

Forward-time simulations using simuPOP, a tutorial

Bo Peng, Ph.D.

Department of Epidemiology
U.T. M.D. Anderson Cancer Center
Houston, TX

June. 6, 2007
Programmers' Cross Training
U.T. M.D. Anderson Cancer Center

outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

- 1 What is simuPOP
- 2 An example
- 3 simuPOP components

simuPOP is ...

**simuPOP
tutorial**

Bo Peng,
Ph.D.

**What is
simuPOP**

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

**simuPOP
components**

A forward-time population genetics simulation environment

simuPOP is ...

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP
components

A forward-time population genetics **simulation** environment

- A population genetics simulation program

simuPOP is ...

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP
components

A **forward-time** population genetics simulation environment

- A population genetics simulation program
- Not coalescent-based

simuPOP is ...

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP
components

A forward-time population genetics simulation **environment**

- A population genetics simulation program
- Not coalescent-based
- Based on an object-oriented scripting language (Python)

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

1 What is simuPOP

- Forward- and backward-time simulation
- Features of simuPOP
- Applications
- Availability

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

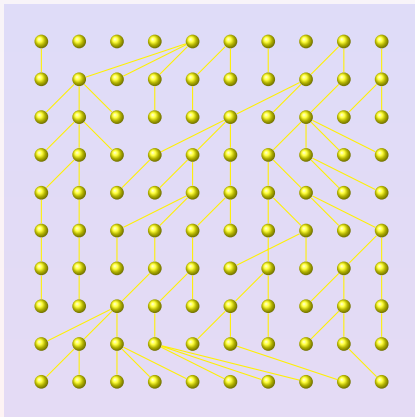
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

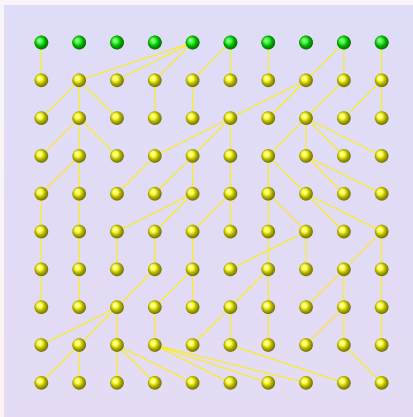
Features of
simuPOP

Applications
Availability

An example

simuPOP components

- Start from an initial population



Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

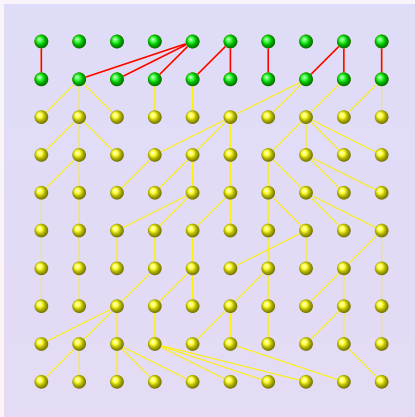
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

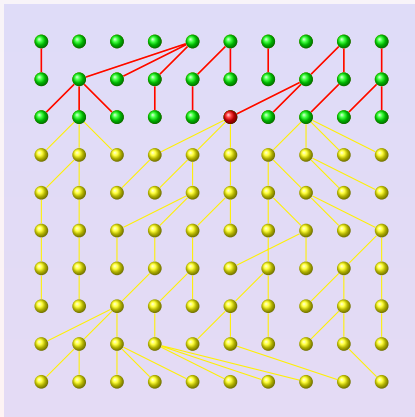
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

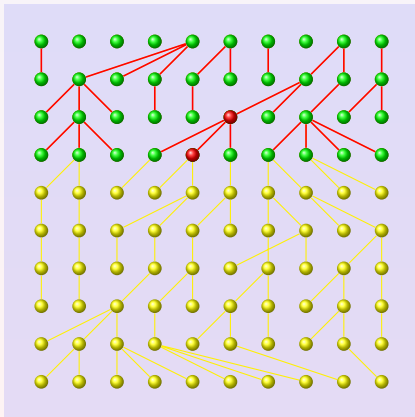
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

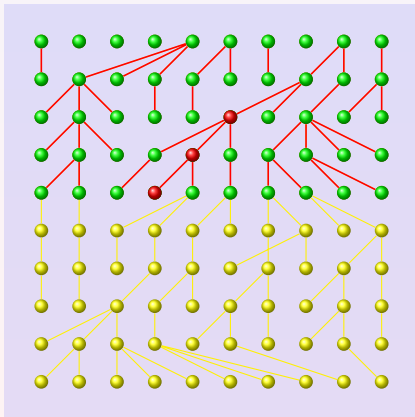
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

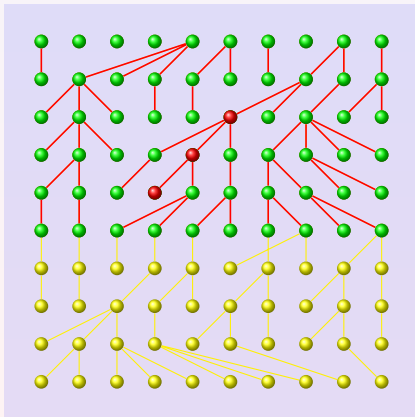
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

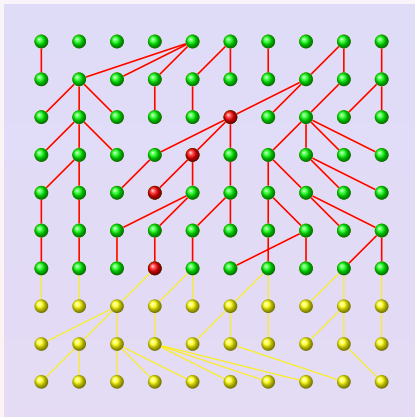
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

