We provide a substitute for figure 2 of the original paper. In that figure, a mismatch exists in panels (a) and (b) between the time span of the data (0 ≤ t ≤ 20) and the markup on the abscissa axes (0 ≤ t ≤ 30). This is fixed in figure 2 of this erratum, where all data are depicted from t = 0 through t = 30. None of the original conclusions has been affected.
Figure 2. Evolution of the probability density of $x_A$ from $x_A(0) = 0.14$ (a), $x_A(0) = 0.15$ (b) and $x_A(0) = 0.16$ (c), for $p = r = 0.1$ and $\lambda = \mu = 0.1$. Data are based on at least $10^4$ independent instances of graph $D$ and are log-binned to the base 1.2. Probability densities are given according to the color-coded logarithmic scale on the right of each panel, ranging from $10^{-3}$ (at the bottom of the scale) to $10^4$ (at the top).
Quasispecies dynamics on a network of interacting genotypes and idiotypes: formulation of the model

Valmir C Barbosa\textsuperscript{1}, Raul Donangelo\textsuperscript{2,3} and Sergio R Souza\textsuperscript{2,4}

\textsuperscript{1} Programa de Engenharia de Sistemas e Computação, COPPE, Universidade Federal do Rio de Janeiro, Caixa Postal 68511, 21941-972 Rio de Janeiro, RJ, Brazil
\textsuperscript{2} Instituto de Física, Universidade Federal do Rio de Janeiro, Caixa Postal 68528, 21941-972 Rio de Janeiro, RJ, Brazil
\textsuperscript{3} Instituto de Física, Facultad de Ingeniería, Universidad de la República, Julio Herrera y Reissig 565, 11.300 Montevideo, Uruguay
\textsuperscript{4} Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, 91501-970 Porto Alegre, RS, Brazil

E-mail: valmir@cos.ufrj.br

Received 30 August 2014
Accepted for publication 13 December 2014
Published 19 January 2015

Abstract. A quasispecies is the stationary state of a set of interrelated genotypes that evolve according to the usual principles of selection and mutation. Quasispecies studies have for the most part concentrated on the possibility of errors during genotype replication and their role in promoting either the survival or the demise of the quasispecies. In a previous work, we introduced a network model of quasispecies dynamics, based on a single probability parameter ($p$) and capable of addressing several plausibility issues of previous models. Here we extend that model by pairing its network with another one aimed at modeling the dynamics of the immune system when confronted with the quasispecies. The new network is based on the idiotypic-network model of immunity and, together with the previous one, constitutes a network model of interacting genotypes and idiotypes. The resulting model requires further parameters and as a consequence leads to a vast phase space. We have focused on a particular niche in which it is possible to observe the trade-offs involved in the quasispecies' survival or destruction. Within this niche, we give simulation results that highlight some
Quasispecies dynamics on a network of interacting genotypes and idiotypes: formulation of the model

key preconditions for quasispecies survival. These include a minimum initial abundance of genotypes relative to that of the idiotypes and a minimum value of $p$. The latter, in particular, is to be contrasted with the stand-alone quasispecies network of our previous work, in which arbitrarily low values of $p$ constitute a guarantee of quasispecies survival.

**Keywords:** co-evolution (theory), mutational and evolutionary processes (theory), random graphs, networks

**ArXiv ePrint:** 1311.4857

---

1. Introduction

The immune system is one of the central regulatory systems of the body, being responsible for the detection and eventual removal of both external agents that may be potentially harmful and body cells whose behavior may have become abnormal. These targets of the immune system are generically referred to as antigens, in allusion to the fact that they trigger the system’s response and ultimate production of antibodies or activation of destroying cells. In order to be able to respond appropriately to a potentially large variety of antigens, the immune system is highly adaptive and along its existence evolves from an initial state of innate immunity through a series of states of acquired immunity. Modeling
the dynamics that gives rise to this type of learning has been challenging for many decades and has elicited the appearance of at least two main explanatory frameworks. One of them is the clonal-selection theory [1, 2], which postulates the preferential survival of those cell types that are more effective for a certain class of antigens. This theory has been successful in many respects but has failed in others, e.g. explaining the existence of any level of innate immunity, which by definition exists in the absence of any antigens.

The other leading framework to model the dynamics of immunity is that of the so-called idiotypic network [3]. The main idea is that the molecular structures capable of being recognized by the immune system, known as epitopes, are found not only in antigens but also in the receptors of the immune-system cells whose task is to recognize those antigens. That is, not only can these cells recognize antigens, they can also recognize one another in much the same way. Readily, the existence of such a network of epitope types, known as idiotypes, has the potential to explain not only how immunity is acquired but also how it can exist before antigens are ever encountered. The idiotypic-network theory has enjoyed both enthusiasm and skepticism along the years, the latter owing mainly to the many difficulties associated with confirming it experimentally. Many of its elements, however, are present in models of a hybrid nature, particularly in those that aim to characterize those phenomena, such as autoimmunity, that are essentially of a systemic nature (see e.g. [4–6] and references therein).

One common element of most mathematical models of the immune system is the overly simplistic manner in which the system’s interaction with antigens is handled. Typically the immune-system model involves a set of time-dependent equations describing the behavior of certain quantities (e.g. the abundance of a given cell type) and including independent terms to account for the presence of antigens. Such terms can be adjusted to account for the simultaneous presence of several antigens, but in general independently of one another. The drawback, of course, is that in important cases such as infections by some viral species, the many different viral strains that may be concomitantly present mutate into one another frequently and confront the immune system in ever-changing ways.

Here we address the problem of how antigens and entities of the immune system interact with one another when both sides already display complex dynamic behavior when left to themselves. We introduce a random-graph model of this interaction that takes into account not only the interaction itself but also the dynamics of genotype mutation on the antigen side and of idiotype recognition on the immune-system side. Our model can be viewed as comprising two subnetworks, one whose nodes represent genotypes that mutate into one another as they replicate, the other whose nodes represent idiotypes that stimulate (are recognized by) one another and proliferate as a consequence. The two subnetworks are put together by the addition of new edges to give idiotypes the further stimuli provided by the genotypes.

