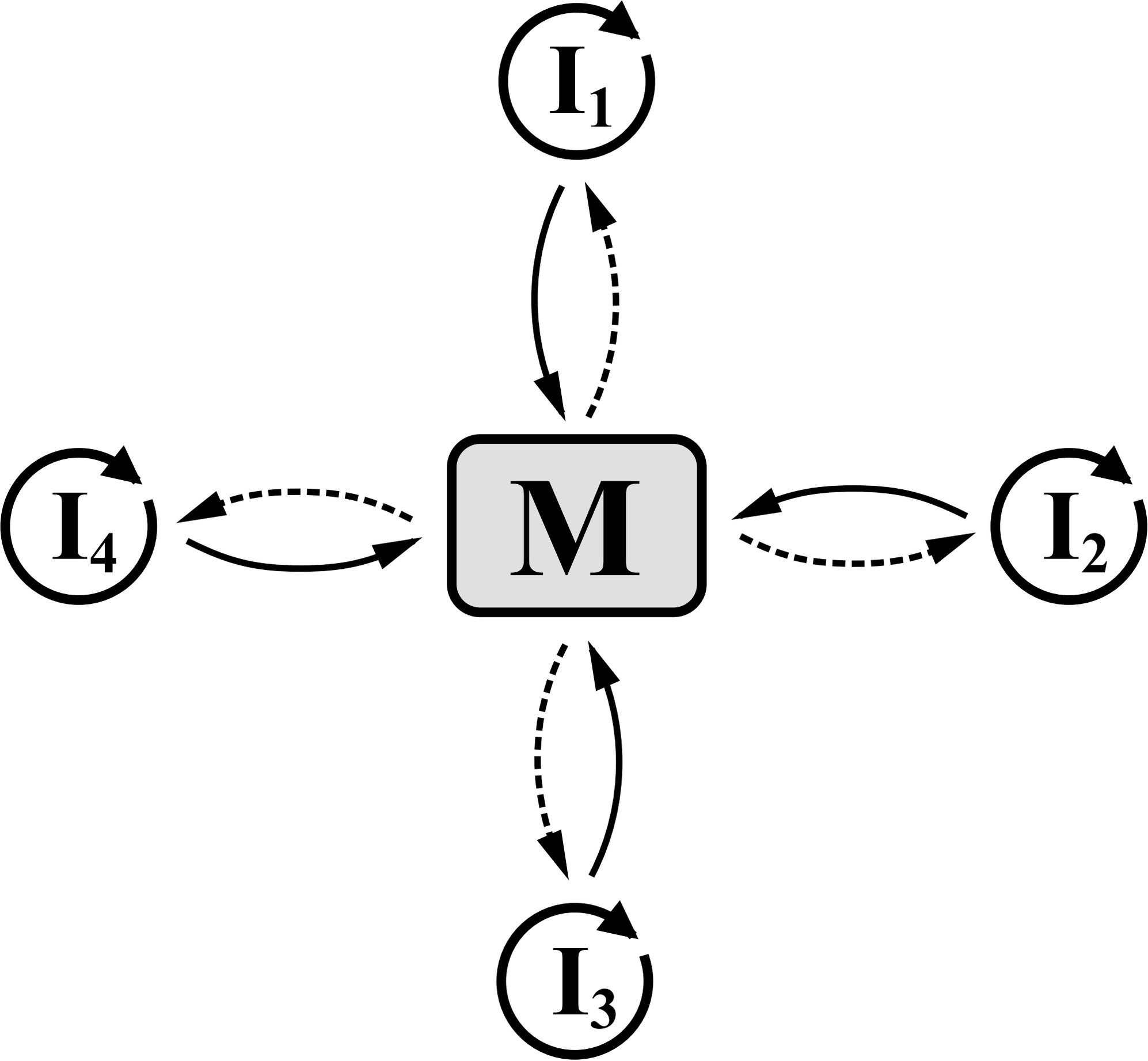
Parameter change	Direct effect	Effect on persistence
NR ↑	↓ Chance that a metabolic neighbourhood contains a complete set of replicators	↓
NHS ↑	↑ Chance that a metabolic neighbourhood contains a complete set of replicators ↓ Advantage of rarity	↑ ↓
DIFF ↑	↑ Spatial mixing (high replicator density) ↑ Spatial mixing (low replicator density)	↑ * ↓ **