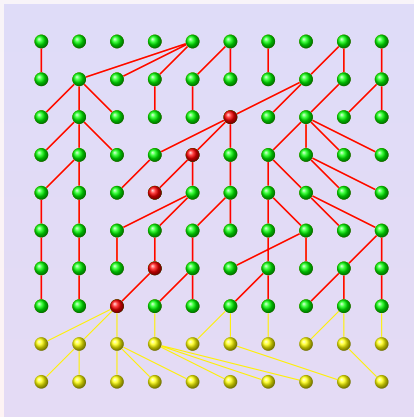
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP
components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

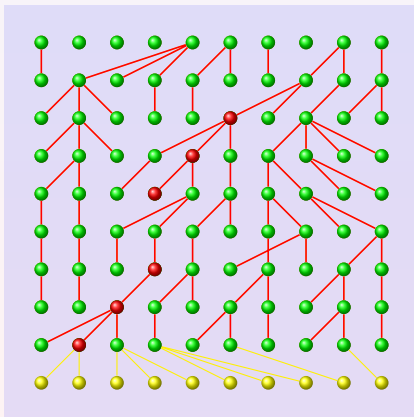
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP
components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects

Forward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

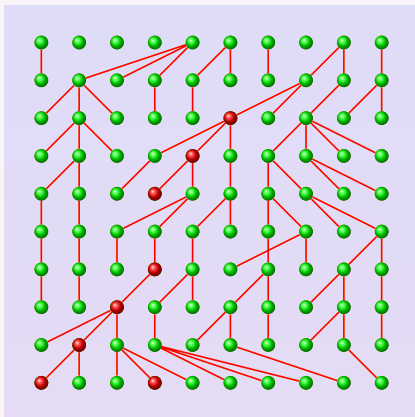
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from an initial population
- Evolve forward in time, generation by generation, subject to certain number of genetic and/or demographic effects
- Samples are collected from the last several generations

Backward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

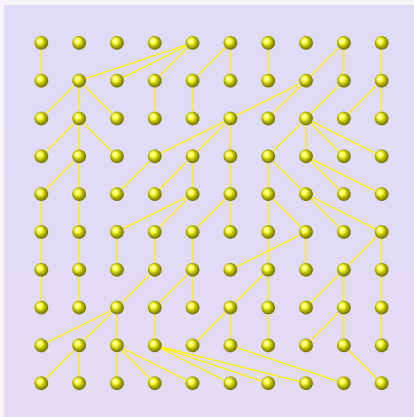
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



Backward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

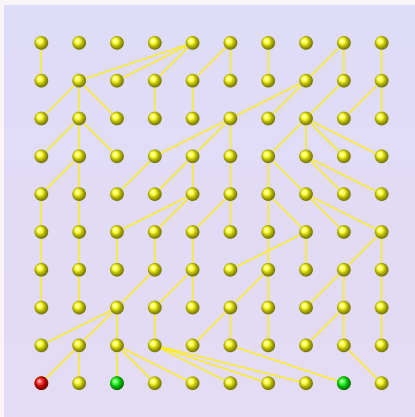
Features of
simuPOP

Applications
Availability

An example

simuPOP components

- Start from a sample with unknown genotype



Backward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

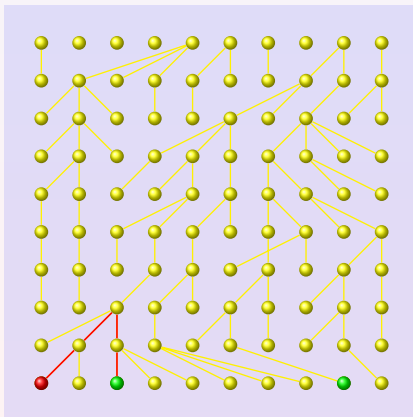
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from a sample with unknown genotype
- Coalesce individuals until the most recent common ancestor of all individuals is found

Backward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

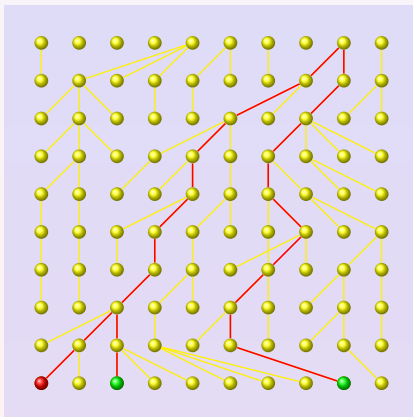
Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components



- Start from a sample with unknown genotype
- Coalesce individuals until the most recent common ancestor of all individuals is found

Backward-time simulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

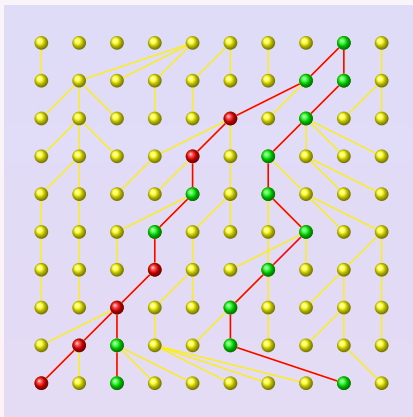
Forward- and backward-time simulation

Features of simuPOP

Applications
Availability

An example

simuPOP components



- Start from a sample with unknown genotype
- Coalesce individuals until the most recent common ancestor of all individuals is found
- Starting from the MRCA, proceed forward in time and fill the genotype of each individual

Forward vs. backward-time simulations

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

Backward-time

- Sample based, efficient.

Forward-time

- Population based,
inefficient.