In building the two subnetworks we have drawn on previous ideas dealing with graph-theoretic representations of both interacting genotypes and idiotypes. On the genotype side the network we use is the same we introduced recently [7] to model the dynamics of the so-called quasispecies. A quasispecies is the stationary state of a set of genotypes that mutate into one another while replicating without recombination based on fitnesses that do not depend on genotype abundance (see e.g. [8–12] and references therein). Our model is based on a single probability parameter, $p$, that regulates both graph connectivity and the occurrence of mutations and can also successfully account for the well-known...
transition from adaptation to degeneracy when mutations become too frequent. For the idiotype side we use a similar network while incorporating some of the essential premises of the idiotypic-network theory. Like our quasispecies network, our idiotypic network is based on a single probability parameter, \( r \), whose function is to regulate both graph connectivity and the occurrence of stimulation/recognition of idiotypes by one another.

Each node in our model is a binary sequence, i.e. a sequence of 0’s and 1’s. This representation is suitable for both genotypes and idiotypes, since it can easily accommodate the types of interaction that are necessary in each subnetwork. Specifically, on the quasispecies side mutations are bit alterations and are more likely to occur between two genotypes that do not differ at too many loci. On the idiotypic side, by contrast, stimulation occurs as a function of how well the two sequences involved can be said to be complementary to each other on a locus-wise basis, being more likely as complementarity increases. This stems directly from the physicochemical nature of the molecular coupling that, in the immune system, leads to stimulation [5]. As the quasispecies and idiotypic networks are combined together into the full model, this same notion of complementarity is used to account for how genotypes stimulate idiotypes, again based on probability \( r \).

The resulting random graph, with its two parameters \( p \) and \( r \), is in principle flexible enough for any genotype to be able to mutate into any other genotype as well as stimulate any idiotype and for any idiotype to stimulate any other idiotype. So, although models have been proposed that incorporate the influence of the immune system on the dynamics of a quasispecies, ours is set apart from them, as they either ignore the fundamental network structure of the idiotypic side [13] or, while taking it into account, severely restrict the connection patterns that can be formed between the quasispecies and idiotypic sides [14].

In structural terms, therefore, our model generalizes both the model in [13] and the one in [14].

Before proceeding, we note that, by virtue of its underlying random-graph model and of the agent interactions this model represents, the subject we study in the present work is related to various other topics of interest. One of them is the wider discipline of complex networks, which in little more than a decade has uncovered many common underpinnings to a great variety of natural and technological systems of interacting agents [15–17]. Another of these topics is that of evolutionary games, which encompass the so-called predator-prey systems and, essentially, are characterized by the mediation of abundance-dependent fitnesses. The reader is referred to [18, 19], for example, as well as to references therein.

In a similar vein, note that the essence of our model is that it portrays the interaction of antagonistic agent populations as they adapt to each other while competing for supremacy. Thus, while in the present study we instantiate such agents by interrelated genotypes invading the body and idiotypes defending it, taking a broader view may come to benefit several other areas. One area with obvious conceptual ties to our invader-defender setting is that of computer security and in fact proposals have been put forward that are based on immune-related notions [20,21]. This has also been the case of other areas with less obvious but equally relevant connections to our study, such as combinatorial optimization (see [22] for an application to the prediction of protein structure), user profiling for adaptive information filtering in online systems [23] and classification of textual documents [24].
Quasispecies dynamics on a network of interacting genotypes and idiotypes: formulation of the model

![Figure 1. An instance of random graph $D$ for $L = 2$, with genotype set $A = \{a00, a01, a10, a11\}$ and idotype set $B = \{b00, b01, b10, b11\}$. Genotype $a00$ is the wild type. Solid edges are related to genotype mutation by similarity (section 2.1); dashed edges are related to idiotype stimulation by complementarity (sections 2.2 and 2.3). This instance has no self-loops on set $B$. The mandatory self-loops on set $A$ (see section 2.1) are not shown, nor are the mandatory edges entailed by full complementarity inside set $B$ (see section 2.2) or from set $A$ to set $B$ (see section 2.3). In-neighbor sets are $I_{a00} = \{a00, a01\}$, $I_{a01} = \{a01\}$, $I_{a10} = \{a01, a10\}$, $I_{a11} = \{a11\}$, $I_{b00} = \{a11, b01, b11\}$, $I_{b01} = \{a01, a10, b00, b10\}$, $I_{b10} = \{a01, b01\}$, and $I_{b11} = \{a00, a01, b00, b01\}$. Out-neighbor sets are $O_{a00} = \{a00, b11\}$, $O_{a01} = \{a00, a01, a10, b01, b10, b11\}$, $O_{a10} = \{a10, b01\}$, $O_{a11} = \{a11, b00\}$, $O_{b00} = \{b01, b11\}$, $O_{b01} = \{b00, b10, b11\}$, $O_{b10} = \{b01\}$, and $O_{b11} = \{b00\}$.](image)

Useful reviews on how these and other areas have come to be influenced by immune-inspired abstractions can be found in [25, 26].

The remainder of the paper is organized in the following manner. First we introduce the model’s details in section 2. Then we discuss some aspects of the computational methodology we have used and present our results, all in section 3. These results elucidate the role played by initial conditions, as well as by one of our main parameters, the probability $p$, in promoting the survival or the demise of the quasispecies as a function of how genotypes interact with one another and with the idiotypes. These issues are discussed in section 4, which is followed by conclusions in section 5.

2. Model

We consider $2^{L+1}$ binary sequences of length $L$ and group them into sets $A$ and $B$, each set comprising all $2^L$ distinct sequences of length $L$. Each sequence in set $A$ stands for a genotype, each sequence in set $B$ for an idotype. Our model is based on a directed random graph $D$ of node set $A \cup B$. For $i \in A \cup B$, we use $I_i$ to denote the set of in-neighbors of node $i$ and $O_i$ to denote its set of out-neighbors. Graph $D$ can have self-loops at all nodes, so it is possible that both $i \in I_i$ and $i \in O_i$. We describe the edge set of graph

doi:10.1088/1742-5468/2015/01/P01022
Quasispecies dynamics on a network of interacting genotypes and idiotypes: formulation of the model

\[ D \] in three separate stages: first the edges involving nodes in \( A \) exclusively, then edges involving nodes in \( B \) exclusively, then the edges used to interconnect the two halves. An example illustrating what is presented is provided in figure 1 for \( L = 2 \). Note that \( I_i \subseteq A \) for every genotype \( i \in A \) and \( O_i \subseteq B \) for every idiotype \( i \in B \). As for the sets \( O_i \) for \( i \in A \) and \( I_i \) for \( i \in B \), they may also intersect sets \( B \) and \( A \), respectively.