* The system approaches a Wilson (1980) type trait group mechanism (group selection on neighbourhood configurations)

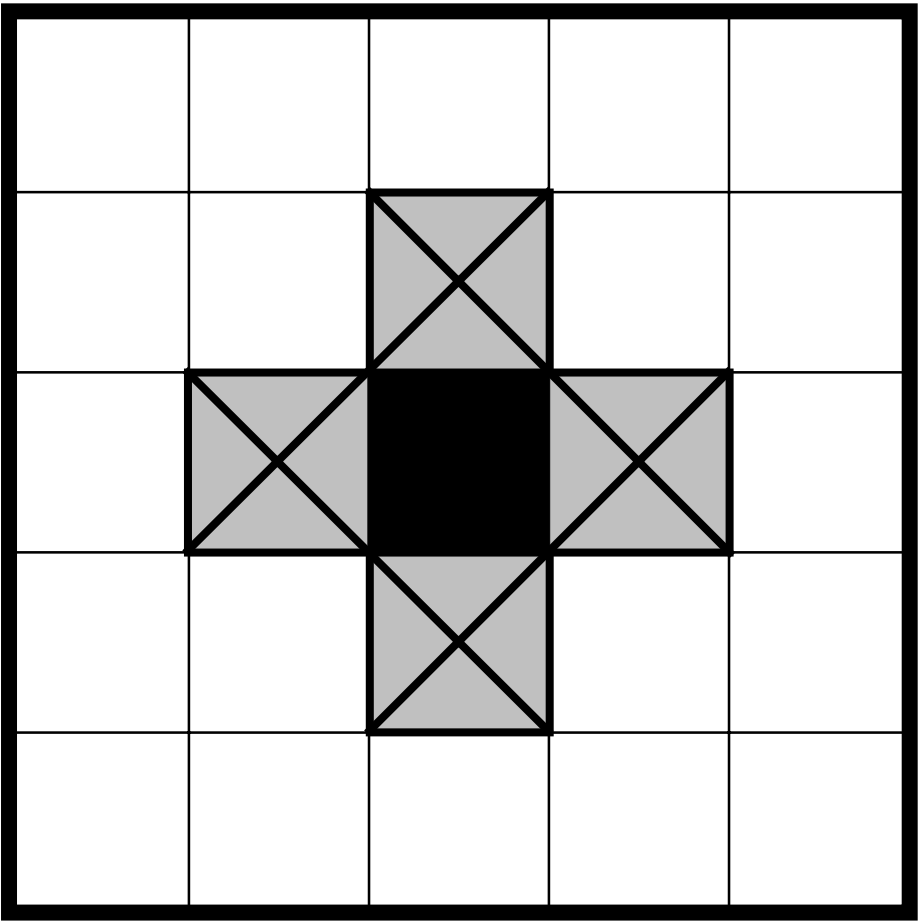
** Replicators are dispersed apart (quickly exterminates hopeless systems)

