J. Appl. Prob. **53,** 606–613 (2016)
*doi:*10.1017/jpr.2016.25

DISTRIBUTION OF THE QUASISPECIES FOR A GALTON–WATSON PROCESS ON THE SHARP PEAK LANDSCAPE

JOSEBA DALMAU,* *Université Paris Sud and ENS Paris*

Abstract

We study a classical multitype Galton–Watson process with mutation and selection. The individuals are sequences of fixed length over a finite alphabet. On the sharp peak fitness landscape together with independent mutations per locus, we show that, as the length of the sequences goes to ∞ and the mutation probability goes to 0, the asymptotic relative frequency of the sequences differing on k digits from the master sequence approaches $(\sigma\mathrm{e}^{-a}-1)(a^k/k!)\sum_{i\geq 1} i^k/\sigma^i$, where σ is the selective advantage of the master sequence and a is the product of the length of the chains with the mutation probability. The probability distribution $\mathcal{Q}(\sigma, a)$ on the nonnegative integers given by the above equation is the quasispecies distribution with parameters σ and a.

Keywords: Quasispecies; Galton–Watson process

2010 Mathematics Subject Classification: Primary 60J80
Secondary 92D25

1. Introduction

Understanding the origin of life on Earth has always been a major objective of science. One of the many important contributions to the subject is the 1971 paper by Eigen [7], which theoretically explored the early stages of life on our planet. As an explanation of how evolution of very simple macromolecules might have occurred, Eigen proposed a model known today as Eigen's model or the quasispecies model. The model aims at describing the evolution of a population of macromolecules, driven both by selection and mutation forces. Due to the simplicity of the reproducing individuals, Eigen's model can be synthesised as a system of differential equations, obtained from the chemico-kinetic reactions the different macromolecules are subject to:

$$x_k'(t)=\sum_{i=1}^{N} f(i)Q(i,k)x_i(t)-x_k(t)\sum_{i=1}^{N} f(i)x_i(t),\qquad 1\leq k\leq N.$$

Here, the different possible genotypes are numbered from 1 to N; $x_k(t)$ represents the concentration of individuals with genotype k at time t; $f(i)$ is the fitness (reproductive rate) of the ith genotype, and $Q(i,k)$ is the probability that an individual with genotype i is transformed into an individual with genotype k by mutation (unfaithful replication during reproduction). Thus, the first term in the differential equation accounts for the production of genotype k individuals, while the second term accounts for the loss of individuals with genotype k; the second term

---

Received 25 November 2014; revision received 22 May 2015.

* Postal address: DMA, École Normale Supérieure, 45, rue d'Ulm, 75230 Paris Cedex 05, France.
Email address: dalmau2718@gmail.com

is proportional to the concentration of genotype k individuals as well as to the population's average fitness, and it helps to keep the total concentration of chains constant.

One of the simplest scenarios we can consider is that of the sharp peak landscape together with independent mutations per locus. In the sharp peak fitness landscape all sequences but one, the master sequence, have the same fitness, while the master sequence has a higher fitness than the rest. Mutations happen during reproduction independently on each locus of the sequence, with equal probability. Eigen studied this simple scenario and found that two major phenomena take place. The first is an error threshold phenomenon: there is a critical mutation probability such that for above-critical mutation probabilities the equilibrium state of the population is a totally disordered one. The second phenomenon is found for below-critical mutation probabilities: in this case the equilibrium state of the population is no longer disordered; it contains a positive concentration of the master sequence, together with a cloud of mutants that closely resemble the master sequence. This kind of distribution has come to be known as a quasispecies distribution.

The concept of error threshold, as well as that of quasispecies, are very appealing to the scientific community, mostly due to their potential for qualitatively explaining the behaviour of a wide range of biological populations. Since Eigen introduced them, it has long been sought to extend the concepts to many other situations, both experimentally and theoretically. From a theoretical point of view, there are two main objections to the applicability of Eigen's model to more complex kinds of populations. The first objection comes from considering at the same time finite chain length and infinite population size: if the individuals we seek to model are fairly complex, the number of possible genotypes largely exceeds the size of any viable population, a feature that Eigen's model fails to account for. The second objection is due to the deterministic nature of Eigen's model: again, for fairly complex individuals, the description of the reproduction mechanism by chemico-kinetic reactions is completely out of reach, and it is typically replaced by some random mechanism. The program is thus settled: to retrieve the error threshold phenomenon and a quasispecies distribution for finite population stochastic models. For a discussion on the several contributions to this program, we refer the reader to [2] and [4].