Forward vs. backward-time simulations

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

Backward-time

- Sample based, efficient.
- Limited selection, recombination models and mating schemes

Forward-time

- Population based, inefficient.
- Can simulate almost arbitrary evolutionary scenarios

Forward vs. backward-time simulations

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

Backward-time

- Sample based, efficient.
- Limited selection, recombination models and mating schemes
- Can not study population properties, or properties of ancestral generations

Forward-time

- Population based, inefficient.
- Can simulate almost arbitrary evolutionary scenarios
- Can study population properties and ancestral generations

Forward vs. backward-time simulations

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

Backward-time

- Sample based, efficient.
- Limited selection, recombination models and mating schemes
- Can not study population properties, or properties of ancestral generations
- Used mostly for sample generation

Forward-time

- Population based, inefficient.
- Can simulate almost arbitrary evolutionary scenarios
- Can study population properties and ancestral generations
- Wider application area

Forward-time simulation programs

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

For specific applications

- Easy to write simple simulations
- Difficult to write complicated simulations
- A few programs are available (EasyPOP, FPG, Nemo, ...), easy to use if they happen to fit your need

Forward-time simulation programs

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

For specific applications

- Easy to write simple simulations
- Difficult to write complicated simulations
- A few programs are available (EasyPOP, FPG, Nemo, ...), easy to use if they happen to fit your need

For general purposes

- Difficult to write
- Easy to set up complicated simulations
- simuPOP fits in this category

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

1 What is simuPOP

- Forward- and backward-time simulation
- Features of simuPOP
- Applications
- Availability

What simuPOP does

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

simuPOP provides

- a large number of functions to manipulate populations
copy, split, merge, manipulate individual genotypes,
determine affection status, save to and load from
various formats, generate sample, ...

What simuPOP does

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

simuPOP provides

- a large number of functions to manipulate populations copy, split, merge, manipulate individual genotypes, determine affection status, save to and load from various formats, generate sample, ...
- and a mechanism to evolve populations forward in time subject to arbitrary demographic and genetic forces such as population size changes, mutation, migration, recombination, selection, ...

What distinguishes simuPOP from others

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications
Availability

An example

simuPOP components

scripting simuPOP is provided as a set of Python modules, and is therefore backed by a full-blown object-oriented programming language.

flexibility simuPOP does not impose any limit on the size of genome, population, demographic model, etc. Using a large number of standard and hybrid (Python-assisted) operators, users can simulate almost arbitrarily complex evolutionary processes.

integration Owing to the 'glue language' nature of Python, it is easy to integrate simuPOP with other languages and programs.

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

1 What is simuPOP

- Forward- and backward-time simulation
- Features of simuPOP
- Applications
- Availability

This is fun, but is it useful?

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

simuPOP can simulate the change of the genetic composition of a population in a complicated evolutionary process. It can be used to

- Demonstrate population genetics phenomena

This is fun, but is it useful?

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

simuPOP can simulate the change of the genetic composition of a population in a complicated evolutionary process. It can be used to

- Demonstrate population genetics phenomena
- Study the impact of genetic and demographic forces on the evolution of a population

This is fun, but is it useful?

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

simuPOP can simulate the change of the genetic composition of a population in a complicated evolutionary process. It can be used to

- Demonstrate population genetics phenomena
- Study the impact of genetic and demographic forces on the evolution of a population
- Study the evolution of (complex) genetic diseases

This is fun, but is it useful?

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

simuPOP can simulate the change of the genetic composition of a population in a complicated evolutionary process. It can be used to

- Demonstrate population genetics phenomena
- Study the impact of genetic and demographic forces on the evolution of a population
- Study the evolution of (complex) genetic diseases
- Simulate samples to validate gene-mapping methods

This is fun, but is it useful?

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

simuPOP can simulate the change of the genetic composition of a population in a complicated evolutionary process. It can be used to

- Demonstrate population genetics phenomena
- Study the impact of genetic and demographic forces on the evolution of a population
- Study the evolution of (complex) genetic diseases
- Simulate samples to validate gene-mapping methods
- Study ascertainment methods in simulated populations

This is fun, but is it useful?

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

simuPOP can simulate the change of the genetic composition of a population in a complicated evolutionary process. It can be used to

- Demonstrate population genetics phenomena
- Study the impact of genetic and demographic forces on the evolution of a population
- Study the evolution of (complex) genetic diseases
- Simulate samples to validate gene-mapping methods
- Study ascertainment methods in simulated populations
- ...

On the simulations of complex human diseases

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

Backward-time

- Haploid simulation only

Forward-time

- No limit on ploidy

On the simulations of complex human diseases

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

Backward-time

- Haploid simulation only
- Additive selection and penetrance models

Forward-time

- No limit on ploidy
- Arbitrary selection and penetrance models

On the simulations of complex human diseases

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

Backward-time

- Haploid simulation only
- Additive selection and penetrance models
- One disease susceptibility locus

Forward-time

- No limit on ploidy
- Arbitrary selection and penetrance models
- Multiple DSL with interaction

On the simulations of complex human diseases

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

Backward-time

- Haploid simulation only
- Additive selection and penetrance models
- One disease susceptibility locus
- Generate independent samples

Forward-time

- No limit on ploidy
- Arbitrary selection and penetrance models
- Multiple DSL with interaction
- Generate multi-generation populations

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

1 What is simuPOP

- Forward- and backward-time simulation
- Features of simuPOP
- Applications
- Availability

I like it, but, oohm, why Python??

