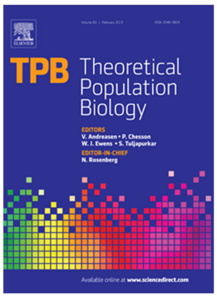

Inference from the stationary distribution of allele frequencies in a family of Wright-Fisher models with two levels of genetic variability

Jake M. Ferguson, Erkan Ozge Buzbas

PII: S0040-5809(18)30060-1
DOI: https://doi.org/10.1016/j.tpb.2018.03.004
Reference: YTPBI 2635

To appear in: *Theoretical Population Biology*

Please cite this article as: Ferguson J.M., Buzbas E.O., Inference from the stationary distribution of allele frequencies in a family of Wright-Fisher models with two levels of genetic variability. *Theoretical Population Biology* (2018), https://doi.org/10.1016/j.tpb.2018.03.004

Inference from the stationary distribution of allele frequencies in a family of Wright-Fisher models with two levels of genetic variability

Jake M. Ferguson[a], Erkan Ozge Buzbas[b]

[a] *Center for Modeling Complex Interactions, University of Idaho*
[b] *Department of Statistical Science, University of Idaho*

Abstract

The distribution of allele frequencies obtained from diffusion approximations to Wright-Fisher models are useful in developing intuition about the population level effects of evolutionary processes. The statistical properties of the stationary distributions of K-allele models have been extensively studied under neutrality or under selection. Here, we introduce a new family of Wright-Fisher models in which there are two hierarchical levels of genetic variability. The genotypes composed of alleles differing from each other at the selected level have fitness differences with respect to each other and evolve under selection. The genotypes composed of alleles differing from each other only at the neutral level have the same fitness and evolve under neutrality. We show that with an appropriate scaling of the mutation parameter with respect to the number of alleles at each level, the frequencies of alleles at the selected and the neutral level are conditionally independent of each other,conditional on knowing the number of alleles at all levels. This conditional independence allows us to simulate from the the joint stationary distribution of the allele frequencies. We use these simulated frequencies to perform inference on parameters of the model with two levels of genetic variability using Approximate Bayesian Computation.

Keywords: Wright-Fisher model, K-allele models, diffusion approximation, balancing selection, Approximate Bayesian Computation

---