Journal of Theoretical Biology 365 (2015) 238–246

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Effect of drift, selection and recombination on the equilibrium frequency of deleterious mutations

Sona John *, Kavita Jain

Theoretical Sciences Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur P.O., Bangalore 560064, India

HIGHLIGHTS

- Equilibrium fitness is studied in a model with back mutations.
- Model takes care of linkage effects.
- Recombination is found to be important in adapting microbial populations.
- Mild effect of recombination is observed in codon usage bias.

ARTICLE INFO

Article history:
Received 8 June 2014
Received in revised form
30 August 2014
Accepted 22 October 2014
Available online 31 October 2014

Keywords:
Back mutations
Linkage
Finite population

ABSTRACT

We study the stationary state of a population evolving under the action of random genetic drift, selection and recombination in which both deleterious and reverse beneficial mutations can occur. We find that the equilibrium fraction of deleterious mutations decreases as the population size is increased. We calculate exactly the steady state frequency in a nonrecombining population when population size is infinite and for a neutral finite population, and obtain bounds on the fraction of deleterious mutations. We also find that for small and very large populations, the number of deleterious mutations depends weakly on recombination, but for moderately large populations, recombination alleviates the effect of deleterious mutations. An analytical argument shows that recombination decreases disadvantageous mutations appreciably when beneficial mutations are rare as is the case in adapting microbial populations, whereas it has a moderate effect on codon bias where the mutation rates between the preferred and unpreferred codons are comparable.

1. Introduction

A large number of population genetic studies assume one-way mutation – in some situations, beneficial mutations are neglected as they occur rarely (Muller, 1964; Felsenstein, 1974; Haigh, 1978; Gordo and Charlesworth, 2000) while in adaptation studies, deleterious mutations are ignored as they are unlikely to fix under strong selection conditions (Gerrish and Lenski, 1998; Rouzine et al., 2008; Seetharaman and Jain, 2014). The assumption of one-way mutation has an important effect on the nature of the state at large times. If the population size is infinite, a time-independent stationary state can be reached due to a balance between mutation and selection even if the mutational forces are unidirectional (Haigh, 1978). However in a finite population, when mutations are completely neglected or only unidirectional mutations are allowed, a population evolving under the influence of other evolutionary forces either does not reach an equilibrium state (Haigh, 1978), or achieves a trivial one in which one of the variants gets fixed at large times (Ewens, 2004). It is when both beneficial and deleterious mutations are taken into account, a finite population reaches a nontrivial stationary state (Wright, 1931).

An example of such a steady state is seen in the context of synonymous codons that represent the same amino acid but do not occur in equal frequencies (Hershberg and Petrov, 2008; Plotkin and Kudla, 2011). In a gene coding for a two-fold degenerate amino acid, while selection favors the preferred codon, reversible mutations between preferred and unpreferred codons and random genetic drift maintain the unpreferred one (Li, 1987; Bulmer, 1991). Assuming that the sites in the sequence evolve independently, analytical results for the equilibrium frequency of unpreferred codons have been obtained (Li, 1987; Bulmer, 1991; McVean and Charlesworth, 1999). However as the evolutionary dynamics at a genetic locus are affected by other loci (Hill and Robertson, 1966), a proper theory of codon usage bias must account for the Hill–Robertson interference between sequence loci (Comeron et al., 1999; McVean and Charlesworth, 2000; Charlesworth et al., 2009; Kaiser and Charlesworth, 2009).

* Corresponding author.
E-mail addresses: sonajohn@jncasr.ac.in (S. John), jain@jncasr.ac.in (K. Jain).

http://dx.doi.org/10.1016/j.jtbi.2014.10.023
0022-5193/

Reverse and compensatory mutations have also been proposed as a possible mechanism to stop the degeneration of asexual populations (Lande, 1998; Whitlock, 2000; Goyal et al., 2012). In a finite nonrecombining population, if beneficial mutations are completely ignored, deleterious mutations accumulate irreversibly due to stochastic fluctuations by a process known as Muller's ratchet (Muller, 1964; Howe and Denver, 2008). But when rare beneficial mutations are taken into account, the population reaches an equilibrium (Estes and Lynch, 2003; Silander et al., 2007; Howe and Denver, 2008). Recently Goyal et al. (2012) calculated the amount of beneficial mutations required to achieve a stationary state. But these authors assumed the mutation rates to be independent of the fitness, contrary to experimental evidence (Silander et al., 2007). Moreover their solution for the equilibrium frequency can become negative in some parameter range.