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example

simuPOP components

- The core of simuPOP is written in C++ for efficiency
- Python is the glue language, a wrapper of the core
- Python is used to write simuPOP extensions (user interface etc)
- The core sometimes calls Python (Python operators) for maximum flexibility

Availability

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

Forward- and
backward-time
simulation

Features of
simuPOP

Applications

Availability

An example simuPOP components

- simuPOP website:
`http://simupop.sourceforge.net`
- Mailing list:
`simupop-list@lists.sourceforge.net`
- License: GPL 2.0
- Platforms: all OS on which Python is available
- Monthly release, currently at 0.7.10
- Documentation: *simuPOP User's Guide* and *simuPOP Reference Manual*

Outline

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

An example

Visualization with R

simuPOP
components

2

An example

- An example
- Visualization with R

A simple example

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example
Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
>>> simu.evolve(
...     preOps = [initByValue([1,2,2,1])],
...     ops = [
...         recombinator(rate=0.1),
...         stat(LD=[0,1]),
...         pyEval(r"'%3d' % gen", rep=0, step=10),
...         pyEval(r"'%f' % LD[0][1]", step=10),
...         pyEval(r"\n", rep=REP_LAST, step=10)
...     ],
...     end=100
... )
```

Output of the example

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example
Visualization with R

simuPOP components

0	0.200605	0.198696	0.200315
10	0.064690	0.073950	0.058101
20	0.033841	0.022740	0.019255
30	0.002388	0.008523	0.002652
40	0.011979	0.025504	0.021945
50	0.000420	0.006112	0.002299
60	0.014547	0.009627	0.000158
70	0.019284	0.000843	0.002107
80	0.014770	0.000300	0.003610
90	0.008108	0.006058	0.000223
100	0.006191	0.004340	0.003232

simuPOP modules

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example
Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
```

Import the default simuPOP module

population

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example

Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
```

Create a **population** of 1000 **diploid** individuals, each having two **loci** on the first chromosome

simulator and mating scheme

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example

Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
```

Create a **simulator** that has one replicate of this population, and a random mating scheme

Operators!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example

Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
>>> simu.evolve(
...     preOps = [initByValue([1,2,2,1])],
...     ops = [
...         recombinator(rate=0.1),
...         stat(LD=[0,1]),
...         pyEval(r"'%3d    ' % gen", rep=0, step=10),
...         pyEval(r"'%f    ' % LD[0][1]", step=10),
...         pyEval(r"'\\n'", rep=REP_LAST, step=10)
...     ],
...     end=100
... )
```

`initByValue` is applied before evolution

Operators!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example
Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
>>> simu.evolve(
...     preOps = [initByValue([1,2,2,1])],
...     ops = [
...         recombinator(rate=0.1),
...         stat(LD=[0,1]),
...         pyEval(r"%3d    ' % gen", rep=0, step=10),
...         pyEval(r"%f    ' % LD[0][1]", step=10),
...         pyEval(r"\n", rep=REP_LAST, step=10)
...     ],
...     end=100
... )
```

recombinator is applied at every generation when an offspring is produced

Operators!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example
Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
>>> simu.evolve(
...     preOps = [initByValue([1,2,2,1])],
...     ops = [
...         recombinator(rate=0.1),
...         stat(LD=[0,1]),
...         pyEval(r"%3d    ' % gen", rep=0, step=10),
...         pyEval(r"%f    ' % LD[0][1]", step=10),
...         pyEval(r"'\\n'", rep=REP_LAST, step=10)
...     ],
...     end=100
... )
```

stat is applied to the offspring generation at every generation

Operators!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example

Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
>>> simu.evolve(
...     preOps = [initByValue([1,2,2,1])],
...     ops = [
...         recombinator(rate=0.1),
...         stat(LD=[0,1]),
...         pyEval(r"'%3d    ' % gen", rep=0, step=10),
...         pyEval(r"'%f    ' % LD[0][1]", step=10),
...         pyEval(r"'\\n'", rep=REP_LAST, step=10)
...     ],
...     end=100
... )
```

pyEval is applied every 10 generations

Outline

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

An example

Visualization with R

simuPOP
components

2

An example

- An example
- Visualization with R

Use R to plot

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example
Visualization with R

simuPOP components

```
>>> from simuPOP import *
>>> from simuRPy import *
>>> simu = simulator(
...     population(size=1000, ploidy=2, loci=[2]),
...     randomMating(),
...     rep = 3)
>>> simu.evolve(
...     preOps = [initByValue([1,2,2,1])],
...     ops = [
...         recombinator(rate=0.1),
...         stat(LD=[0,1]),
...         varPlotter('LD[0][1]', numRep=3, step=10,
...             saveAs='ld', ylim=[0,.25],
...             lty=range(1, 4), col=range(2, 5),
...             xlab='generation', ylab='D',
...             title='LD Decay'),
...     ],
...     end=100
... )
True
>>>
```

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

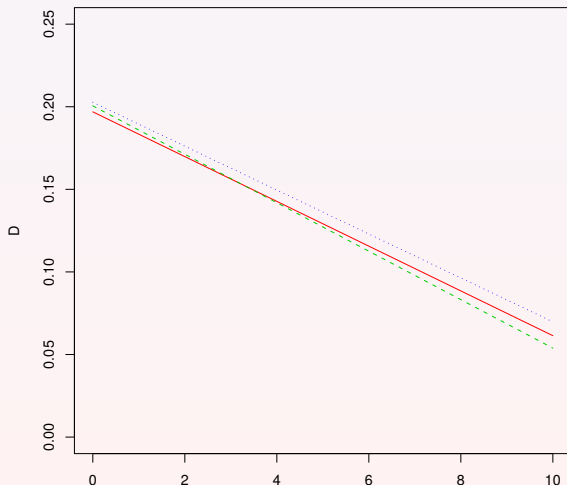
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

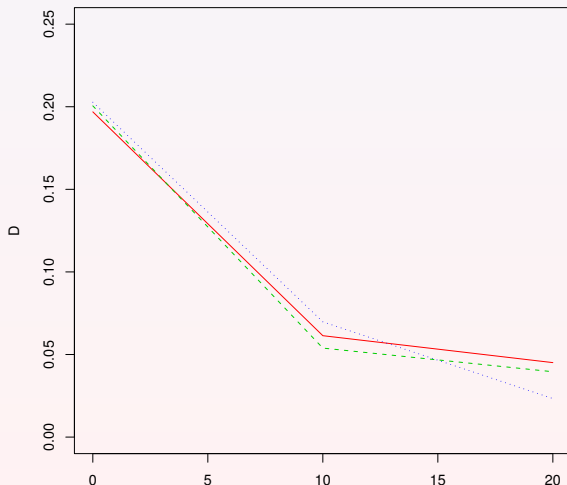
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