In the series of papers [2]–[5], the authors studied the classical Moran and Wright–Fisher models, recovering both the error threshold phenomenon and a quasispecies distribution for mutation rates below the error threshold. Furthermore, the quasispecies distribution happens to be the same for both models, and an explicit expression was found: the concentration of sequences differing in exactly k digits from the master sequence is given by

$$(\sigma \mathrm{e}^{-a} - 1)\frac{a^k}{k!}\sum_{i=1}^{\infty}\frac{i^k}{\sigma^i},$$

where $\sigma > 1$ is the reproductive advantage of the master sequence and a is the product of the mutation probability with the length of the sequences. We call this distribution the quasispecies distribution with parameters σ and a, and we denote it by $\mathcal{Q}(\sigma, a)$. Both the Moran and the Wright–Fisher models are constant population models, since their aim is to describe a sufficiently large population which has stabilised in its environment. However, we might be interested in studying the evolution of a population in its early stages. The size of such a population is very likely to undergo significant fluctuations, the classical stochastic model for this situation is the Galton–Watson branching process. The aim in this paper is to study a Galton–Watson branching process, with selection and mutation, in order to recover the phase transition phenomenon and the quasispecies distribution.

Demetrius *et al.* [6] already pursued this task in a more general context: a general fitness landscape as well as a general mutation kernel. In [1], Antoneli *et al.* generalised the work in [6] by studying a multivariate branching process, which incorporates neutral, deleterious, and beneficial mutations. Our setting is closer to that of [6] than [1]; our aim is to show that for the sharp peak landscape along with per-locus independent mutations, the quasispecies distribution is again the one obtained for the Moran model and for the Wright–Fisher model. In [6], it was proved that the relative frequencies of the genotypes converge to those given by the stationary solution of Eigen's system of differential equations. However, the quasispecies distribution is a distribution on the Hamming classes of the sequence space, which arises in a particular asymptotic regime. Thus, we cannot apply the results in [6] directly. Along the lines of [2]–[5], we develop our argument from scratch. We start by defining the Galton–Watson process on the genotypes, with selection and mutation. We formally show how to pass from the process on the genotypes to a Galton–Watson process on the Hamming classes. The relative frequencies of the classes are shown to converge to the stationary solution of the corresponding Eigen's system, as shown in [6]. Finally, the stationary solution to this particular Eigen's system is shown to converge to the quasispecies distribution.

Our paper is organised as follows. First we define a multitype Galton–Watson process in order to model the evolution of a finite population. We state next the main result of this paper, and all the remaining sections are devoted to the proof of the main result.

2. The Galton–Watson process

In this section we define a multitype Galton–Watson process driving the dynamics of a finite population, which incorporates both selection and mutation effects. Let us begin by introducing the individuals that will form the population.

Individuals. Let $\mathcal{A}$ be a finite alphabet of cardinality $\kappa \geq 1$, and consider sequences of fixed length $\ell \geq 1$ over the alphabet $\mathcal{A}$. A sequence in $\mathcal{A}^\ell$ represents the genotype of a haploid individual. We study the evolution of a population of such individuals, with selection and mutation.

Sharp peak landscape. The selection mechanism is given by a fitness function $A\colon \mathcal{A}^\ell \to \mathbb{R}_+$. Many fitness landscapes might be considered, but we choose to work with the sharp peak landscape: there is a particular sequence $w^* \in \mathcal{A}^\ell$, called the master sequence, whose fitness is $\sigma \geq 1$, while every other sequence in $\mathcal{A}^\ell$ has fitness 1. So the fitness function in this case is given by, for all $u \in \mathcal{A}^\ell$,

$$A(u) = \begin{cases} \sigma & \text{if } u = w^*, \\ 1 & \text{if } u \neq w^*. \end{cases}$$

Independent mutations per locus. Mutations happen randomly due to unfaithful replication of the chains, independently on each locus of the chain, with equal probability $q \in (0, 1)$ for all loci. When an allele mutates, it does so to a randomly chosen letter, uniformly from the $\kappa - 1$ letters still available in the alphabet $\mathcal{A}$. This mutation mechanism can be encoded into a mutation kernel in the following manner:

$$M(u, v) = \prod_{i=1}^{\ell} \biggl((1-q)\, \mathbf{1}_{\{u(i)=v(i)\}} + \frac{q}{\kappa - 1}\, \mathbf{1}_{\{u(i)\neq v(i)\}} \biggr) \qquad \text{for all } u, v \in \mathcal{A}^\ell.$$