2.1. Quasispecies network

For \( i, j \in A \) (i.e. \( i \) and \( j \) are both genotypes), possibly identical, an edge exists directed from \( i \) to \( j \) with probability \( p_{ij} = p^{H_{ij}} \), where \( p \) is a probability parameter and \( H_{ij} \) is the Hamming distance between \( i \) and \( j \). In particular, this implies the existence of self-loops at all genotypes. If the edge does exist then it is possible for genotype \( i \) to mutate into genotype \( j \) during replication, provided \( i \neq j \). The probability that such mutation occurs is \( q_{ij} \), assumed proportional to \( p_{ij} \) in such a way that \( \sum_{j \in O_i \cap A} q_{ij} = 1 \), where the intersection \( O_i \cap A \) is meant to exclude any edges leading from genotype \( i \) to an idiotype (see section 2.3) and \( q_{ii} \) is the probability that genotype \( i \) remains unchanged during replication. Note that larger Hamming distances entail smaller connection probabilities and, consequently, smaller mutation probabilities as well.

Letting \( X_i \) denote the time-dependent abundance of genotype \( i \), we write

\[
\dot{X}_i = \sum_{j \in I_i} f_j q_{ji} X_j,
\]

where \( f_j \) is the fitness of genotype \( j \) and reflects its replication rate. We assume throughout that \( f_j = 2^{-d_j} \), where \( d_j \) is the number of 1’s in genotype \( j \). That is, a genotype’s fitness decays exponentially with its Hamming distance to the genotype having no 1’s (the fittest genotype, or wild type). Equation (1) is the well-known quasispecies equation [18, 27], now written for graph \( D \) as in the uniform case of [7].

2.2. Idiotypic network

For \( i, j \in B \) (i.e. \( i \) and \( j \) are both idiotypes), again possibly identical, an edge exists directed from \( i \) to \( j \) with probability \( r_{ij} = r^{L-H_{ij}} \), where \( r \) is another probability parameter. The existence of this edge indicates that idiotype \( i \) stimulates idiotype \( j \) during the proliferation phase of the idiotypic dynamics, provided \( i \neq j \). Note that edges between fully complementary idiotypes \( (H_{ij} = L) \) exist with probability 1. The probability that this stimulation occurs is \( s_{ij} \), assumed proportional to \( r_{ij} \) in such a way that \( \sum_{j \in O_i} s_{ij} = 1 \), where \( s_{ii} \) is the probability that idiotype \( i \) stimulates no other idiotype during proliferation (so long as a self-loop exists at idiotype \( i \)). Now larger Hamming distances entail larger connection and stimulation probabilities, which indeed are expected to be larger if the idiotypes involved are more complementary to each other (i.e. differ at more loci).

Letting \( X_i \) denote the time-dependent abundance of idiotype \( i \) and assuming that \( X_i \) grows in proportion by some rate \( \lambda > 0 \) to the total stimulus received by \( i \), yields

\[
\dot{X}_i = \lambda \sum_{j \in I_i \cap B} s_{ji} X_j,
\]

where the intersection \( I_i \cap B \) is meant to exclude, for the time being, any stimuli that do not originate from other idiotypes (see section 2.3 for a relaxation of this). This constitutes a very simple model of the idiotypic network in the absence of antigens.

doi:10.1088/1742-5468/2015/01/P01022
2.3. Genotype-idiotype interaction network

We are now in position to describe the entirety of graph $D$, of node set $A \cup B$, which is intended to model the interaction of the quasispecies and idiotypic networks, of node sets $A$ and $B$, respectively. We do this by allowing the existence of edges directed from nodes in $A$ to nodes in $B$, indicating that idiotypes are stimulated not only by one another but also by the genotypes against which they are supposed to constitute a defense. This, in effect, brings antigens into the idiotypic dynamics and lets graph $D$ work as a model of how invading genotypes and defending idiotypes interact with one another. These edges affect the $O_i$ sets for $i \in A$ and the $I_i$ sets for $i \in B$, thus justifying the intersection $O_i \cap A$ in the quasispecies-network constraint $\sum_{j \in O_i \cap A} q_{ij} = 1$ and the intersection $I_i \cap B$ in equation (2). It remains for us to specify how a genotype $i \in A$ stimulates an idiotype $j \in B$.

We begin by assuming that the length-$L$ sequences representing genotypes and idiotypes are all relative to the same support, i.e. they all share the same representation space, $\{0, 1\}^L$. In other words, the Hamming distance is well defined also between a genotype and an idiotype. Our approach is then to handle idiotype stimulation independently of whether it is effected by a genotype or another idiotype. That is, we let the edge from genotype $i \in A$ to idiotype $j \in B$ exist with the same probability $r_{ij}$ as above. This implies that there is always an edge from each genotype to its fully complementary idiotype. When an edge exists, stimulation occurs with the same probability $s_{ij}$ as above, so every node $i \in A$ is now subject to the additional constraint that $\sum_{j \in O_i \cap B} s_{ij} = 1$. Equation (2), therefore, becomes

$$\dot{X}_i = \lambda \sum_{j \in I_i} s_{ji} X_j,$$

where $i \in B$. Once coupled in this way, the $2^{L+1}$ differential equations given in equations (1) and (3) mandate an unbounded exponential growth of both genotype and idiotype abundances from any nontrivial initial values. We therefore lack further coupling in order for the possibility of genotype removal by the idiotypes to be explicitly taken into account. We achieve this by rewriting equation (1) as

$$\dot{X}_i = \sum_{j \in I_i} f_{ji} q_{ji} X_j - \mu \sum_{j \in O_i \cap B} s_{ij} X_j,$$

where $i \in A$ and $\mu > 0$ is a rate parameter. This modification to equation (1) lets the abundance of genotype $i$ be decreased at a rate that is proportional to how strongly $i$ stimulates each idiotype $j$.