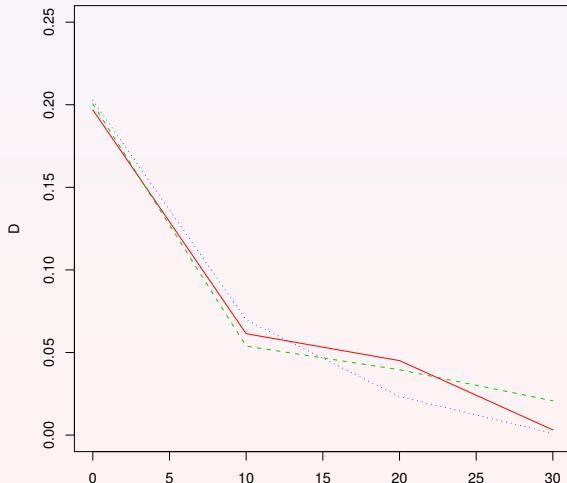
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

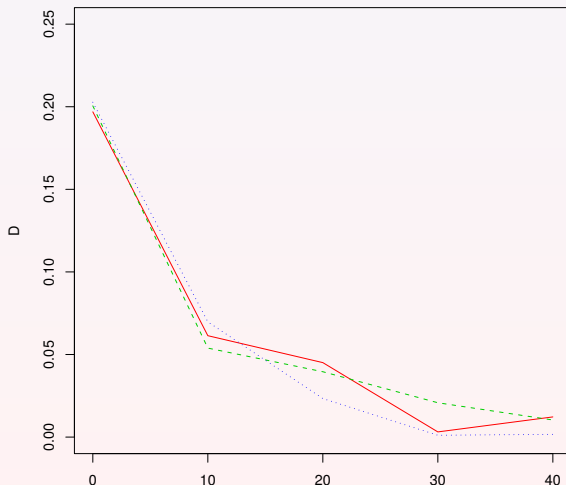
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

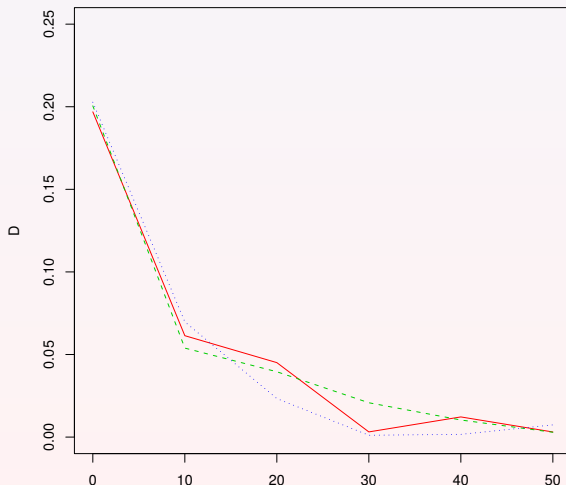
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

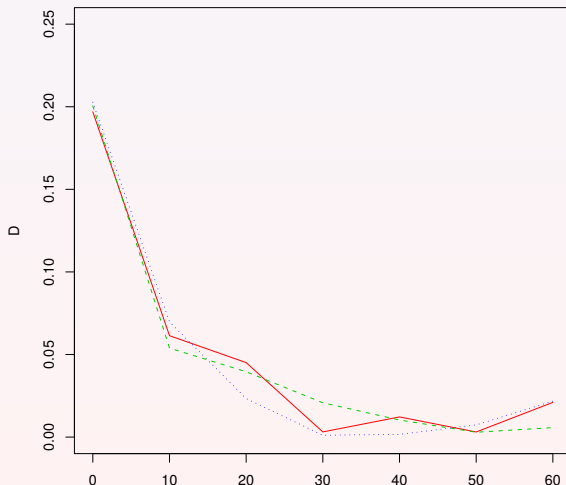
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

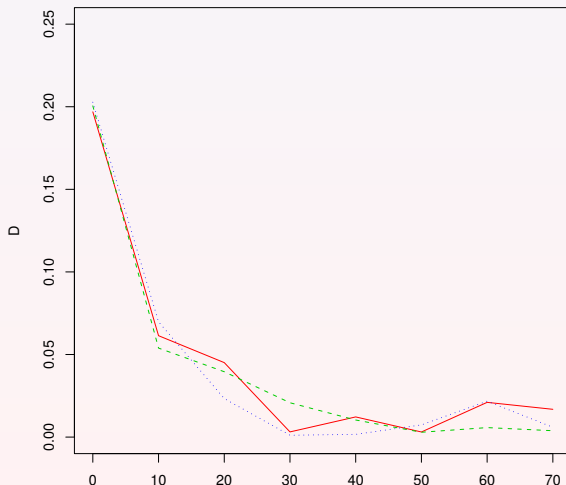
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

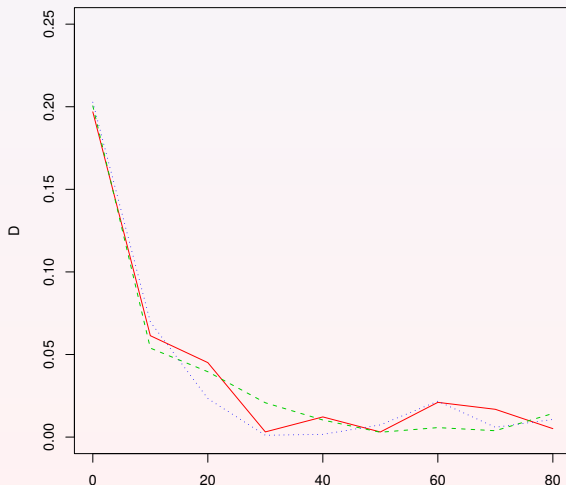
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

