**PROGRAM NOTE**

**BOTTLESIM: a bottleneck simulation program for long-lived species with overlapping generations**

C.-H. KUO and F. J. JANZEN

339 Science II, Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, USA

**Abstract**

Population bottlenecks reduce genetic diversity and thus cause great concern in conservation biology. Previous theoretical studies often assume discrete generations in projecting declines in genetic diversity caused by bottlenecks. This assumption creates complexities when applying the models to long-lived species with overlapping generations. **BOTTLESIM** is a program for simulating bottlenecks to estimate the impact on genetic diversity; the novelties include an overlapping-generation model, a wide range of reproductive systems, and flexible population size settings. With these features, **BOTTLESIM** will be a useful tool for estimating the genetic consequences of bottlenecks, evaluating conservation plans, and performing power analysis.

**Keywords:** computer simulation, conservation genetics, genetic diversity, heterozygosity, long-lived species, population bottleneck

*Received 10 July 2003; revision accepted 30 August 2003*
model leave large gaps between generations for long-lived species. Most of the simulation programs available to population geneticists are designed under a discrete-generation model (e.g. EASYPOP (Balloux 2001) and GENELoss (England & Osler 2001); for the ones that accommodate overlapping generations, e.g. MANAGEDPOP (Birnbaum et al. 2002) and META SIM (Strand 2002), constraints on population size settings render those programs unsuitable for simulating population bottlenecks.

BOTTLE SIM is a program specifically designed for simulating the genetic consequences of bottlenecks and postbottleneck population growth for long-lived species. In addition to an overlapping-generation model, BOTTLE SIM allows users to specify an arbitrary population size each year. Compared to other programs that are limited to a constant population size throughout the simulation (e.g. EASYPOP, GENELoss, MANAGEDPOP), this feature enables users to simulate a wide range of scenarios (e.g. gradual vs. rapid population decline, repeated bottlenecks, exponential vs. logistic population growth), provides a more realistic model, and eliminates the need to calculate effective population size when population size fluctuates. Though the harmonic mean of successive generations generally provides a good approximation of effective population size (Nei et al. 1975), significant deviations can occur if populations experience cyclic changes in size (Motro & Thomson 1982).

BOTTLE SIM also provides a wide range of reproductive system settings. The reproductive system can be set to asexual, monoecious (strict selfing, random mating with selfing, and random mating without selfing), or dioecious (random mating, single reproducing male each year, and single reproducing pair each year). Sex ratio can be specified by users when a dioecious reproductive system is chosen. In cases where reproductive skew is simulated, a single reproducing male or pair is randomly chosen from the population each year and reproduces all the new individuals in that year. With these options and those involving overlapping generations and population size, BOTTLE SIM is expected to be a useful tool for population and conservation geneticists.

The current version (version 2.6) of BOTTLE SIM includes four simulation modules (single locus with constant population size, single locus with variable population sizes, multilocus with constant population size, and multilocus with variable population sizes). The diploid single locus modules accept arbitrary allele frequencies at a single locus and are intended for exploring the genetic consequences of population bottlenecks under different settings. The diploid multilocus modules accept multilocus genotypic data of codominant markers, and are useful for evaluating conservation plans of different population sizes based on the empirical genotypic data. The input files must be ASCII text files (e.g. MS-DOS text or text-only); the input file format is described in detail in the users’ manual.

User-defined parameters for simulations include the degree of generation overlap (0–100%, 0 = discrete generations, 100 = completely overlapping generations), reproductive system, expected longevity of the organism, age of reproductive maturation, population sizes, the number of years to simulate, and the number of iterations. Observed number of alleles and allele frequencies in the founder population can be specified in the input file. All parameters are only limited by the amount of memory available. Selection, migration, and mutation are not included in the simulation model.

Each iteration starts by generating a founder population with the population size, observed number of alleles and allele frequencies specified by users. An age value is assigned to each individual; the degree of generation overlap specifies the percentage of individuals assigned a random age value within the limit of expected longevity. The age of individuals that are not assigned randomly is set to zero. When the degree of overlap is set to zero, all individuals start at age zero and reach the longevity limit in the same year, resulting in a complete population turnover equivalent to a discrete-generation model. When the degree of overlap is set to 100, all individuals in the population start with a random age value. The expected number of individuals in each age group is (N/L), where N is the founder population size and L is the length of expected longevity.

