Journal of Theoretical Biology 394 (2016) 18–31

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Evolutionary escape on complex genotype–phenotype networks

Esther Ibáñez-Marcelo [a,b,*], Tomás Alarcón [e,a,c,d]

[a] Centre de Recerca Matemàtica, Edifici C, Campus de Bellaterra, 08193 Bellaterra (Barcelona), Spain
[b] Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, 08028 (Barcelona), Spain
[c] Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain
[d] Barcelona Graduate School of Mathematics (BGSMath), (Barcelona), Spain
[e] ICREA (Institució Catalana de Recerca i Estudis Avançats), Spain

HIGHLIGHTS

- We perform a comparative analysis between escape on regular hypercube (*H*-graphs) and complex genotype–phenotype spaces (*B*-graphs).
- We find that the distribution of distances between phenotypes in *B*-graphs exhibits a much larger degree of heterogeneity than in *H*-graphs.
- This property causes heterogeneous behaviour in all results associated to the escape problem.
- Our main result is that the escape probability can be underestimated by assuming a regular hypercube genotype network.

ARTICLE INFO

Article history:
Received 25 March 2015
Received in revised form
22 December 2015
Accepted 25 December 2015
Available online 21 January 2016

Keywords:
Escape
Genotype–phenotype map
Resistance

ABSTRACT

We study the problem of evolutionary escape that is the process whereby a population under sudden changes in the selective pressures acting upon it try to evade extinction by evolving from previously well-adapted phenotypes to those that are favoured by the new selective pressure. We perform a comparative analysis between results obtained by modelling genotype space as a regular hypercube (*H*-graphs), which is the scenario considered in previous work on the subject, to those corresponding to a complex genotype–phenotype network (*B*-graphs). In order to analyse the properties of the escape process on both these graphs, we apply a general theory based on multi-type branching processes to compute the evolutionary dynamics and probability of escape. We show that the distribution of distances between phenotypes in *B*-graphs exhibits a much larger degree of heterogeneity than in *H*-graphs. This property, one of the main structural differences between both types of graphs, causes heterogeneous behaviour in all results associated to the escape problem. We further show that, due to the heterogeneity characterising escape on *B*-graphs, escape probability can be underestimated by assuming a regular hypercube genotype network, even if we compare phenotypes at the same distance in *H*-graphs. Similarly, it appears that the complex structure of *B*-graphs slows down the rate of escape.

1. Introduction

Evolutionary escape is the process whereby a population under sudden changes in the selective pressures acting upon it try to evade extinction by evolving from previously well-adapted phenotypes to those that are favoured by the new selective pressure. This evolutionary process is driven by gene mutations. Examples of biological situations where this process is relevant include viruses evading anti-microbial therapy, emergence of drug resistance in cancer, parasites trying to infect a new host, or species attempting to invade a new ecological niche (Iwasa et al., 2003).

Previous models of evolutionary escape have been developed by Iwasa et al. (2003), Iwasa et al. (2004). They base their approach on the assumption that n point mutations in some crucial parts of the genome are necessary for escape. They further assume that the different mutants can be described by binary strings (with entries of $+1$ or -1) of length n. There are 2^n-1 such mutants. It is assumed that the new selective pressures, such as those associated to the administration of a drug, reduce the proliferation ratios, R, of sensitive genotypic variants, $R<1$; whereas resistant genotypes are such that $R>1$. The corresponding evolutionary dynamics is modelled in terms of Galton–Watson multi-type branching process (GWMBP) (Kimmel and Axelrod, 2002), where at each generation each individual of each type has a given (in general, genotype-dependent) probability of mutating and

* Corresponding author at: Centre de Recerca Matemàtica, Edifici C, Campus de Bellaterra, 08193 Bellaterra (Barcelona), Spain. Tel.: +34 93 581 1081.
E-mail address: eibanez@crm.cat (E. Ibáñez-Marcelo).

http://dx.doi.org/10.1016/j.jtbi.2015.12.033
0022-5193/

producing offspring belonging to a different type. The problem is to calculate the probability that an escape genotype is reached starting from a single individual of non-resistant genotype. The model proposed by Iwasa et al. has been analysed in more detail by Serra and coworkers (Serra, 2006; Serra and Haccou, 2007; Sagitov and Serra, 2009). These authors have thus considered the process of evolutionary escape as a random search on a genotype space modelled by a hypercube: individuals concentrate in a given non-escape genotype and they must reach a well-adapted genotype (the so-called escape genotype) before the population undergoes extinction. An alternative escape mechanism has been proposed in Alarcon and Jensen (2010) whereby escape is achieved by means of a growth-restricted (quiescent) phenotype that is insensitive to the selective pressure (e.g. a drug). This escape mechanism is relevant in cancer treatment of hypoxic tumours (Alarcon et al., 2005; Brikci et al., 2008; Bristow and Hill, 2008) and drug resistance in bacterial populations which exhibit persistence (Balaban et al., 2004; Lewis, 2007).