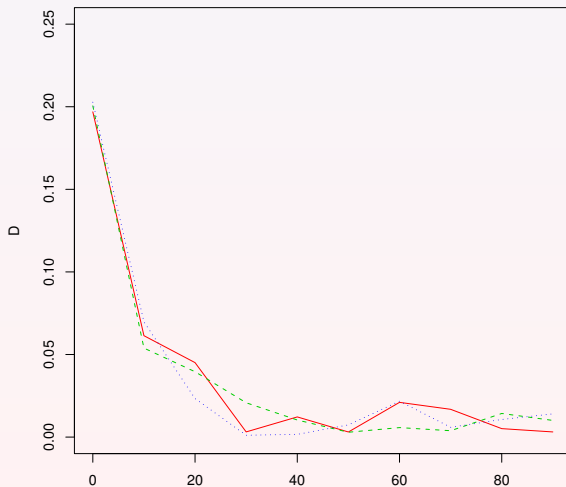
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Evolve!

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

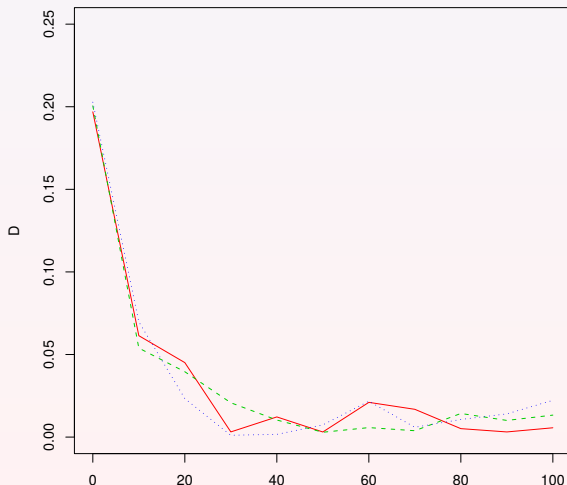
An example

An example

Visualization with R

simuPOP components

LD Decay



- Update at every 10 generations
- $LD=0.25$ before generation 0
- LD calculated at the end of each generation

Exercise time

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

An example
Visualization with R

simuPOP components

- Start python
- Load simuPOP
- Create a population and run

```
pop.ploidyName( )
```

- run tutorial_example1.py

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

3 simuPOP components

- Population
- Individual
- Operator
- Mating scheme
- Simulator

Structure of a population

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

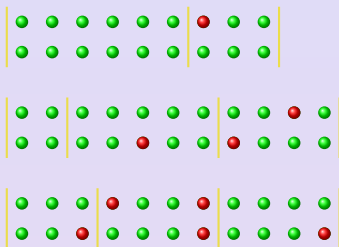
Individual

Operator

Mating scheme

Simulator

- Unaffected
- Affected



gen = 2, numAffected = 5, ...