Each simulation year includes two major temporal steps: (i) from previous year’s end to current year’s beginning; (ii) from current year’s beginning to current year’s end. In the first major step, the program checks if population decline occurs and generates a list of surviving individuals. Age of all surviving individuals is increased by one. The second major step includes three substeps: (i) identify the individuals that reach reproductive maturity and generate a list of possible reproducing individuals; (ii) identify the individuals that reach the longevity limit, replace them with new genotypes according to the reproductive system setting, and reset age to zero; and (iii) check if population growth occurs and generate new individuals accordingly. All genetic diversity measurements are calculated at the end of each simulation year.

The output files generated by BOTTLE SIM are in ASCII text format with space delimiters, and can be viewed directly by word-processing software or imported into spreadsheet software for any further analysis. The summary output file contains all simulation settings, fixation probability, observed number of alleles (OA), effective number of alleles (EA), observed heterozygosity (H_O), expected heterozygosity (H_E), and the fixation index (F = (H_E - H_O) / H_E). The optional genotypic data output file contains the raw genotypic data from the last year of each iteration. The
genotypic data output is in *Genepop* (Raymond & Rousset 1995) format. Availability of raw genotypic data allows users to perform power analysis of statistical genetic tests.

To validate the simulation algorithm, we compared the simulation output to analytical expectation for the loss of heterozygosity (Fig. 1) and to another simulation program (England & Osler 2001) for the loss of alleles (Fig. 2) under the discrete-generation model. Excellent agreement was found under all simulation settings tested. All validating simulations were performed on an iMac with an 800-MHz PowerPC G4 processor and 512 MB of memory. Each of the validating simulations was completed within 1 min. Figure 3 demonstrates the flexibility of the program, including various settings of the degree of generation overlap (Fig. 3a), representative reproductive systems (Fig. 3b), and fluctuating population sizes (Fig. 3c). One example (Fig. 3c, ▲) also shows that using the harmonic mean as the effective population size can lead to significant deviation when a population experiences cyclical changes in size (Motro & Thomson 1982).

The source code of *bottlesim* is written in C++ programming language. All the pseudorandom numbers used in the simulation process were obtained by using the rand() function in the C++ mathematical library. The random number generator function is seeded with the starting time of each run to ensure the uniqueness of each simulation. The source code, sample input files, users’ manual, and compiled executables for Mac OS 9, Mac OS X, and Microsoft Windows operating systems are available free of charge at http://www.public.iastate.edu/~fjanzen/.

**Acknowledgements**

This work was supported by grants from the Chelonian Research Foundation Linnaeus Fund and the Society for Integrative and Comparative Biology Grants-in-Aid of Research program to C-HK and NSF IBN-0212935 to FJJ. We thank R. Fernando, G. Luikart, J. Nason, L. Nunney, and M. Whitlock for helpful comments.
Fig. 2 Validating the loss of alleles. (a) Five alleles at equal frequencies. (b) Five alleles at frequencies 0.6, 0.1, 0.1, 0.1, 0.1. (c) Ten alleles at equal frequencies. Top curves show simulation result from a constant population size of 100 (bottlesim: solid lines, geneloss: ●). Bottom curves show simulation results from a constant population size of 10 (bottlesim: solid lines, geneloss: ●). Other simulation parameters were set as follows: degree of generation overlap = 0, monoecy with random mating (with selfing), expected longevity = 1 year, age of reproductive maturation = 1, number of years simulated = 100, and number of iterations = 1000 (error bars indicating 95% confidence intervals are not shown because they are not discernible).

Fig. 3 The loss of observed number of alleles under different simulation settings. (a) Degree of generation overlap = 0 (●), 50 (■), and 100 (■). Monoecy with random mating (with selfing), expected longevity = 10 years, age of reproductive maturation = 1, constant population size of 20. Note that higher degree of generation overlap accelerates loss of alleles. (b) Representative reproductive systems: asexual reproduction (●), monoecy with complete selfing (▲), dioecy with random mating (●), dioecy with single reproducing pair each year (■). Degree of generation overlap = 0, expected longevity = 1 years, age of reproductive maturation = 1, constant population size of 20. (c) Fluctuating population sizes: 90 generations of population size = 100 with 10 generations of bottleneck (population size = 10) in generation 1–10 (■), generation 46–55 (●), generation 91–100 (▲), and one bottleneck every 10 generations (▲). The smooth line shows a constant population size of 53 (approximation from harmonic mean). Degree of generation overlap = 0, monoecy with random mating (with selfing), expected longevity = 1 year, age of reproductive maturation = 1. Other simulation parameters were set as follows: 10 alleles at equal frequencies, number of years simulated = 100, and number of iterations = 1000 (error bars indicating 95% confidence intervals are not shown because they are not discernible).
References