The multitype Galton–Watson process is a Markov chain with values in $\mathbb{N}^{\kappa^\ell}$,

$$X_n = (X_n(u),\ u \in \mathcal{A}^\ell), \qquad n \geq 0.$$

For each $u \in \mathcal{A}^\ell$ and $n \geq 0$, $X_n(u)$ represents the number of individuals with genotype u present in the population at time n. At each generation, each individual in the population gives birth to a random number of children, independently of the other individuals and of the past of the process. The number of offspring of an individual $u \in \mathcal{A}^\ell$ is distributed as a Poisson random variable with mean $A(u)$. The newborn individuals then mutate according to the kernel M. The new generation is formed by all the offspring, after mutation.

Generating functions. The classical tool for studying the Galton–Watson process we just described is generating functions, which are also useful for formally defining the transition mechanism of the process. Let $u \in \mathcal{A}^\ell$ and define the function $f^u\colon [-1,1]^{\mathcal{A}^\ell} \to \mathbb{R}$ by

$$f^u(s) = \sum_{r \in \mathbb{N}^{\mathcal{A}^\ell}} p^u(r) \prod_{v \in \mathcal{A}^\ell} s(v)^{r(v)} \quad \text{for all } s \in [-1,1]^{\mathcal{A}^\ell},$$

where $p^u(r)$ represents the probability that an individual with genotype u has $r(v)$ children with genotype v. For each $v \in \mathcal{A}^\ell$,

$$p^u(r) = \mathrm{e}^{-A(u)} \frac{A(u)^{|r|_1}}{|r|_1!} \prod_{v \in \mathcal{A}^\ell} M(u,v)^{r(v)} \quad \text{for all } r \in \mathbb{N}^{\mathcal{A}^\ell}.$$

Here $|r|_1$ represents the usual 1-norm of the vector r, that is, the sum of its components. For an initial population X_0 consisting of one genotype u individual only, X_1 is a random vector having generating function f^u. In general, for $n \geq 0$, if $X_n = r \in \mathbb{N}^{\mathcal{A}^\ell}$ then X_{n+1} is the sum of $|r|_1$ random vectors, where, for each $u \in \mathcal{A}^\ell$, $r(u)$ of the random vectors have generating function f^u. Note that the null vector is an absorbing state.

3. Main result

Since we work with the sharp peak landscape fitness function, we can classify the sequences in $\mathcal{A}^\ell$ according to the number of digits they differ from the master sequence. Precisely, the Hamming distance between two sequences $u, v \in \mathcal{A}^\ell$ is defined as the number of digits where the two sequences differ:

$$d_{\mathrm{H}}(u,v) = \operatorname{card}\{i \in \{1,\dots,\ell\}\colon u(i) \neq v(i)\}.$$

For each $k \in \{0,\dots,\ell\}$, let $\mathcal{C}_k$ be the set of the sequences in $\mathcal{A}^\ell$ at Hamming distance k from the master sequence:

$$\mathcal{C}_k = \{u \in \mathcal{A}^\ell\colon d_{\mathrm{H}}(u,w^*) = k\}.$$

We refer to the set $\mathcal{C}_k$ as the kth Hamming class. Our aim is to study the concentration of the individuals of X_n which are in the class k in the following asymptotic regime:

$$\ell \to \infty, \qquad q \to 0, \qquad \ell q \to a \in [0,\infty].$$

We have the following result.

Theorem 1. *The process $(X_n)_{n\geq 0}$ has a positive probability of survival. Conditioned on the event of nonextinction, if $\sigma \mathrm{e}^{-a} \leq 1$ then*

$$\lim_{\ell\to\infty,\, q\to 0,\, \ell q\to a} \; \lim_{n\to\infty} \frac{1}{|X_n|_1} \sum_{u \in \mathcal{C}_k} X_n(u) = 0 \quad \textit{for all } k \geq 0.$$

If $\sigma \mathrm{e}^{-a} > 1$ *then*

$$\lim_{\ell\to\infty,\, q\to 0,\, \ell q\to a}\ \lim_{n\to\infty} \frac{1}{|X_n|_1} \sum_{u\in\mathcal{C}_k} X_n(u) = (\sigma\mathrm{e}^{-a}-1)\frac{a^k}{k!}\sum_{i\geq 1}\frac{i^k}{\sigma^i} \quad \textit{for all } k\geq 0.$$

The right-hand side in this equation is the concentration of the kth Hamming class in the distribution of the quasispecies $\mathcal{Q}(\sigma, a)$ with parameters σ and a. We devote the rest of the paper to the proof of this result.