Table II. Summary of the effects of parameter changes on persistence

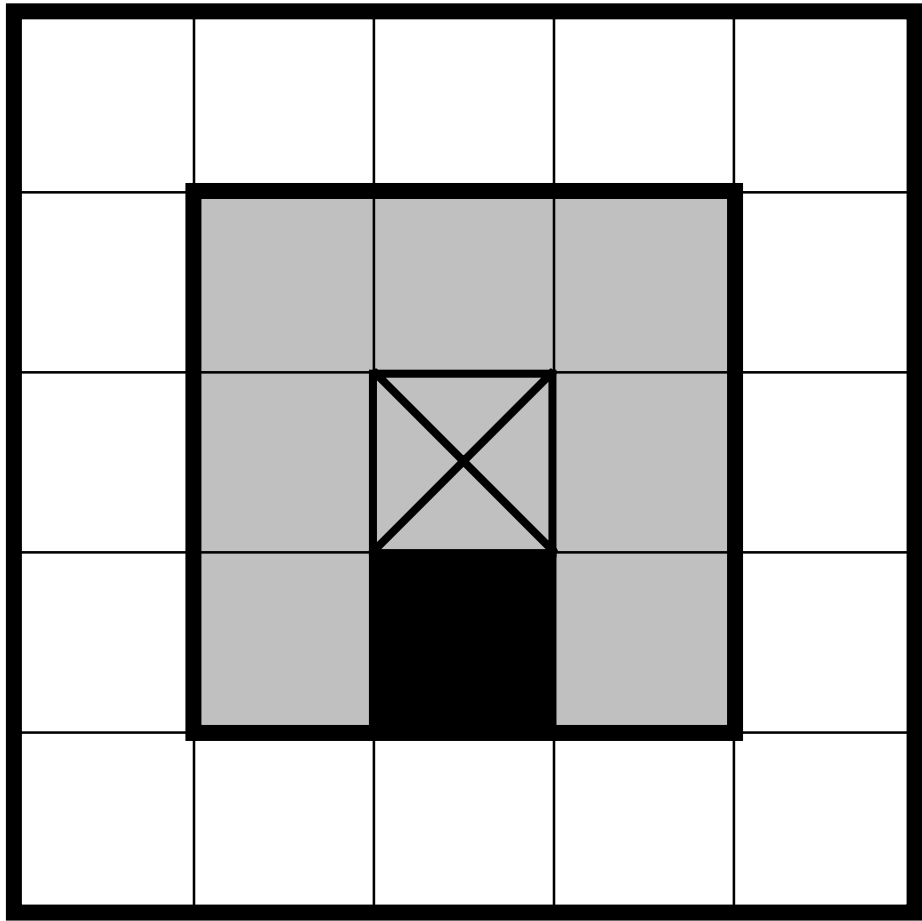

I_1
