

Simulation of partially clonal diploid populations with simuPOP v.1.0.8

Zhian N. Kamvar

June 24, 2013

1 Methods

2 Introduction

SimuPOP v.1.0.8 is a python scripted, forward-time, individual-based simulator that can simulation populations with any marker type. It works not by drawing data points from a distribution, but rather by initializing representatives of individuals and having these individuals undergo mating with user specified population parameters. Often, it is used for traditional population genetics analyses such as pedigree or genetic drift simulations. Simulations through simupop undergo a two step process, they are first initialized by user defined or random processes, then ‘evolved’ through a specific mating and mutation regiment. We simulated 100 replicates of populations with increasing levels of clonal reproduction over 10 loci with 6 to 10 alleles per locus and 10,000 individuals each. These populations were evolved under the different levels of clonal reproduction and then sub sampled 10 times with the sample sizes of 10, 25, 50, and 100 individuals.

12 Population Initialization

All of the clonal scenarios were simulated from the same 100 seed populations. This was to ensure that variation in the treatment was not due to stochasticity in the initial populations. The seed populations were initialized by first determining the number of individuals (10,000) and number of microsatellite loci (10). Microsatellites often exhibit more than two allelic states in a population and are rapidly evolving. For this reason, at each locus, the number of alleles were randomly assigned to be between 6 and 10 alleles. After the assignment of alleles, they were randomly assigned allele frequencies summing up to one.

The populations were then initialized with 10,000 individuals, half of which were assigned “Male” status and half of which were assigned “Female” status. In terms of Oömycetes, this represents a heterothallic system. Each individual was then assigned a genotype at all loci using the frequencies previously determined. Since the locus parameters were randomly assigned, there was no way to guarantee that these populations would be in equilibrium from the beginning, so all of the 100 seed populations underwent a ”burn in” period of 1,000 generations of random mating with no mutation in an attempt to reach equilibrium.

25 Evolution

There were ten levels of sexual reproduction tested: 0%, 0.01%, 0.05%, 0.1%, 1%, 5%, 10%, 20%, 50%, and 100%. Each level represents the proportion of offspring produced via sexual reproduction.