While the assumption that point mutations drive evolutionary escape is preserved in our study, we investigate whether considering the more realistic mechanism whereby selective pressures act on phenotypes rather than genotypes alters the properties of evolutionary escape in significant ways. To this end we describe evolutionary escape in terms of a population dynamic that accounts for the genotype–phenotype map. This modification alters the approach proposed by Iwasa et al. in two relevant aspects. First, due to evolved robustness in populations with genotype–phenotype map (Wagner, 1996; Ciliberti et al., 2007, 2007; Wagner, 2007, 2008, 2008; Jaeger and Monk, 2014), not every gene mutation necessarily generates a new phenotype. As a consequence, many gene mutations are neutral as far as the evolutionary escape process is concerned. Furthermore, it has been shown that the topology of genotype–phenotype networks is far from that of the hypercube lattice assumed by Aguirre et al. (2011) and Ibáñez-Marcelo and Alarcon (2014). In fact, we have recently shown that the corresponding phenotype network exhibits the small-world phenomenon and that, as a consequence, accelerated evolvability (relative to that of a system with no genotype–phenotype map) may emerge. The question naturally arises as to whether these properties, i.e. phenotypic robustness and evolvability typical of genotype–phenotype networks, have an influence on the process of evolutionary escape. To address this issue, we apply a general theory, based on multitype branching processes (Kimmel and Axelrod, 2002), to compute the evolutionary dynamics and probabilities of escape which takes into account the structure of the genotype–phenotype space.

The structure and properties of systems with genotype–phenotype maps have been studied by considering several model systems (Wagner, 2012): RNA, circuits of gene regulation and metabolic networks. In the RNA model of the genotype–phenotype map (Fontana and Schuster, 1998), the *genotype* of each RNA molecule is its sequence of nucleotides. There are four such nucleotides, so for sequences of length L, the size of the genotype space is 4^L. The folded structure of an RNA sequence, which is a proxy for its *phenotype* (although it still lies far from defining its function), is determined by the sequence (genotype) in a many-to-one way, i.e. many different genotypes have the same associated phenotype. Such non-uniqueness has led to the concept of the neutral network, first introduced in Lipman and Wilbur (1991) and Schuster et al. (1994), which is defined as the network whose nodes correspond to genotypes, all with the same phenotypes, with edges between those nodes which differ by only one nucleotide (Wagner, 2007). This concept has been used extensively in the study of the genotype–phenotype map, in particular, those issues regarding its evolutionary properties, such as the role of phenotypic robustness in evolvability and adaptation (Wagner, 2011, 2012). Recently, the topology of the RNA genotype–phenotype space, composed by an intermingled set of neutral networks, has been analysed (Aguirre et al., 2011).

Gene regulatory networks (GRNs) have also been extensively used as models of the genotype–phenotype map, in particular several variants of the model of phenotype plasticity originally introduced by Wagner (1996). These models are dynamical systems for the expression levels of the corresponding genes and are characterised by two elements: a matrix whose entries specify the character of the interaction between two genes (usually, activation or inhibition) and, possibly, the intensity of such interaction, and a series of rules for the time evolution of the expression levels of the genes involved. The entries of the corresponding matrix are defined as the genotype of the GRN. The associated phenotype is the steady-state gene expression yielded by the dynamics. There are many such genotypes associated to the same phenotype, which allows to extend the concept of neutral network to GRNs, where nodes correspond to different matrices (producing the same steady-state) and links exist between nodes if the corresponding matrices differ only in one regulatory interaction. GRNs have been used to study phenotype plasticity (Wagner, 1996), robustness and innovation in circuits of gene regulation (Ciliberti et al., 2007, 2007), and canalisation (Siegal and Bergman, 2002; Bergman and Siegal, 2003), among other issues.

Metabolic networks are a third class of systems that have been used to assess properties regarding robustness and innovation (Ndifon et al., 2009; Rodrigues and Wagner, 2009; Samal et al., 2010; Rodrigues and Wagner, 2011). They are formed by thousands of enzyme-catalysed chemical reactions. These networks are responsible for supplying cells with energy (i.e. ATP) and the molecular building blocks that cells need to grow. The *genotype* space for this system consists of the space of all the possible metabolic networks, whereas the *phenotype* corresponds to the secondary metabolites the metabolic network is able to synthesise, the molecules they can use as energy sources, the ability to detoxify certain waste products, etc. (Wagner, 2012). Innovations in these aspects not always appear as the result of gene mutations that give rise to new enzymes. They can also arise through novel combinations and utilisation of existing elements.

Genotype–phenotype networks have been extensively used and analysed in RNA and GRN models of the genotype–phenotype map. These are networks whose set of nodes corresponds to the set of (viable) genotypes (e.g. each nucleotide sequence that yields a properly folded molecule). A link between a pair of genotypes exist if they are separated by one single mutation (e.g. an RNA molecule is linked to all its (viable) single-point mutants). The ensemble of genotypes associated to the same phenotype form its neutral network. In Ibáñez-Marcelo and Alarcon (2014), we added to this picture the so-called phenotype nodes which correspond to the set of (viable) phenotypes. Links between genotype and phenotype nodes exist if the genotype node belongs to the neutral network of that particular phenotype.

Our aim is to extend the theory of evolutionary escape by analysing the effects on the probability of escape and the escape rate of considering that the evolutionary dynamics occurs on a complex genotype–phenotype network rather than on a regular hypercube. This paper is organised as follows. Section 2 is devoted to a detailed description of our model and a summary of the mathematical background involved in its analysis. A full account of the mathematical apparatus is included in the Supplementary Materials. In Section 3, we report the results of our analysis and our main findings. Finally, in Section 4, we present our conclusions and discuss our results as well as future directions for further research.

2. Mathematical model

The classical escape model (Iwasa et al., 2003, 2004; Serra, 2006; Serra and Haccou, 2007; Sagitov and Serra, 2009) can be summarised as follows. Each of the 2^n nodes of an n-dimensional hypercube is assumed to represent a genotype. Fitness values,