4. The occupancy process

In this section we build an occupancy process,

$$(Z_n)_{n\geq 0} = (Z_n(0), \dots, Z_n(\ell))_{n\geq 0},$$

to keep track of the number of sequences in each of the Hamming classes. Here $Z_n(l)$ represents the number of individuals in X_n that are at distance l from the master sequence. In order to build the occupancy process formally, we use the classical lumping technique [9].

Fitness. The fitness function A can be factorised into Hamming classes: define the function $A_{\mathrm{H}}\colon \{0, \dots, \ell\} \to \mathbb{R}_+$ by, for all $l \in \{0, \dots, \ell\}$,

$$A_{\mathrm{H}}(l) = \begin{cases} \sigma & \text{if } l = 0, \\ 1 & \text{if } 1 \leq l \leq \ell. \end{cases}$$

Then, for each $u \in \mathcal{A}^\ell$, we have $A(u) = A_{\mathrm{H}}(d_{\mathrm{H}}(u, w^*))$.

Mutations. The mutation matrix M can also be factorised into the Hamming classes. Indeed, for each $u \in \mathcal{A}^\ell$ and $c \in \{0, \dots, \ell\}$, the value

$$\sum_{v\in\mathcal{C}_c} M(u, v)$$

depends on u through its Hamming class only; see [2, Lemma 6.1]). For $b, c \in \{0, \dots, \ell\}$, let us call $M_{\mathrm{H}}(b, c)$ this common value for u in $\mathcal{C}_b$. The coefficient $M_H(b, c)$ can be analytically expressed as

$$\sum_{\substack{0\leq k\leq \ell-b,\\ 0\leq l\leq b,\\ k-l=c-b}} \binom{\ell-b}{k}\binom{b}{l} q^k (1-q)^{\ell-b-k}\left(\frac{q}{\kappa-1}\right)^l\left(1-\frac{q}{\kappa-1}\right)^{b-l}.$$

Lumping. Let $\mathcal{Z}\colon \mathbb{N}^{\mathcal{A}^\ell} \to \mathbb{N}^{\ell+1}$ be the map that associates to each population $r \in \mathbb{N}^{\mathcal{A}^\ell}$ the corresponding occupancy distribution

$$\mathcal{Z}(r)(l) = \sum_{u\in\mathcal{C}_l} r(u) \quad \text{for all } r\in\mathbb{N}^{\mathcal{A}^\ell},\ l\in\{0, \dots, \ell\}.$$

The occupancy process $(Z_n)_{n\geq 0}$ is defined by

$$Z_n = \mathcal{Z}(X_n) \quad \text{for all } n \geq 0.$$

We next check that the occupancy process is again a Galton–Watson process. Let $k \in \{0, \dots, \ell\}$, $u \in \mathcal{C}_k$, and $z \in \mathbb{N}^{\ell+1}$. We have

$$\begin{aligned}\sum_{r \in \mathbb{N}^{\mathcal{A}^\ell},\, \mathcal{Z}(r)=z} p^u(r) &= \sum_{r \in \mathbb{N}^{\mathcal{A}^\ell},\, \mathcal{Z}(r)=z} \mathrm{e}^{-A(u)} \frac{A(u)^{|r|_1}}{|r|_1!} \prod_{v \in \mathcal{A}^\ell} M(u, v)^{r(v)} \\ &= \mathrm{e}^{-A_{\mathrm{H}}(k)} \frac{A_{\mathrm{H}}(k)^{|z|_1}}{|z|_1!} \sum_{r \in \mathbb{N}^{\mathcal{A}^\ell},\, \mathcal{Z}(r)=z} \prod_{v \in \mathcal{A}^\ell} M(u, v)^{r(v)}.\end{aligned}$$

Decomposing the last sum and product into Hamming classes, we obtain

$$\begin{aligned}\sum_{r \in \mathbb{N}^{\mathcal{A}^\ell},\, \mathcal{Z}(r)=z} \prod_{v \in \mathcal{A}^\ell} M(u, v)^{r(v)} &= \sum_{l=0}^{\ell} \sum_{t \in \mathbb{N}^{\mathcal{C}_l},\, |t|_1 = z(l)} \prod_{v \in \mathcal{C}_l} M(u, v)^{t(v)} \\ &= \prod_{l=0}^{\ell} \bigg(\sum_{v \in \mathcal{C}_l} M(u, v) \bigg)^{z(l)} \\ &= \prod_{l=0}^{\ell} M_H(k, l)^{z(l)}.\end{aligned}$$