It is important to note that, in equation (4), the $\mu$-dependent term that accounts for the removal of genotype $i$ by the combined action of the idiotypes does not directly depend on the genotype’s abundance, $X_i$. The reason for this has been detailed elsewhere in relation to a model of the immune system [4] and essentially has to do with the ‘inhibitory potential’ generated by the idiotypes being largely unaltered for most values of $X_i$. This, however, breaks down as $X_i$ approaches 0, thus requiring a slight correction to equation (4). We return to this issue in section 3.

doi:10.1088/1742-5468/2015/01/P01022
2.4. Relative abundances

Further modifications to all equations might still be considered for the removal of genotypes as they mutate into other genotypes or the removal of idiotypes as they stimulate (and are therefore recognized and sought for destruction by) other idiotypes. We approach this last step by henceforth considering relative, rather than absolute, abundances. Proceeding in this way is equivalent to imposing that the various abundances sum up to a constant at all times, which automatically leads to the appearance of terms that account for the desired sources of both genotype and idiotype removal.

For \( i \in A \cup B \), let \( x_i = \frac{X_i}{\sum_{k \in A \cup B} X_k} \) be the relative abundance of genotype or idiotype \( i \). Thus, \( \sum_{i \in A \cup B} x_i = 1 \) at all times. It follows that

\[
\dot{x}_i = \frac{\dot{X}_i}{\sum_{k \in A \cup B} X_k} - x_i \frac{\dot{X}_k}{\sum_{k \in A \cup B} X_k} = \frac{\dot{X}_i}{\sum_{k \in A \cup B} X_k} - x_i (\phi - \mu \psi + \lambda),
\]

(5)

where \( \phi = \sum_{k \in A} f_k x_k \) and \( \psi = \sum_{k \in B} x_k \sum_{l \in I_k \cap A} s_{lk} \). In this expression for \( \psi \), the rightmost summation represents the stimulatory influence of all genotypes upon idiotype \( k \) and is here referred to as that idiotype’s proliferability. Note, additionally, that \( \phi / \sum_{k \in A} x_k \) and \( \psi / \sum_{k \in B} x_k \) are, respectively, the average genotype fitness and idiotype proliferability.

Equation (5) yields the final expressions for \( \dot{x}_i \):

\[
\dot{x}_i = \sum_{j \in I_i} f_{ji} q_{ji} x_j - \mu \sum_{j \in O_i \cap B} s_{ij} x_j - x_i (\phi - \mu \psi + \lambda) \quad (6)
\]

for \( i \in A \) and

\[
\dot{x}_i = \lambda \sum_{j \in I_i} s_{ji} x_j - x_i (\phi - \mu \psi + \lambda) \quad (7)
\]

for \( i \in B \).

2.5. Missing edges

Being a random graph, \( D \) has the potential to connect any genotype out to any other genotype or to any idiotype and likewise to connect any idiotype out to any other idiotype. However, once \( D \) is instantiated and an actual graph is obtained, any of these edges may be missing from it. So, in essence, choosing a random graph as the basis for our model is a means to provide for the general situation in which some connections do exist while others do not. In the previous work introducing our quasispecies network [7], we argued for this generality on the grounds that it leads to a more plausible model simply by acknowledging that in many cases not every combination of loci can be simultaneously involved in a mutation.

That argument remains the same in the present work and in fact can be extended to justify the possibility that edges involving idiotypes may be absent as well. In fact, as we noted above, edges leading out from a genotype to an idiotype or connecting two different idiotypes stand for the possibility that two molecules bind to each other as a result of some complementarity in their physicochemical properties. Complementarity is thus necessary, but is far from being sufficient because the actual coupling at one or more
binding loci depends on many more factors, not least the spatial structure of the larger molecule in which the loci are contained (e.g. peptides comprising the several amino acids through which binding can actually take place) [5].

2.6. Expected connectivity

In graph $D$, a genotype $i$ may have out-neighbors both in set $A$ (other genotypes) and in set $B$ (idiotypes). If one ignores all self-loops, then the expected number of out-neighbors of the former type can be obtained by considering every other genotype $j$ and counting the edge from $i$ to $j$ with weight $p_{ij}^H$ [28]. That is, a genotype’s expected number of out-neighbors that are genotypes other than itself is

$$\sum_{h=1}^{L} \binom{L}{h} p^h = (1 + p)^L - 1. \tag{8}$$

As for a genotype’s expected number of out-neighbors that are idiotypes, we have

$$\sum_{h=0}^{L} \binom{L}{h} r^{L-h} = (1 + r)^L. \tag{9}$$

An idiotype, on the other hand, can only have other idiotypes as out-neighbors. Its expected number of out-neighbors that are also idiotypes, ignoring itself, is

$$\sum_{h=1}^{L} \binom{L}{h} r^{L-h} = (1 + r)^L - r^L. \tag{10}$$

3. Results

We begin by revisiting equation (6) and recognizing that it is possible for $\dot{x}_i$ to be negative when $x_i = 0$, thus leading the implicit constraint that $x_i \geq 0$ for all $i \in A \cup B$ at all times to be violated. We fix this by noting that it suffices that this equation’s predecessor for absolute abundances, equation (4), be modified to

$$\dot{X}_i = \sum_{j \in I_i} f_j q_{ji} X_j - \mu H(X_i) \sum_{j \in O_i \cap B} s_{ij} X_j, \tag{11}$$

where $H(z)$ is the Heaviside step function, adapted to yield 1 if and only if $z > 0$ (yield 0, otherwise). This modification forbids $\dot{X}_i$ to be negative when $X_i = 0$ and affects both equations (6) and (7), since the expression for $\psi$ is itself affected.

