The distribution of the quasispecies for the Wright–Fisher model on the sharp peak landscape

Joseba Dalmau*

Université Paris Sud, France
ENS Paris, France

Abstract

We consider the classical Wright–Fisher model with mutation and selection. Mutations occur independently in each locus, and selection is performed according to the sharp peak landscape. In the asymptotic regime studied in Cerf (2014), a quasispecies is formed. We find explicitly the distribution of this quasispecies, which turns out to be the same distribution as for the Moran model.
© 2014 Elsevier B.V. All rights reserved.

Keywords: Quasispecies; Wright–Fisher model; Error threshold

1. Introduction

The concept of quasispecies first appeared in 1971, in Manfred Eigen's celebrated paper [7]. Eigen studied the evolution of a population of macromolecules, subject to both selection and mutation effects. The selection mechanism is coded in a fitness landscape; while many interesting landscapes might be considered, some have been given more attention than others. One of the most studied landscapes is the sharp peak landscape: one particular sequence – the master sequence – replicates faster than the rest, all the other sequences having the same replication rate. A major discovery made by Eigen is the existence of an error threshold for the mutation rate on the sharp peak landscape: there is a critical mutation rate q_c such that, if $q > q_c$ then the population evolves towards a disordered state, while if $q < q_c$ then the population evolves so as to form a quasispecies, i.e., a population consisting of a positive concentration of the master sequence, along with a cloud of mutants which highly resemble the master sequence.

* Correspondence to: DMA/ENS, 45 rue d'Ulm, 75005 Paris, France. Tel.: +33 788130124.
  E-mail address: dalmau@dma.ens.fr.

Eigen's model is a deterministic model, the population of macromolecules is considered to be infinite and the evolution of the concentrations of the different genotypes is driven by a system of differential equations. Therefore, when trying to apply the concepts of error threshold and quasispecies to other areas of biology (e.g. population genetics or virology), Eigen's model is not particularly well suited; a model for a finite population, which incorporates stochastic effects, is the most natural mathematical approach to the matter.

Several works have tackled the issue of creating a finite and stochastic version of Eigen's model [1,5,6,9–14]. Some of these works have recovered the error threshold phenomenon in the case of finite populations: Alves and Fontanari [1] find a relation between the error threshold and the population size by considering a finite version of Eigen's model on the sharp peak landscape. Demetrius, Schuster and Sigmund [5] generalise the error threshold criteria by modelling the evolution of a population via branching processes. Nowak and Schuster [12] also find the error threshold phenomenon in finite populations by making use of a birth and death chain. Some other works have tried to prove the validity of Eigen's model in finite populations by designing algorithms that give similar results to Eigen's theoretical calculations [9], while others have focused on proposing finite population models that converge to Eigen's model in the infinite population limit [6,11].

The Wright–Fisher model is one of the most classical models in mathematical evolutionary theory, it is also used to understand the evolution of DNA sequences. In [3], some counterparts of the results on Eigen's model were derived in the context of the Wright–Fisher model. The Wright–Fisher model describes the evolution of a population of m chromosomes of length ℓ over an alphabet with κ letters. Mutations occur independently at each locus with probability q. The sharp peak landscape is considered: the master sequence replicates at rate $\sigma > 1$, while all the other sequences replicate at rate 1. The following asymptotic regime is studied:

$$\ell \to +\infty, \qquad m \to +\infty, \qquad q \to 0,$$
$$\ell q \to a, \qquad \frac{m}{\ell} \to \alpha.$$

In this asymptotic regime the error threshold phenomenon present in Eigen's model is recovered, in the form of a critical curve $\alpha\psi(a) = \ln\kappa$ in the parameter space (a, α). If $\alpha\psi(a) < \ln\kappa$, then the equilibrium population is totally random, whereas a quasispecies is formed when $\alpha\psi(a) > \ln\kappa$. In the regime where a quasispecies is formed, the concentration of the master sequence in the equilibrium population is also found. The aim of this paper is to continue with the study of the Wright–Fisher model in the above asymptotic regime in order to find the distribution of the whole quasispecies. It turns out that the resulting distribution is the same as the one found for the Moran model in [4]. Nevertheless, the techniques we use to prove our result are very different from those of [4]. The study of the Moran model relied strongly on monotonicity arguments, and the result was proved inductively. The initial case and the inductive step boiled down to the study of birth and death Markov chains, for which explicit formulas could be found. The Wright–Fisher model is a model with no overlapping generations, for which this approach is no longer suitable. In order to find a more robust approach, we rely on the ideas developed by Freidlin and Wentzell to investigate random perturbations of dynamical systems [8], as well as some techniques already used in [3]. Our setting is essentially the same as the one in [3], the biggest difference being that we work in several dimensions instead of having one dimensional processes. The main challenge is therefore to extend the arguments from [3] to the multidimensional case. This is achieved by replacing the monotonicity arguments employed in [3] by uniform estimates.

We present the main result in the next section. The rest of the paper is devoted to the proof.