Structure of a population

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

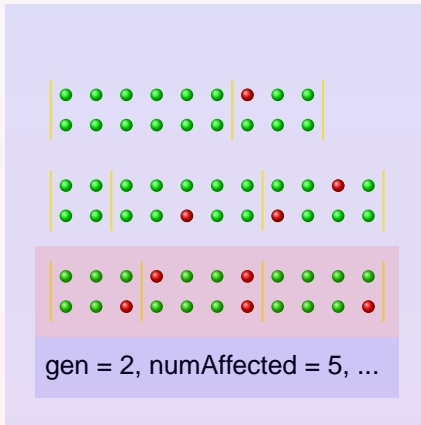
Individual

Operator

Mating scheme

Simulator

- Unaffected
- Affected



Current generation

Structure of a population

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

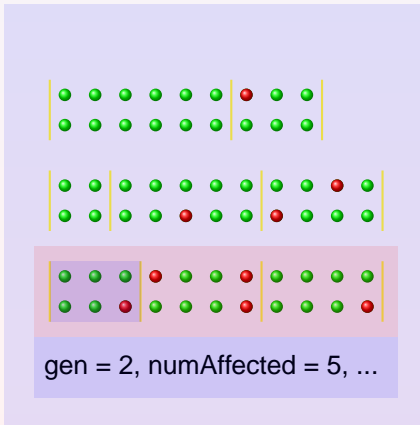
Individual

Operator

Mating scheme

Simulator

- Unaffected
- Affected



Current generation

Structure of a population

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

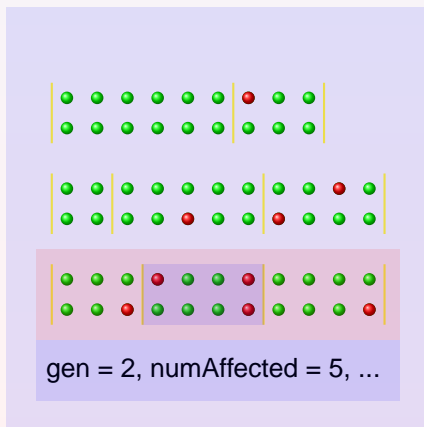
Individual

Operator

Mating scheme

Simulator

- Unaffected
- Affected



Current generation

Structure of a population

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

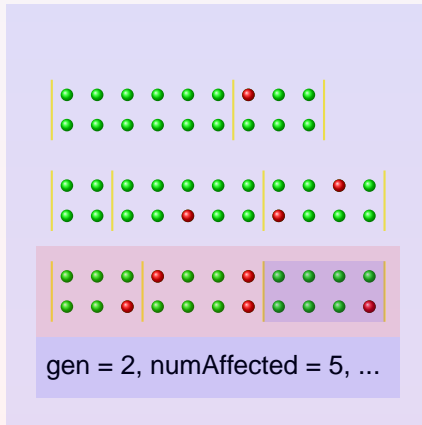
Individual

Operator

Mating scheme

Simulator

- Unaffected
- Affected



Current generation

Structure of a population

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

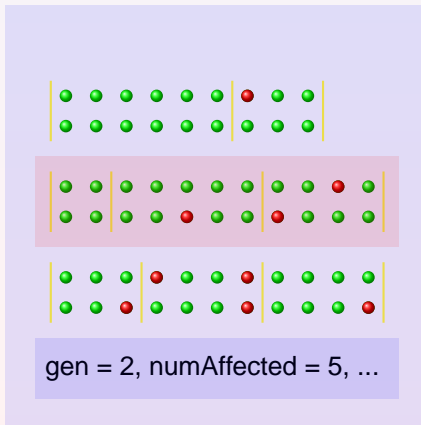
Individual

Operator

Mating scheme

Simulator

- Unaffected
- Affected



Ancestral generation 1

Current generation

Structure of a population

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

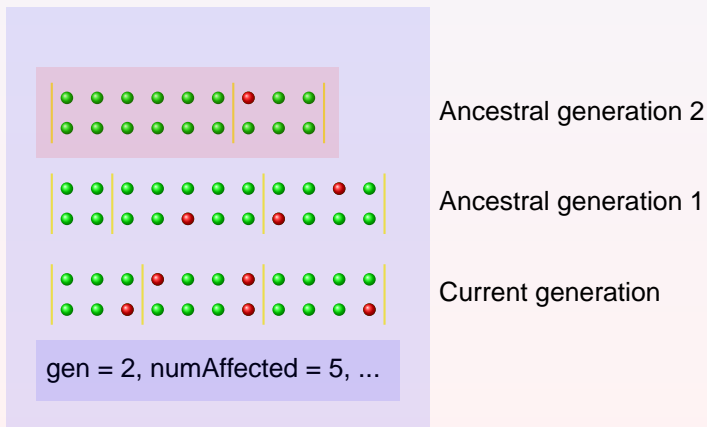
Individual

Operator

Mating scheme

Simulator

- Unaffected
- Affected



Structure of a population

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

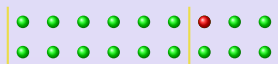
Individual

Operator

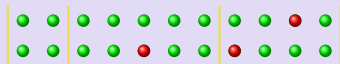
Mating scheme

Simulator

- Unaffected
- Affected



Ancestral generation 2



Ancestral generation 1



Current generation

gen = 2, numAffected = 5, ...

Population variables

Create and manipulate populations

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> pop = population(size=10, loci=[2, 3])
>>> Dump(pop)
Ploidy:                2
Number of chrom:       2
Number of loci:        2 3
Maximum allele state:   255
Loci positions:
                1 2
                1 2 3
Loci names:
                loc1-1 loc1-2
                loc2-1 loc2-2 loc2-3
population size:       10
Number of subPop:      1
Subpop sizes:         10
Number of ancestral populations: 0
individual info:
sub population 0:
    0: MU    0 0 0 0 0 | 0 0 0 0 0
    1: MU    0 0 0 0 0 | 0 0 0 0 0
    2: MU    0 0 0 0 0 | 0 0 0 0 0
    3: MU    0 0 0 0 0 | 0 0 0 0 0
```


Genotypic structure

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> pop = population(subPop=[200, 300], loci=[3, 2],
...     maxAllele=3, ploidy=3,
...     lociPos=[[1, 3, 5], [2.5, 4]],
...     alleleNames=['A', 'C', 'T', 'G'])
>>> pop.numLoci(0)
3
>>> pop.totNumLoci()
5
>>> pop.locusPos(4)
4.0
>>> pop.subPopSize(1)
300
>>> pop.popSize()
500
>>> pop.ploidyName()
'triploid'
>>> pop.individual(1).allele(1, 2)
0
>>>
```

Population manipulation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> # make a copy of pop
>>> pop1 = pop.clone()
>>> # remove loci 2, 3, 4
>>> pop.removeLoci(keep=[0, 1])
>>> # pop2 will have 3 chromosomes, with loci 2, 3, 2
>>> pop2 = MergePopulationsByLoci(pops=[pop, pop1])
>>> # randomly assign alleles using given allele frequencies
>>> InitByFreq(pop2, [0.8, .2])
>>> # calculate population allele frequency
>>> Stat(pop2, alleleFreq=range(pop2.totNumLoci()))
>>> # print allele frequency
>>> print pop2.dvars().alleleFreq
[[0.81066666666666665, 0.18933333333333333], [0.7853333333333333, 0.21466666666666667]]
>>> # assign affection status using a penetrance model
>>> MapPenetrance(pop2, locus=1,
...               penetrance={'0-0': 0.05, '0-1': 0.2, '1-1': 0.8})
>>> # draw case control sample
>>> (sample,) = CaseControlSample(pop2, cases=5, controls=5)
>>> # save sample in Merlin QTDT format
>>> from simuUtil import SaveQTDT
>>> SaveQTDT(sample, output='sample', affectionCode=['U', 'A'],
...          fields=['affection'])
```

Population manipulation (cont.)

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> # have a look at the sample in Merlin-QTDT Format
>>> print open('sample.map').read()
CHROMOSOME MARKER POSITION
1      loc1-1  1.000000
1      loc1-2  3.000000
2      loc1-1_1      1.000000
2      loc1-2_1      3.000000
2      loc1-3  5.000000
3      loc2-1  2.500000
3      loc2-2  4.000000

>>> print open('sample.dat').read()
A      affection
M      loc1-1
M      loc1-2
M      loc1-1_1
M      loc1-2_1
M      loc1-3
M      loc2-1
M      loc2-2
```

Population manipulation (cont.)