The actual expressions we use are then as follows, where all appearances of the $H$ function have been shifted by a suitably small $\delta > 0$ to avoid numerical instabilities. For

$$\psi = \sum_{k \in B} x_k \sum_{\ell \in I_k \cap A} s_{\ell k} H(x_\ell - \delta), \tag{12}$$

we have

$$\dot{x}_i = \sum_{j \in I_i} f_j q_{ji} x_j - \mu H(x_i - \delta) \sum_{j \in O_i \cap B} s_{ij} x_j - x_i(\phi - \mu \psi + \lambda) \tag{13}$$

doi:10.1088/1742-5468/2015/01/P01022
for $i \in A$ and
\[ \dot{x}_i = \lambda \sum_{j \in I_i} s_{ji} x_j - x_i (\phi - \mu \psi + \lambda) \] (14)
for $i \in B$.

While equations (13) and (14) make it difficult for the instantaneous value of $\dot{x}_i$ to be negative when $x_i \leq \delta$, this can still happen. Thus, in order to ensure that every $x_i$ is always nonnegative, further control is required when actually solving the equations. Specifically, when the value of any given $x_i$ is to be updated, say at time $t$, one must aim to ensure that the time step to be used, $\Delta t$, entails $x_i + \dot{x}_i \Delta t \geq \delta$ whenever possible. That is, for each $i \in A \cup B$ one must first solve $x_i + \dot{x}_i (t_i - t) = \delta$ for $t_i$ and then choose $\Delta t$ to be the least value of $t_i - t$ for which $t_i > t$. What we do is solve $x_i + \dot{x}_i (t_i - t) = \delta/10$ instead (again, to avoid numerical instabilities), using $\delta = 10^{-10}$ throughout.

We give results for $L = 10$ (i.e. for 1 024 genotypes and 1 024 idiotypes) in figures 2–8. These results reflect our exploration of a specific parameter niche in which it is possible to observe a rich variety of behaviors for the entire quasispecies (figures 2–5) and, particularly, for the wild type (figures 6–8). We follow [7] and let $x_1$ denote the relative abundance of the wild type. Moreover, we let the relative abundance of genotypes be denoted by $x_A$, i.e.
\[ x_A = \sum_{i \in A} x_i, \] (15)
We use $x_A(0)$ to denote the initial value of $x_A$.

All figures give results based on at least $10^4$ instances of graph $D$. Solving equations (13) and (14) for each instance starts at uniform relative abundances for genotypes and likewise for idiotypes. That is, initially $x_i = x_A(0)/2^L$ for $i \in A$ and $x_i = [1 - x_A(0)]/2^L$ for $i \in B$. All $2^{L+1}$ equations are then time-stepped through $t = 50$, which empirically we found to suffice for all relative abundances to reach a stationary state.

4. Discussion

The three plots in figure 2 show the probability density of the genotypes’ relative abundance, $x_A$, as $t$ is varied from $t = 0$ through $t = 30$. All cases correspond to $p = r = 0.1$, meaning that on average both the genotype and the idiotype subnetworks have densities of the same order of magnitude and so does the set of edges directed from the genotype subnetwork to the idiotype subnetwork to account for the removal of genotypes by idiotypes (see equations (8)–(10)). It also means that the dynamics of genotype mutation and of idiotype stimulation are based on the same underlying probability. The three cases also have in common that $\lambda = \mu = 0.1$, so the rate at which idiotypes proliferate due to stimulation by other idiotypes and the rate at which genotypes are eliminated as they stimulate idiotypes are the same. What distinguishes the three cases from one another is the initial relative abundance of genotypes, $x_A(0)$, which is varied from 0.14 in panel (a) through 0.16 in panel (c).

These three values of $x_A(0)$ were singled out, despite being so close to one another, because they allow us to qualitatively zoom in on what appears to be a transition from
Quasispecies dynamics on a network of interacting genotypes and idiotypes: formulation of the model

Figure 2. Evolution of the probability density of $x_A$ from $x_A(0) = 0.14$ (a), $x_A(0) = 0.15$ (b) and $x_A(0) = 0.16$ (c), for $p = r = 0.1$ and $\lambda = \mu = 0.1$. Data are based on at least $10^4$ independent instances of graph $D$ and are log-binned to the base 1.2. Probability densities are given according to the color-coded logarithmic scale on the right of each panel, ranging from $10^{-3}$ (at the bottom of the scale) to $10^4$ (at the top).
a regime in which the genotypes may either disappear or endure in the long run (i.e. the quasispecies may perish or survive) to another in which they endure almost certainly. In fact, examining the three panels of figure 2 reveals that, even though in all three cases the genotypes are driven toward a sharp decrease in relative abundance up to about $t = 2$, regardless of the particular instance of graph $D$ on which the dynamics is taking place, thereafter network structure begins to matter and does so in a manner that depends on what the relative genotype abundance was to begin with.

For the lowest of the three values used in the figure ($x_A(0) = 0.14$) and notwithstanding the fact that there exist network topologies continuing to push the genotypes toward the demise of the quasispecies, there are also cases in which the underlying network topology supports the survival of the quasispecies to the point that it seriously threatens the idiotype population, i.e. $x_A$ approaches 1. This turn of events becomes significantly more pronounced as $x_A(0)$ is slightly increased, quickly reaching the situation in which almost no network topology supports the destruction of the quasispecies (for $x_A(0) = 0.16$).

Interestingly, a similar situation occurs when $x_A(0)$ remains fixed, along with $p$ and $r$, while $\lambda$ and $\mu$ are made to vary relative to each other. This is shown in figure 3, in whose panels we have $p = r = 0.1$ and $x_A(0) = 0.1$ but $\lambda$ and $\mu$ are given values so that the ratio $\lambda/\mu$ increases by one order of magnitude at each new combination, from $\lambda/\mu = 10^{-1}$ to $\lambda/\mu = 10$ ($\lambda = 0.01$ and $\mu = 0.1$ in panel (a), $\lambda = \mu = 0.1$ in panel (b), $\lambda = 0.1$ and $\mu = 0.01$ in panel (c)). Moving from the scenario of panel (a) toward the one in panel (c) reveals an initial situation of uncertain survival of the quasispecies and a final situation in which it almost certainly survives. The intermediate scenario (shown in panel (b)), in particular, is to be compared with the ones in figure 2. This scenario differs from the ones in that figure simply in that the value of $x_A(0)$ is lower. What we see is the same uncertainty that is present in panels (a) and (b) of figure 2 regarding the eventual survival of the quasispecies, but now those network topologies for which the quasispecies perishes are such that this happens even earlier in the evolution.