In this article, we are interested in understanding the stationary state of a multilocus model, which is described in detail in the following section. We consider a class of non-epistatic fitness landscapes where the fitness depends only on the number of deleterious mutations in a sequence (*fitness class*). As in previous works (Li, 1987; Comeron et al., 1999; McVean and Charlesworth, 2000), we assume that the beneficial mutations are back mutations, the probability of whose occurrence depends on the fitness class. More precisely, if the mutation probability per site is small, the total probability of a beneficial (deleterious) mutation increases (decreases) linearly with the fitness class. We consider the evolution of both infinitely large and finite populations, and to analyse the effect of linkage amongst the loci, we allow recombination to occur. We are primarily interested in the population size dependence of the average number of disadvantageous mutations at equilibrium. We obtain analytical results when the sites are completely linked, and compare them with the known results for a freely recombining population. For intermediate recombination rates, we obtain numerical results.

We find that the number of deleterious mutations decreases in a reverse sigmoidal fashion, as the population size is increased. For small populations, the fraction of disadvantageous mutations is seen to be roughly independent of population size and recombination rate. An understanding of this behaviour is obtained from an exact solution and numerical simulations for a neutral finite population. For very large populations that can be described by a deterministic model, we find the stationary state exactly which is also unaffected by recombination. However for moderately large populations, recombination is found to alleviate the effect of deleterious mutations (Hill and Robertson, 1966; Felsenstein, 1974; Barton and Charlesworth, 1998; Charlesworth et al., 2009), and the extent to which it does so depends on the beneficial mutation rate relative to the deleterious one. We find that when beneficial mutations are rare, the equilibrium frequency of disadvantageous mutations decreases logarithmically with population size when the loci are completely linked, but exponentially fast when linkage is absent. On the other hand, when disadvantageous mutations are rare, the deleterious mutation fraction drops exponentially fast, irrespective of the recombination rate. Thus we expect that the linkage has a weak effect on codon bias where the rates at which mutations between preferred and unpreferred codons occur are of the same order (Zeng, 2010; Schrider et al., 2013). But in adapting microbial populations where beneficial mutations are rare (Sniegowski and Gerrish, 2010), recombination may be expected to reduce the frequency of disadvantageous mutations significantly.

2. Models

We consider a haploid population of size N in which each individual carries a biallelic (either zero or one) sequence of finite length L, where zero represents the wild type allele and one denotes the deleterious mutation. The population is evolved in computer simulations using a Wright–Fisher process in which recombination followed by mutation and selection occurs in discrete, non-overlapping generations. To create an offspring, two parent individuals are chosen at random with replacement. With probability $r \leq 1/2$, a single crossover event occurs in the parent sequences at one of the $L-1$ equally likely break points to form two recombinant sequences, while with probability $1-r$, the parent sequences are copied to the offspring sequences. In either case, one of the offspring is chosen with probability half to undergo mutations and selection, and the other one is discarded. In the offspring sequence, a deleterious mutation occurs at a locus with a wild type allele with probability μ and a reverse beneficial mutation on mutant allele with probability ν. The resulting sequence is allowed to survive with a probability equal to its fitness, where the fitness of a sequence with j deleterious mutations is assumed to be a nonepistatic, and given by $w(j)=(1-s)^j$, $0 \leq s < 1$. This process is repeated until N individuals in the next generation are obtained.

We have been able to implement the procedure described above for sequences of length up to 500 and population sizes of the order 10^3. For larger populations with long nonrecombining sequence, the computational difficulties were overcome by tracking only the number of deleterious mutations (fitness class) carried by the individual since the fitness of a sequence depends only on the number of deleterious mutations in the sequence. Here a parent chosen at random produces a clone of itself, and the offspring may undergo mutations with a probability that depends on its fitness class. In a sequence with j deleterious mutations, as a deleterious (beneficial) mutation can happen at any one of the $L-j$ (j) sites, the rate of deleterious and beneficial mutations is given by $(L-j)\mu$ and $j\nu$ respectively. To find the number of beneficial (b) and deleterious (d) mutations acquired by the offspring, random variables were drawn from Poisson distribution with mean $j\nu$ and $(L-j)\mu$ respectively. The total number of deleterious mutations in the offspring is then given by $j' = j+d-b$. If j' turns out to be greater than L or less than zero, the offspring individual is produced with $j' = j$ mutations. As before, the offspring is allowed to survive with probability $w(j')$, and the process is repeated until N individuals in the next generation are obtained.

All the numerical results presented here are obtained with an initial condition in which none of the individuals carry deleterious mutations. In each stochastic run, the Wright-Fisher process was implemented for about 10^4 generations and it was ensured that the stationary state is reached. In the equilibrium state of each run, we measured the number of deleterious mutations present in the population and averaged them over another 10^4 generations. The data were also averaged over 100 independent stochastic runs. Although all the simulation results presented here are obtained using the Wright–Fisher process, we will also use a continuous time Moran model for some analytical calculations which is described in a later section. If the population is infinitely large, the dynamics and equilibrium state of the population fraction can be described by a deterministic equation, which we discuss next.

3. Results

3.1. Infinite population

3.1.1. Nonrecombining population

For small selection coefficient and mutation rates, the population fraction $X(j,t)$ in the jth fitness class at time t evolves in