Genetic Simulation Library

John S. Conery and Michael Lynch

Computational Science Institute, University of Oregon, Eugene, OR 97403, USA

Abstract

Summary: *The Genetic Simulation Library (GSL) is a set of C++ programs and classes that can be used in individual-based simulation models of plant and animal populations. The classes in this library build representations of genes for a wide variety of different types of organisms, from asexually reproducing single-cell organisms to chromosome-structured sexually reproducing species.*
Availability: *http://www.csi.uoregon.edu/projects/genetics/GSL. Free for non-commercial applications; others should contact the author.*

An individual-based ecological model is a computer simulation in which each member of a population is represented as a distinct object (DeAngelis and Gross, 1992; Bart, 1996). Rather than summarizing average case behavior with systems of differential equations, an individual based model captures the essential features of each individual plant or animal in a distinct object. An object state is updated over time as the individual represented by that object interacts with other individuals or with its environment.

The Genetic Simulation Library (GSL) is a collection of stand-alone programs and C++ classes that can be used to incorporate genetic information with individual-based models. By adding a genotype to the representation of each individual, it is possible to track the progress of new mutations introduced into a population, or to use the relative fitness of an individual as a factor in its interactions with the rest of the population, or any of a number of other potential operations. The library has been used by the authors in simulation studies of 'mutational meltdown' in small populations (Lynch *et al.*, 1995a,b; Conery *et al.*, 1995). We are currently using the library in simulations of Coho salmon populations.

The class hierarchy is illustrated in Figure 1. The major component is the Genome class, which is an abstract class with definitions of functions that create and manipulate representations of the genes of a single individual. The Individual class is a template for the representation of a single organism. The Generation and Population classes are container classes that manage collections of individuals. Additional modules we developed for the mutational meltdown and Coho salmon models are also available in the library; these include an extensive set of random number generators (the class named RNG) and a class for computing summary statistics over populations.

There are two ways to use the class library, depending on whether one wants to add genetics to an existing model or to create a new model from scratch.

To add genetics to an existing C++ program, include the GSL header file and add a Genome field to the definition of an individual. The existing model can then be extended by invoking operations on genotypes at the appropriate places, e.g. combining genotypes from two parents when a new individual is created. The library includes three different implementations of the Genome class to allow users to choose a representation that is most efficient for each application.

A new simulation can be built using methods defined in the Population class, which contains the necessary procedures to create one or more populations and track the genotypes of the individuals in the populations over a series of generations. The Population class includes methods for transferring individuals between population objects so users can build meta-population models. The mutational meltdown simulator in the GSL application directory is an example of a program based on the Population class; one way of developing a new model would be to extend the mutational meltdown application with demographic and environmental factors for the population to be studied.

The class library is available at the project web site, http://www.csi.uoregon.edu/projects/genetics/GSL. The web site contains instructions for downloading precompiled binaries, complete program source code, extensive on-line documentation, and several demo programs and complete applications built with GSL classes as examples of how the library can be used.
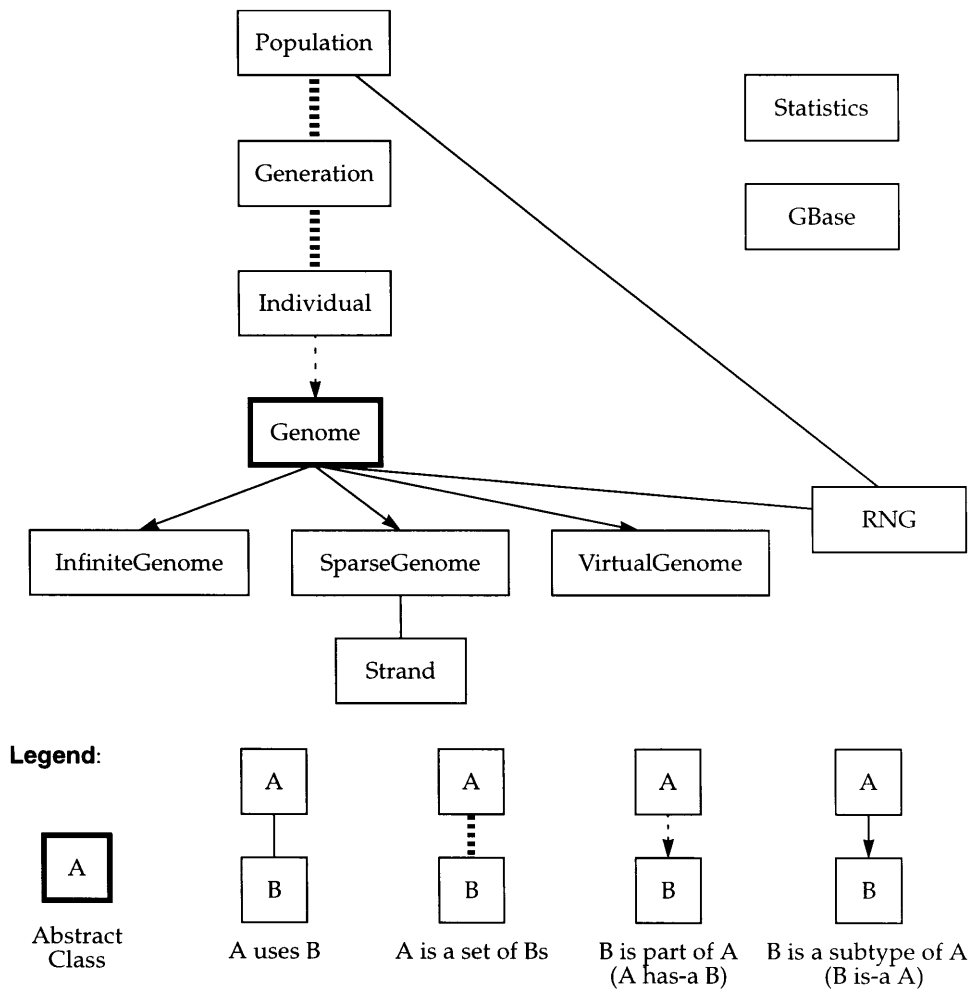
Fig. 1. GSL class hierarchy. The Population, Generation, Individual, and Genome classes and their subtypes all create representations of simulation objects. The other three classes implement general utilities: statistics computes summary statistics of various events, GBase is a genotype database useful for initializing simulation runs, and RNG is a collection of random number generators.

References

Bart,J. (1996) Acceptance criteria for using individual-based models to make management decisions. *Ecologic. Applic.*, **5,** 411–420.

Conery,J.S., Lynch,M. and Hovland,T. (1995) Optimizing irregular computations on SIMD machines: A case study. In *Proceedings of the Fifth Symposium on Frontiers of Massively Parallel Computation*, IEEE Press, pp. 222–230.

DeAngelis,D.L. and Gross,L.J. (1992) *Individual Based Models and Approaches in Ecology.* Chapman and Hall.

Lynch,M., Conery,J.S. and Bürger,R. (1995a) Mutation accumulation and the extinction of small populations. *Am. Naturalist*, **146,** 489–518.

Lynch,M., Conery,J.S. and Bürger,R. (1995b) Mutational meltdowns in sexual populations. *Evolution*, **49,** 1067–1080.