The ratio $\lambda/\mu$ indicates how much more responsive the immune system is in reorganizing itself to confront the threat of the genotypes than it is in actually destroying genotypes. What we see in figure 3 is that a heavy imbalance toward the former almost certainly leads to dominance by the attacking quasispecies. An imbalance in the other direction makes quasispecies survival dubious but still leaves room for it to happen, thus raising the issue of how to set $\lambda$ and $\mu$ for the immune system to prevail almost certainly. Simply setting $\lambda = \mu = 0.1$ is not enough, as shown in panel (b) of figure 3, but increasing both $\lambda$ and $\mu$ does suffice. This is illustrated in figure 4, with $\lambda = \mu = 0.15$ and all other settings as in figure 3.

Throughout the transition depicted in figures 2 and 3, the bifurcation that the genotype population undergoes at about $t = 2$ clearly depends on the particular instance of graph $D$ being used. We study the role of network topology on the quasispecies by initially continuing to focus on the $p = r = 0.1$ setting that is common to figures 2–4. In this setting, the stationary-state distribution of a genotype’s number of out-neighbors is the one shown in figure 5, where only out-neighbors in set $A$ are counted (i.e. out-neighbors that are genotypes as well, so only the value of $p$ matters) and self-loops are ignored (so that 0 is a possibility). This distribution is clearly concentrated on the lowest numbers of out-neighbors and in this range it seems to be possible to approximate it by the Poisson distribution of mean given by equation (8). However, the Poisson distribution
Figure 3. Evolution of the probability density of $x_A$ from $x_A(0) = 0.1$ for $p = r = 0.1$. Panel (a) has $\lambda = 0.01$ with $\mu = 0.1$; panel (b) has $\lambda = \mu = 0.1$, as in figure 2; panel (c) has $\lambda = 0.1$ with $\mu = 0.01$. Note that the ratio $\lambda/\mu$ increases by one order of magnitude from panel (a) to panel (b) and from panel (b) to panel (c). Data are based on at least $10^4$ independent instances of graph $D$ and are log-binned to the base 1.2. Probability densities are given according to the color-coded logarithmic scale on the right of each panel, ranging from $10^{-3}$ (at the bottom of the scale) to $10^4$ (at the top).
Quasispecies dynamics on a network of interacting genotypes and idiotypes: formulation of the model

Figure 4. Evolution of the probability density of $x_A$ from $x_A(0) = 0.1$ for $p = r = 0.1$, as in figure 3 and $\lambda = \mu = 0.15$. Data are based on at least $10^4$ independent instances of graph $D$ and are log-binned to the base 1.2. Probability densities are given according to the color-coded logarithmic scale on the right of each panel, ranging from $10^{-3}$ (at the bottom of the scale) to $10^4$ (at the top).

Figure 5. Probability distribution of a genotype’s number of out-neighbors in set $A$, ignoring self-loops. Data are given for $p = 0.1$ and are based on $10^4$ independent instances of graph $D$. As defined, the probability that an edge exists between two genotypes in $D$ depends on the Hamming distance between them. If this probability were independent of the two genotypes, the resulting distribution would be the Poisson distribution, here shown as a solid line for the same mean of 1.59 that is given for $p = 0.1$ by equation (8). If $z$ is the expected number of out-neighbors given by this equation, then the Poisson distribution of mean $z$ requires that edges exist with the fixed probability $z/(2^L - 1) = [(1 + p)^L - 1]/(2^L - 1)$, which approaches the probabilities we use only as $p$ nears 1.

arises only when edges exist with the same probability for all node pairs (as in the original Erdős–Rényi model [29] and its directed variant [28]) and in fact we see in the figure that the two distributions differ markedly for the higher numbers of out-neighbors.

doi:10.1088/1742-5468/2015/01/P01022
Quasispecies dynamics on a network of interacting genotypes and idiotypes: formulation of the model

Figure 6. Stationary-state relative abundance of the wild type ($x_1$) as a function of its number of out-neighbors in set $A$, ignoring the self-loop. Data are given for $p = r = 0.1$ and $\lambda = \mu = 0.1$ and are based on $10^5$ independent instances of graph $D$.

The distribution shown in figure 5 indicates that a randomly chosen genotype is likely to be able to mutate only into a small number of other genotypes. If idiotypes were altogether absent this would entail excellent survival chances for the quasispecies, meaning in particular that the wild type would be able to climb from its initial relative abundance of $1/2^L \approx 10^{-3}$ toward a stationary-state relative abundance of the order of $10^{-1}$ [7]. In the presence of idiotypes, however, one must look at how network topology affects the quasispecies from a closer perspective, paying special attention to how the wild type’s number of out-neighbors affects its stationary-state relative abundance, $x_1$.

Figure 6 shows the average value of $x_1$ as a function of the wild type’s number of out-neighbors in set $A$ and ignoring the self-loop. All data in this figure are still relative to $p = r = 0.1$ and $\lambda = \mu = 0.1$. Each plot corresponds to a different value of the initial genotype population, beginning at $x_A(0) = 0.095$ and proceeding through $x_A(0) = 0.15$ (i.e. up to inside the transition qualitatively depicted in figure 2). In those instances of graph $D$ in which the wild type has no out-neighbors, the value of $x_1$ is influenced from outside exclusively by mutations into the wild type and by the action of the idiotypes. Under these circumstances, it is clear from the figure that the wild type recovers from its initial situation of uniform dilution with respect to the other genotypes, i.e. from an initial relative abundance of $x_A(0)/2^L \approx 10^{-4}$. Instances with larger numbers of out-neighbors at the wild type, on the other hand, reflect the effect of mutations of the wild type into other genotypes as well, leading to progressively smaller values of $x_1$ until a precipitous drop to some value between $\delta/10$ and $\delta$ is reached (see section 3), indicating the total absence of $D$ instances in which the wild type has any of the corresponding numbers out-neighbors. Some of this figure’s plots are repeated in figure 7 alongside new plots for the same values of $x_A(0)$ but ten times as many instances of graph $D$, which postpones the aforementioned drop as $D$ instances appear in which the wild type has a larger number of out-neighbors. Taken together, the two figures seem to support the presence of an exponential decay of $x_1$ with the wild type’s number of out-neighbors. However, as already suggested by the

doi:10.1088/1742-5468/2015/01/P01022
data in figure 2, such decay is ever slower as \( x_A(0) \) increases, so the wild type is ever more resilient to the possibility of mutation into a larger number of genotypes.