Let

$$p^k(z) = \sum_{r \in \mathbb{N}^{\mathcal{A}^\ell},\, \mathcal{Z}(r)=z} p^u(r) = \mathrm{e}^{-A_{\mathrm{H}}(k)} \frac{A_{\mathrm{H}}(k)^{|z|_1}}{|z|_1!} \prod_{l=0}^{\ell} M_H(k, l)^{z(l)}.$$

Since this expression depends on u only through k, the sum

$$\sum_{r' \in \mathbb{N}^{\ell+1},\, \mathcal{Z}(r')=z'} \mathbb{P}(X_{n+1} = z' \mid X_n = r)$$

depends on r only through $z = \mathcal{Z}(r)$. Thus, by the classical lumping theorem, the process $(Z_n)_{n \ge 0}$ is a Markov chain (the classical lumping theorem is stated in [9] for finite state space Markov chains, but both the result and the proof carry over word by word to the case of denumerable Markov chains). Let us define, for $k \in \{0, \dots, \ell\}$ and $s \in [0, 1]^{\ell+1}$,

$$f^k(s) = \sum_{z \in \mathbb{N}^{\ell+1}} p^k(z) \prod_{l=0}^{\ell} s(l)^{z(l)}.$$

The process $(Z_n)_{n \ge 0}$ is in fact a Galton–Watson process with $\ell + 1$ types, having the following transition mechanism: for all $n \ge 0$ and $z \in \mathbb{N}^{\ell+1}$, if $Z_n = z$ then Z_{n+1} is the sum of $|z|_1$ independent random vectors, where, for each $k \in \{0, \dots, \ell\}$, $z(k)$ of the vectors have generating function f^k.

Proof of Theorem 1. We use now the classical theory of branching processes [8] in order to study the process $(Z_n)_{n \ge 0}$. The mean matrix W of $(Z_n)_{n \ge 0}$ is the matrix with coefficients $W(i, j)$, $0 \le i, j \le \ell$, given by the expected number of class j individuals in the first generation, when the process starts with a population consisting of just one individual in the class i. The mean matrix is thus given by

$$W(i, j) = A_{\mathrm{H}}(i) M_{\mathrm{H}}(i, j) \quad \text{for all } i, j \in \{0, \dots, \ell\}.$$

The entries of the matrix W are all positive. By the Perron–Frobenius theorem, there exists a unique largest eigenvalue λ of W and a unique positive and unitary eigenvector ρ associated to λ. By the general theory of multitype Galton–Watson processes [8, Chapter II, Theorems 7.1 and 9.2]), if $\lambda \leq 1$ then the population goes extinct with probability 1. If $\lambda > 1$ there is a positive probability of survival, and conditioned on the event of nonextinction, we have

$$\lim_{n\to\infty} \frac{Z_n(k)}{Z_n(0)+\cdots+Z_n(\ell)} = \rho(k), \qquad 0 \leq k \leq \ell.$$

From the definition in [8], λ and ρ satisfy

$$\lambda\rho(k) = \sum_{i=0}^{\ell} \rho(i) A_{\mathrm{H}}(i) M_{\mathrm{H}}(i,k), \qquad 0 \leq k \leq \ell.$$

Summing the above expression over k, since ρ is a unitary vector, we deduce that

$$\lambda = \sum_{i=0}^{\ell} \rho(i) A(i) = (\sigma - 1)\rho(0) + 1.$$

Thus, the eigenvalue λ is equal to the average fitness of a population whose concentrations are given by the vector ρ. We remark that solving the above system of equations is equivalent to finding the stationary solutions of the corresponding Eigen's system of differential equations. From the above equation, we see that, in particular, $\lambda \in (1, \sigma)$, and this implies the first statement of the theorem: the process $(Z_n)_{n\geq 0}$ has a positive probability of survival. It remains to study the asymptotic behaviour of λ and ρ when ℓ goes to ∞, q goes to 0, and ℓq goes to a. In this asymptotic regime the mutation kernel M_{H} converges to the limiting expression, for all $i, k \geq 0$,

$$\lim_{\ell\to\infty,\, q\to 0,\, \ell q\to a} M_{\mathrm{H}}(i,k) = \begin{cases} \mathrm{e}^{-a} \dfrac{a^{k-i}}{(k-i)!} & \text{if } k \geq i, \\ 0 & \text{if } k < i. \end{cases}$$