I_2
I_3
I_4

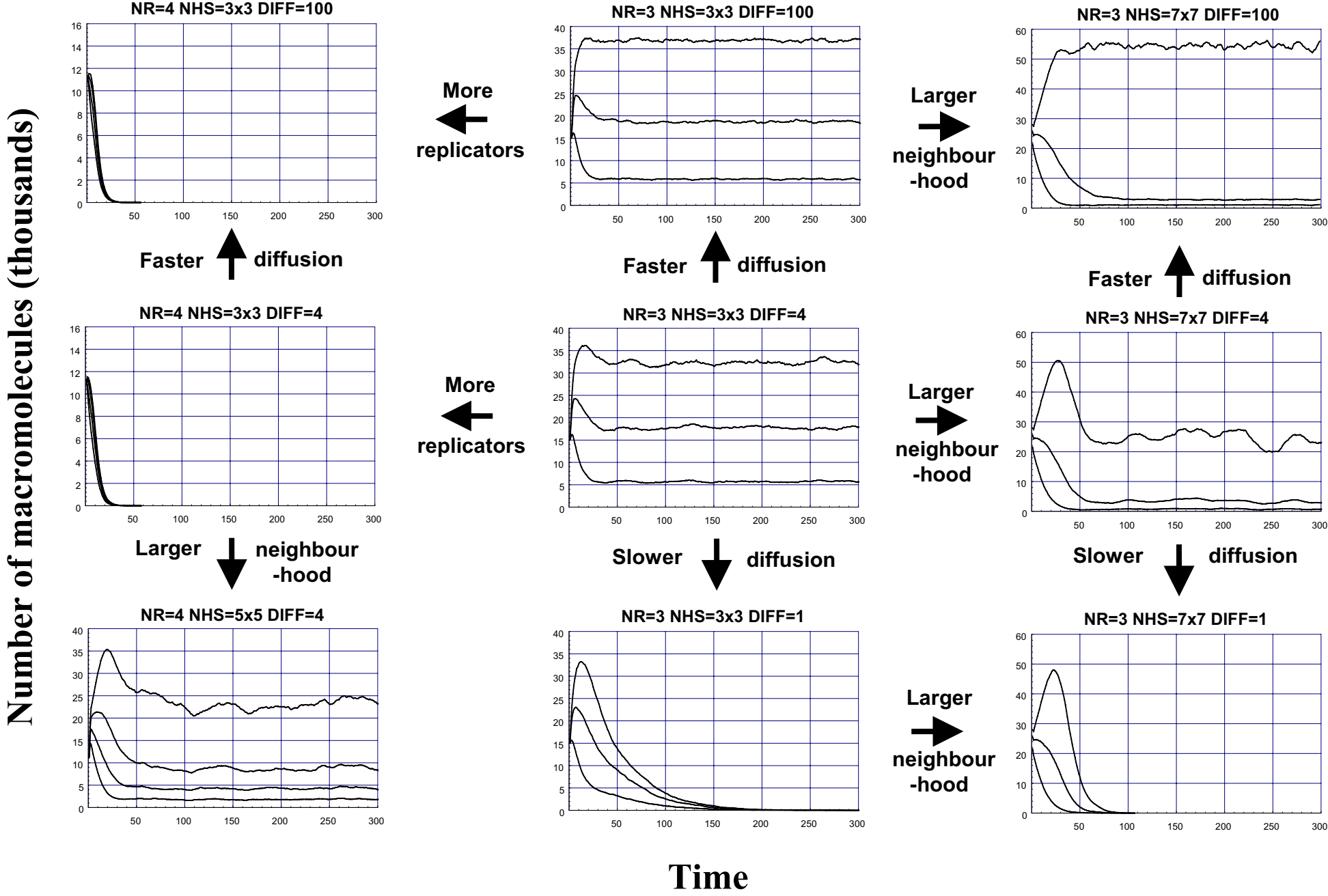

I1
I4
M
I2
I3

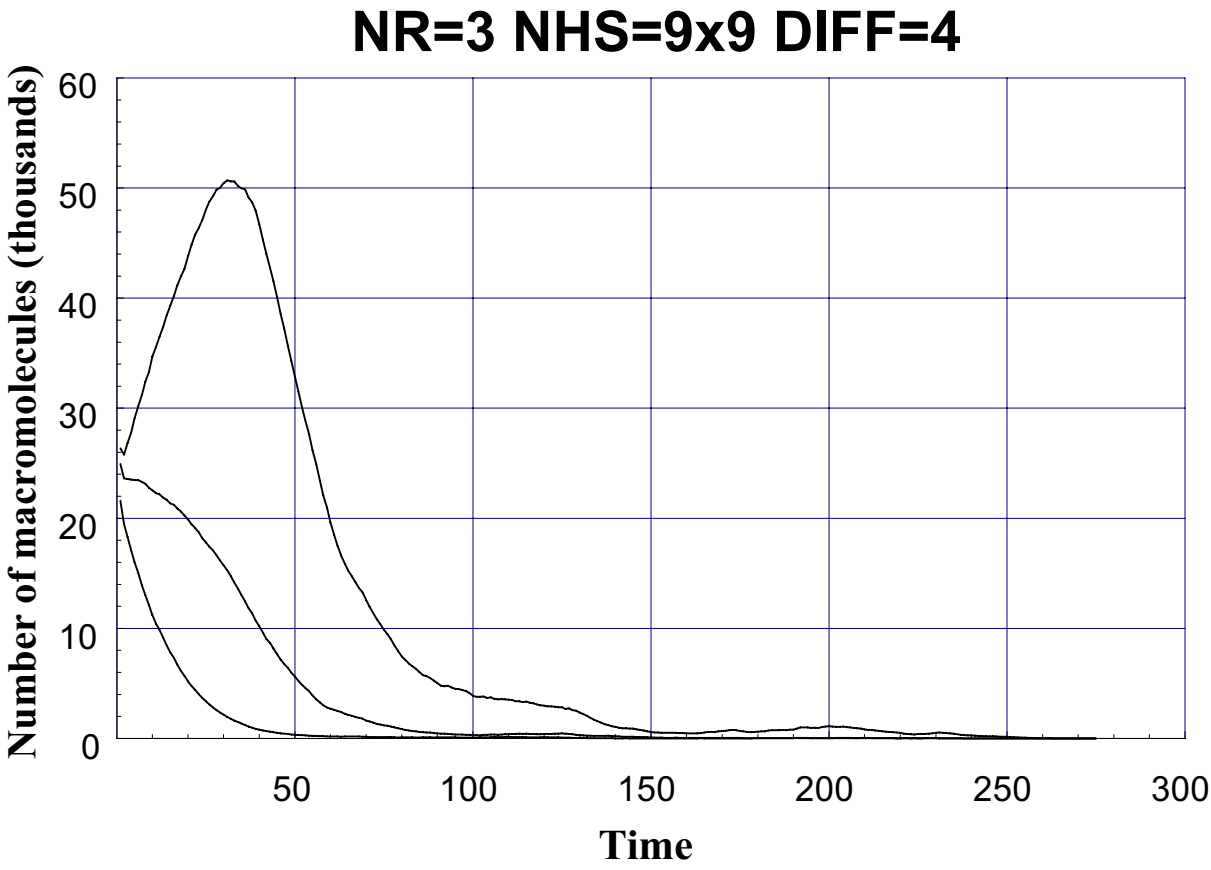
(a)