Varying the value of \( p \) and \( r \) relative to each other can provide further insight into wild-type survival, since these are the parameters regulating edge density and also mutation or stimulation on those edges that do exist in a particular instance of graph \( D \). We give data for this in figure 8, which continues to be relative to \( \lambda = \mu = 0.1 \) but now focuses exclusively on the \( x_A(0) = 0.1 \) case of figures 6 and 7. Note, first, that using \( p = 0.1 \) with \( r = 0.01 \) has no relevant effect with respect to the \( p = r = 0.1 \) case. The reasons for this can be grasped from equations (9) and (10): according to these equations, decreasing \( r \) from 0.1 to 0.01, though entailing a difference by one order of magnitude in the value of this probability, impacts the expected number of out-neighbors a genotype or an idotype has inside set \( B \) much more modestly. In fact, these variations are in both cases from about 2.59 to about 1.01 and seem to affect genotype-idiotype interaction very little.

On the other hand, using \( p = 0.01 \) with \( r = 0.1 \), while affecting only the quasispecies side of the network in comparison to the \( p = r = 0.1 \) case, leads to very different results. By equation (8), now a genotype’s expected number of out-neighbors in set \( A \) drops from 1.59 to 0.1, therefore by one order of magnitude as well. Genotypes are then very sparsely interconnected to one another, but in spite of this we see in figure 8 that the wild type manages to survive for those instances of graph \( D \) in which it has a very small number of out-neighbors. The pattern of survival is significantly different from the previous one, which may at first seem as a surprise given that the pull of idiotypes on genotypes continues to be the same and all that has changed is the density in genotype interconnection and the probabilities that mutations occur. However, the resulting expected sparsity does not rule out those relatively rare \( D \) instances in which the wild type is fed by mutations from at least one other genotype. Combined with the very low number of genotypes into which the wild type itself may mutate, this unexpected existence of edges incoming to the wild type ensures survival.
Quasispecies dynamics on a network of interacting genotypes and idiotypes: formulation of the model

Figure 8. Stationary-state relative abundance of the wild type ($x_1$) as a function of its number of out-neighbors in set $A$, ignoring the self-loop. Data are given for $\lambda = \mu = 0.1$ and $x_A(0) = 0.1$ and are based on $10^5$ independent instances of graph $D$ when $p = 0.1$, $10^6$ instances when $p = 0.01$.

It follows from these analyses of the data in figure 8 that, as in the case of the isolated quasispecies network [7], probability $p$ is instrumental in determining the fate of the wild type. Contrasting with that case, however, the wild type is no longer guaranteed to survive for arbitrarily low values of $p$. Instead, it seems that some minimum value is required for the quasispecies to resist the action of the idiotypes, i.e. for the wild type to escape being diluted into the remainder of the quasispecies.

Of course, the above remarks on the relative impact of the $p$ and $r$ probabilities must always be interpreted within the constraints of the number $L$ of loci. By equations (8)–(10), all three expected numbers of out-neighbors become the same for genotypes and idiotypes alike in the limit of large $L$, provided only that $p = r$. That is, in these circumstances a genotype is expected to have as many out-neighbors in set $A$ as it does in set $B$ and this number is the same as the expected number of out-neighbors of an idiotype. The three numbers become roughly the same for smaller values of $L$ if $p$ and $r$ are relatively high (e.g. $p = r = 0.1$) and for larger values of $L$ if $p$ and $r$ are relatively low (e.g. $p = r = 0.01$). For $L = 70$ (which, e.g., seems to be well within the range of values reported for how many binding loci the influenza A/H3N2 virus has [30]), the three expected numbers of out-neighbors are about the same for $p = r = 0.1$ ($\approx 790$) but do differ for $p = r = 0.01$, though within the same order of magnitude (1 for a genotype’s out-neighbors in $A$, 2 for a genotype’s or an idiotype’s out-neighbors in $B$). Thus, while using $p = 0.01$ with $r = 0.1$ impacts a genotype’s number of other genotypes it connects out to similarly to the $L = 10$ case (a reduction by at least one order of magnitude), using $p = 0.1$ with $r = 0.01$ is now markedly different in regard to the number of edges involving idiotypes (as it gets reduced by orders of magnitude as well).

For larger values of $L$, therefore, we should expect both probability-reduction scenarios to have an impact on wild-type survival. In fact, it seems reasonable to expect both scenarios to make it harder for the wild type to survive, since reducing $p$ thins out the inter-genotype connections and reducing $r$ has a similar effect on the connections that...
involve idiotypes. This corroborates the conclusion drawn above for $L = 10$ regarding the role of $p$, but further calculations seem necessary for our expectation regarding $r$ to be confirmed. These, regrettably, seem computationally infeasible at this time.

5. Conclusion

Quasispecies studies have invariably concentrated on characterizing two distinct regimes for the evolution of genotypes through time, one in which the wild type survives in the quasispecies, another in which the wild type becomes diluted and no more abundant than any of the quasispecies’ other genotypes. Distinguishing between the two regimes has been a matter of selecting the right perspective from which to model replication errors, or mutations. In our own previous model, for example, the single parameter $p$ can be used to characterize quasispecies survival or demise: Increasing $p$ progressively leads the wild type to a situation in which its relative abundance cannot rise above those of the other, less fit genotypes [7]. This focus on the two extremal regimes of survival and demise has been motivated by the theory’s purported use in the modeling of viral populations, along with its appeal as a potential aid in the discovery of therapies [10, 11].