Up to extraction of a subsequence, we can suppose that the following limits exist:

$$\lambda^* = \lim_{\ell\to\infty,\, q\to 0,\, \ell q\to a} \lambda, \qquad \rho^*(k) = \lim_{\ell\to\infty,\, q\to 0,\, \ell q\to a} \rho(k), \qquad k \geq 0.$$

Writing down the first equation of the system $\lambda\rho = \rho^{\top} W$, we see that

$$\sigma\rho M_{\mathrm{H}}(0,0) < \lambda\rho(0) < \sigma\rho M_{\mathrm{H}}(0,0) + \max_{1\leq i\leq \ell} M_{\mathrm{H}}(i,0).$$

Since we also know that $\lambda > 1$, we conclude that $\lambda^* \geq \max\{1, \sigma\mathrm{e}^{-a}\}$. As we have already pointed out

$$\lambda = (\sigma - 1)\rho(0) + 1.$$

Thus, taking the limits in the above two equations we deduce that

$$\lambda^* = (\sigma - 1)\rho^*(0) + 1 \quad \text{and} \quad \lambda^*\rho^*(0) = \sigma\rho^*(0)\mathrm{e}^{-a}.$$

Since $\lambda^* \geq \max\{1, \sigma \mathrm{e}^{-a}\}$, we conclude that

- If $\sigma \mathrm{e}^{-a} \leq 1$ then $\lambda^* = 1$ and $\rho^*(0) = 0$.
- If $\sigma \mathrm{e}^{-a} > 1$ then
$$\lambda^* = \sigma \mathrm{e}^{-a} \quad \text{and} \quad \rho^*(0) = \frac{\sigma \mathrm{e}^{-a} - 1}{\sigma - 1}.$$

Finally, writing down the kth equation of the system $\lambda\rho = \rho^\top W$, we see that

$$\sigma\rho_0 M_{\mathrm{H}}(0,k) + \sum_{i=1}^{k} \rho(i) M_{\mathrm{H}}(i,k) < \lambda\rho(k)$$

$$< \sigma\rho_0 M_{\mathrm{H}}(0,k) + \sum_{i=1}^{k} \rho(i) M_{\mathrm{H}}(i,k) + \max_{k<i\leq\ell} M_{\mathrm{H}}(i,k).$$

Thus, taking the limit we obtain the recurrence relation

$$\sigma \mathrm{e}^{-a} \rho^*(k) = \sigma\rho^*(0)\mathrm{e}^{-a}\frac{a^k}{k!} + \sum_{i=1}^{k} \rho^*(i)\mathrm{e}^{-a}\frac{a^{k-i}}{(k-i)!}, \qquad k \geq 1.$$

We conclude that if $\sigma \mathrm{e}^{-a} \leq 1$ then $\rho^*(k) = 0$ for all $k \geq 0$, and if $\sigma \mathrm{e}^{-a} > 1$ then

$$\rho^*(k) = (\sigma \mathrm{e}^{-a} - 1)\frac{a^k}{k!}\sum_{i\geq 1}\frac{i^k}{\sigma^i}, \qquad k \geq 0.$$

This can be seen by solving the recurrence relation by the method of generating functions, see, for example, [4]. □

References

[1] Antoneli, F., Bosco, F., Castro, D. and Janini, L. M. (2013). Virus replication as a phenotypic version of polynucleotide evolution. *Bull. Math. Biol.* **75,** 602–628.

[2] Cerf, R. (2015). Critical population and error threshold on the sharp peak landscape for a Moran model. *Mem. Amer. Math. Soc.* **233,** No. 1096.

[3] Cerf, R. (2015). Critical population and error threshold on the sharp peak landscape for the Wright–Fisher model. *Ann. Appl. Prob.* **25,** 1936–1992.

[4] Cerf, R. and Dalmau, J. (2016). The distribution of the quasispecies for a Moran model on the sharp peak landscape. *Stoch. Process. Appl.* **126,** 1681–1709.

[5] Dalmau, J. (2015). The distribution of the quasispecies for a Wright–Fisher model on the sharp peak landscape. *Stoch. Process. Appl.* **125,** 272–293.

[6] Demetrius, L., Schuster, P. and Sigmund, K. (1985). Polynucleotide evolution and branching processes. *Bull. Math. Biol.* **47,** 239–262.

[7] Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58,** 465–523.

[8] Harris, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.

[9] Kemeny, J. G. and Snell, J. L. (1960). *Finite Markov Chains*. Van Nostrand, Princeton, NJ.