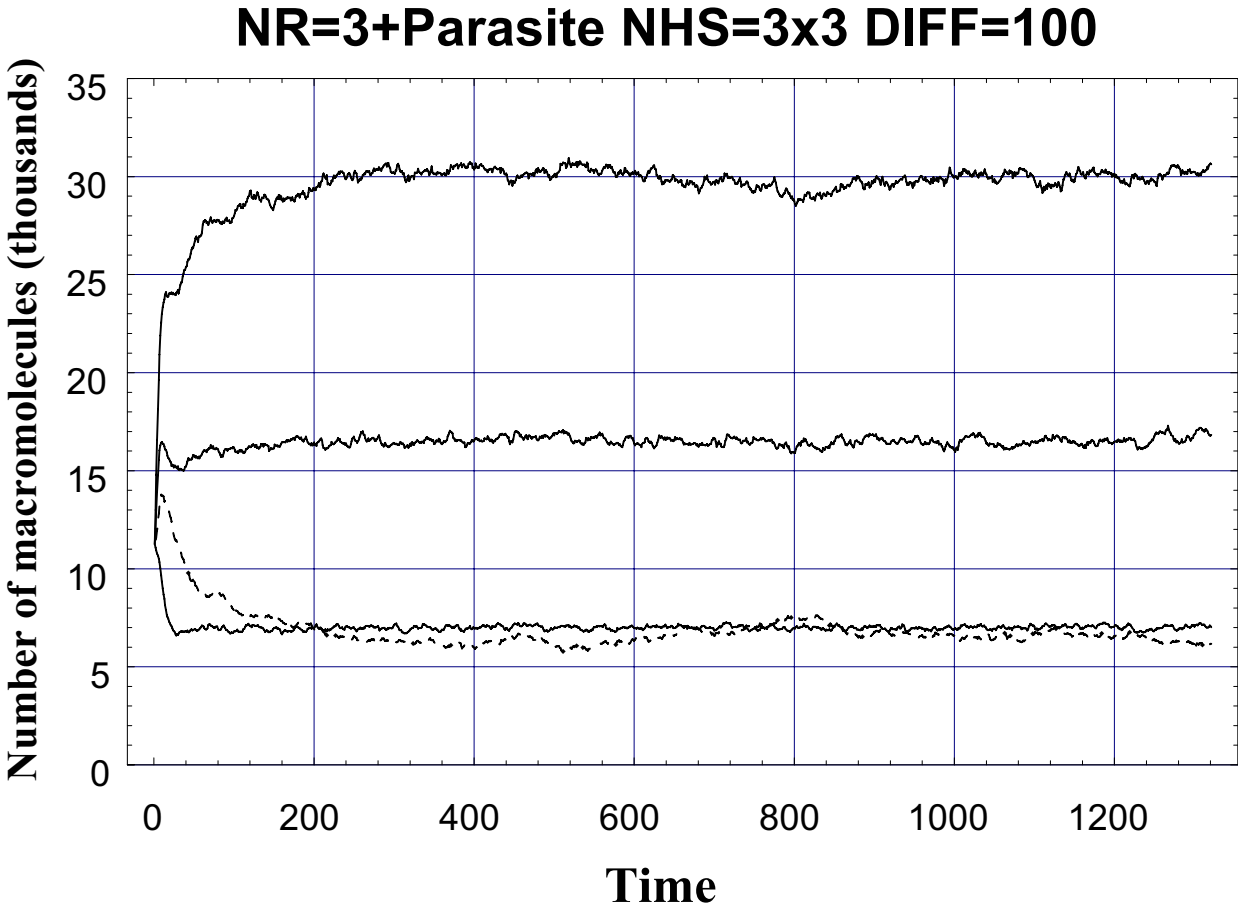


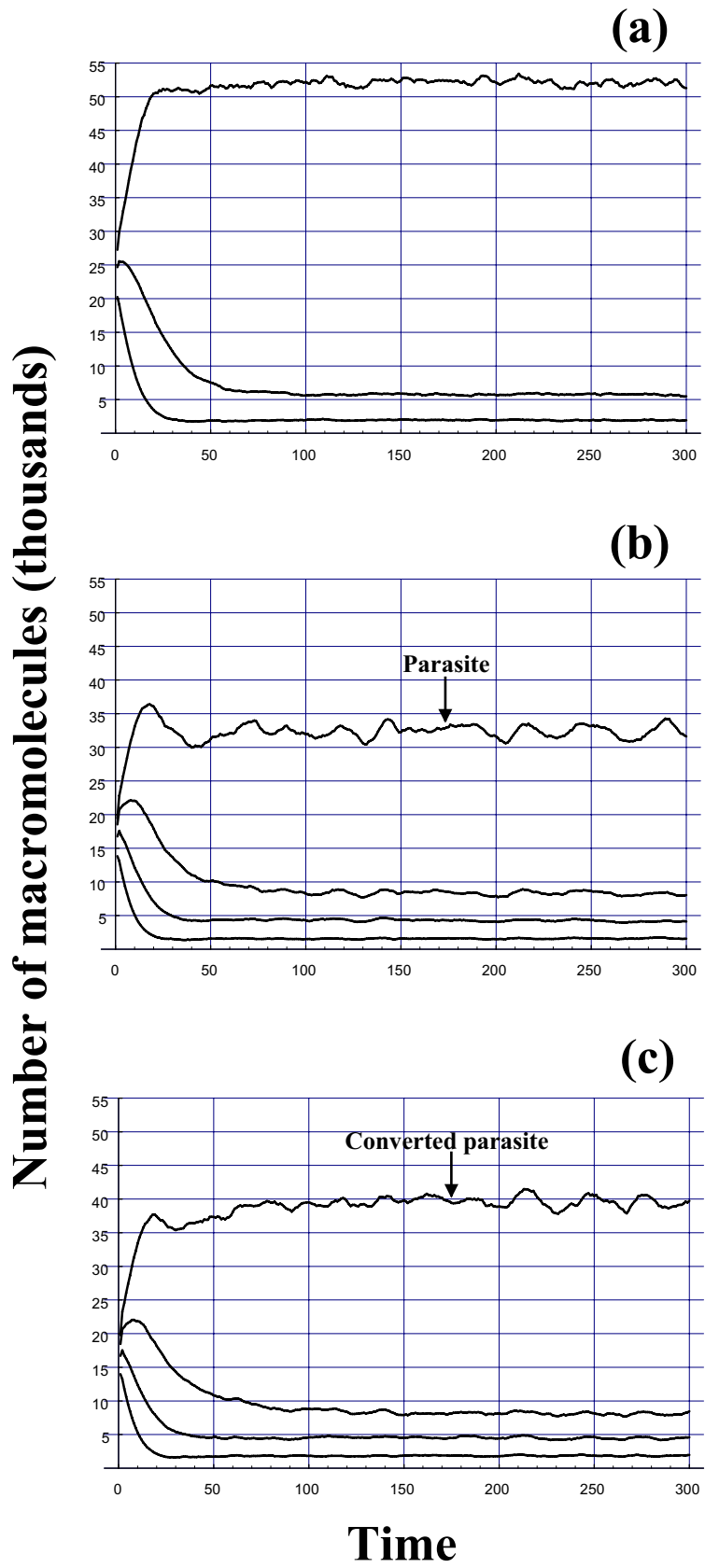
(b)




NR=4 NHS=3x3 DIFF=100
NR=3 NHS=3x3 DIFF=100
NR=3 NHS=7x7 DIFF=100
More replicators
Larger neighbour-hood
Faster diffusion
Faster diffusion
Faster diffusion
NR=4 NHS=3x3 DIFF=4
NR=3 NHS=3x3 DIFF=4
NR=3 NHS=7x7 DIFF=4
More replicators
Larger neighbour-hood
Larger neighbour-hood
Slower diffusion
Slower diffusion
NR=4 NHS=5x5 DIFF=4
NR=3 NHS=3x3 DIFF=1
NR=3 NHS=7x7 DIFF=1
Larger neighbour-hood
Number of macromolecules (thousands)
Time


NR=3 NHS=9x9 DIFF=4
Number of macromolecules (thousands)
60
50
40
30
20
10
0
50
100
150
200
250
300
Time


NR=3+Parasite NHS=3x3 DIFF=100
Number of macromolecules (thousands)
35
30
25
20
15
10
5
0
0
200
400
600
800
1000
1200
Time


(a)
(b)
(c)
Parasite
Converted parasite
Number of macromolecules (thousands)
Time
0
50
100
150
200
250
300
5
10
15
20
25
30
35
40
45
50
55