However, in our view the picture has clearly been incomplete, particularly when viral pathogens are the motivation, because the quasispecies’ interaction with the host’s immune system seems to have been left completely aside. The present work constitutes an attempt to take the immune system into account and study its effect on quasispecies behavior. Our model uses the same network as [7] to represent the interacting genotypes and interconnects it to another network built on top of the so-called idiotypes, some of the fundamental motifs of the host’s immune response. Like the quasispecies network, this new, idiotypic network is self-adapting (based on the idea of complementarity among idiotypes) and potentially effective in destroying genotypes (based on the idea that genotypes and idiotypes are mutually complementary to some degree).

The resulting network is a continuous-time dynamical system whose variables, representing the relative abundances of genotypes and idiotypes, are coupled with one another according to a random-graph model. The system’s behavior depends on four parameters (the probabilities $p$ and $r$ and the rates $\lambda$ and $\mu$), as well as on the initial relative abundance of genotypes, $x_A(0)$, which together give rise to a phase space of vast proportions. In view of this, we have concentrated on analyzing a specific niche inside which both survival and destruction of the quasispecies can be examined side by side.

Our exploration of this particular niche has highlighted the existence of two main factors influencing quasispecies survival. The first one is $x_A(0)$ itself, or equivalently the interplay between $\lambda$ and $\mu$. If genotypes are reasonably abundant in the beginning with respect to idiotypes, or if idiotypes eliminate genotypes much more slowly than the idiotypes adapt their abundances to meet the challenge of the mutating genotypes (i.e. $\mu \ll \lambda$), then it is possible for the quasispecies to survive. The second factor is the topology of the quasispecies network, governed by the $p$ parameter. Unlike the case of the isolated quasispecies network, in which very low values of $p$ are practically a guarantee that the wild type will nearly dominate the quasispecies even from an initial situation of total dilution, the presence of the idiotypic network makes things quite different. Specifically, a non-negligible value of $p$ seems necessary to ensure that the quasispecies network is sufficiently
dense and mutations sufficiently likely, for genotypes to change into one another and help the wild type evade the action of the idiotypic network.

Further research will concentrate on revisiting equations (3) and (4) with the goal of modeling the action on the immune system of those quasispecies, such as that of HIV, that work toward depleting the system’s supply of idiotypes. While the required changes may turn out to be conceptually simple, it seems unavoidable that at least one new rate parameter will be needed, thus enlarging the model’s phase space even further. Finding an appropriate parameter niche to explore, however, may provide interesting insight into how genotype mutation, idiotype stimulation and idiotype destruction affect one another.

Acknowledgments

We acknowledge partial support from CNPq, CAPES, a FAPERJ BBP grant and the joint PRONEX initiative of CNPq/FAPERJ under contract No. 26-111.443/2010.

References

[1] Burnet F M 1959 The Clonal Selection Theory of Acquired Immunity (Cambridge: Cambridge University Press)
[2] Forsdyke D R 1995 FASEB J. 9 164–6
[3] Jerne N K 1974 Ann. Inst. Pasteur Imm. C125 373–89
[4] Flores L E, Aguilar E J, Barbosa V C and Carvalho I A V 2004 J. Theor. Biol. 229 311–25
[5] Behn U 2007 Immunol. Rev. 216 142–52
[6] Madi A, Kenett D Y, Bransburg-Zabary S, Merbl Y, Quintana F J, Tauber A I, Cohen I R and Ben-Jacob E 2011 PLoS One 6 e17445
[7] Barbosa V C, Donangelo R and Souza S R 2012 J. Theor. Biol. 312 114–9
[8] Eigen M 1971 Naturwissenschaften 58 465–523
[9] Eigen M and Schuster P 1977 Naturwissenschaften 64 541–65
[10] Domingo E 2009 Contrib. Sci. 5 161–8
[11] Lauring A S and Andino R 2010 PLoS Pathog. 6 e1001005
[12] Más A, López-Galíndez C, Cacho I, Gómez J and Martínez M A 2010 J. Mol. Biol. 397 865–77
[13] Woo H J and Reifman J 2012 Proc. Natl Acad. Sci. USA 109 12980–5
[14] Kamp C and Bornholdt S 2002 Phys. Rev. Lett. 88 068104
[15] Bornholdt S and Schuster H G (ed) 2003 Handbook of Graphs and Networks (Weinheim: Wiley)
[16] Newman M, Barabási A L and Watts D J (ed) 2006 The Structure and Dynamics of Networks (Princeton, NJ: Princeton University Press)
[17] Bollobás B, Kozma R and Miklós D (ed) 2009 Handbook of Large-Scale Random Networks (Berlin: Springer)
[18] Nowak M A 2006 Evolutionary Dynamics (Cambridge, MA: Harvard University Press)
[19] Knebel J, Krüger T, Weber M F and Frey E 2013 Phys. Rev. Lett. 110 168106
[20] Harmer P K, Williams P D, Gunsch G H and Lamont G B 2002 IEEE Trans. Evol. Comput. 6 252–80
[21] Drozda M, Schaub S and Schüdt S 2011 Nat. Comput. 10 243–74
[22] Cutello V, Nicosia G, Pavone M and Timmis J 2007 IEEE Trans. Evol. Comput. 11 101–17
[23] Nanas N, Vavalis M and De Roeck A 2010 Swarm Intell. 4 275–300
[24] Abi-Haidar A and Rocha L M 2011 Evol. Intell. 4 69–80
[25] Garrett S M 2005 Evol. Comput. 13 145–78
[26] Timmis J, Hone A, Stibor T and Clark E 2008 Theor. Comput. Sci. 403 11–32
[27] Biebricher C K and Eigen M 2006 What is a quasispecies? Quasispecies: Concept and Implications for Virology (Current Topics in Microbiology and Immunology vol 299) ed E Domingo (Berlin: Springer) pp 1–31
[28] Karp R M 1990 Random Struct. Alg. 1 73–93
[29] Erdős P and Rényi A 1959 Publ. Math. (Debrecen) 6 290–7
[30] Bui H H, Peters B, Assarsson E, Mbawuike I and Sette A 2007 Proc. Natl Acad. Sci. USA 104 246–51
doi:10.1088/1742-5468/2015/01/P01022