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> print open('sample.ped').read()
1 1 0 0 2 A 1 1 1 1 2 1 1 2 1 1 1 1 1 1
2 1 0 0 1 A 1 1 1 2 1 1 1 1 1 2 2 1 1 1
3 1 0 0 2 A 1 1 1 2 2 1 1 1 2 1 1 2 1 1
4 1 0 0 1 A 2 1 1 2 1 1 1 2 1 2 1 1 2 1
5 1 0 0 1 A 2 1 2 1 1 1 2 1 1 1 1 1 1 1
6 1 0 0 1 U 2 1 2 1 1 1 1 1 1 2 2 1 2 1
7 1 0 0 1 U 1 1 1 1 1 2 2 2 1 1 2 1 1 1
8 1 0 0 1 U 2 2 1 1 1 1 1 1 1 1 1 1 2
9 1 0 0 2 U 2 1 1 2 1 1 2 1 1 1 2 1 2 1
10 1 0 0 2 U 2 1 1 1 1 2 1 1 1 1 1 1 1 1

>>>
```

Population variables

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> pop = population(subPop=[5, 10], loci=[5])
>>> InitByFreq(pop, [.6, .3, .1])
>>> Stat(pop, alleleFreq=[1], genoFreq=[2])
>>> print pop.dvars().alleleFreq[1][0]
0.4666666666667
>>> from simuUtil import ListVars
>>> ListVars(pop.dvars(), useWxPython=False)
grp : -1
rep : -1
alleleNum :
  [1]
    [0]      14
    [1]      12
    [2]       4
genoFreq :
  [2]
    [0]
      0 : 0.333333333333
      1 : 0.333333333333
      2 : 0.133333333333
    [1]
      2 : 0.133333333333
    [2]
      2 : 0.0666666666667
genoNum :
  [2]
    [0]
      0 : 5.0
      1 : 5.0
      2 : 2.0
  [1]
```

Population variables (cont.)

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
[1]
[0]      0.46666666666667
[1]      0.4
[2]      0.13333333333333
subPop
[0]
  alleleNum :
    [1]
      [0]  4
      [1]  5
      [2]  1
  genoNum :
    [2]
      [0]
        0 : 2.0
        1 : 2.0
      [1]
      [2]
        2 : 1.0
  genoFreq :
    [2]
      [0]
        0 : 0.4
        1 : 0.4
      [1]
      [2]
```

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

3 simuPOP components

- Population
- Individual
- Operator
- Mating scheme
- Simulator

Structure of Individuals

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

Assume ploidy = 2, maxAllele = 1

0	1	2	3	4	5	6
0	1	1	1	0	0	1
0	0	1	1	1	0	1

0	1	2	3	4	5
0	1	0	0	0	1
1	0	1	1	0	0

Male

● Affected

fitness | father_id | ...

Structure of Individuals

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

Assume ploidy = 2, maxAllele = 1

0	1	2	3	4	5	6
0	1	1	1	0	0	1
0	0	1	1	1	0	1

Chromosome 0

0	1	2	3	4	5
0	1	0	0	0	1
1	0	1	1	0	0

Male

● Affected

fitness	father_id	...
---------	-----------	-----

Structure of Individuals

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

Assume ploidy = 2, maxAllele = 1

0	1	2	3	4	5	6
0	1	1	1	0	0	1
0	0	1	1	1	0	1

Chromosome 0

0	1	2	3	4	5
0	1	0	0	0	1
1	0	1	1	0	0

Chromosome 1

Male

● Affected

fitness	father_id	...
---------	-----------	-----

Structure of Individuals

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

Assume ploidy = 2, maxAllele = 1

0	1	2	3	4	5	6
0	1	1	1	0	0	1
0	0	1	1	1	0	1

Chromosome 0

0	1	2	3	4	5
0	1	0	0	0	1
1	0	1	1	0	0

Chromosome 1

Male

Sex

● Affected

fitness | father_id | ...

Structure of Individuals

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

Assume ploidy = 2, maxAllele = 1

0	1	2	3	4	5	6
0	1	1	1	0	0	1
0	0	1	1	1	0	1

Chromosome 0

0	1	2	3	4	5
0	1	0	0	0	1
1	0	1	1	0	0

Chromosome 1

Male

Sex

● Affected

Affection status

fitness | father_id | ...

Structure of Individuals

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

Assume ploidy = 2, maxAllele = 1

0	1	2	3	4	5	6
0	1	1	1	0	0	1
0	0	1	1	1	0	1

Chromosome 0

0	1	2	3	4	5
0	1	0	0	0	1
1	0	1	1	0	0

Chromosome 1

Male

Sex

● Affected

Affection status

fitness | father_id | ...

Information
fields

Individuals

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> pop = population(subPop=[5, 8], loci=[5],
...                   infoFields=['penetrance'])
>>> InitByFreq(pop, [.6, .3, .1])
>>> MaPenetrance(pop, locus=2, penetrance=[0.05, 0.2, 0.5],
...              wildtype=[0], infoFields=['penetrance'])
>>> # iterate through all individuals in subPop 1
>>> for ind in pop.individuals(1):
...     print 'Aff: %d Fit: %.3f Geno: %d %d' % \
...           (ind.affected(), ind.info('penetrance'), \
...            ind.allele(2, 0), ind.allele(2, 1))
...
Aff: 0 Fit: 0.200 Geno: 0 1
Aff: 0 Fit: 0.500 Geno: 1 1
Aff: 1 Fit: 0.200 Geno: 0 1
Aff: 0 Fit: 0.200 Geno: 0 2
Aff: 0 Fit: 0.200 Geno: 0 1
Aff: 1 Fit: 0.500 Geno: 1 1
Aff: 0 Fit: 0.050 Geno: 0 0
Aff: 0 Fit: 0.500 Geno: 2 1
>>>
```

Information fields

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> pop = population(100, loci=[5, 8],  
...     infoFields=['father_idx', 'mother_idx'])  
>>> simu = simulator(pop, randomMating(numOffspring=2))  
>>> simu.evolve(ops=[parentsTagger()], end=5)  
True  
>>> ind = simu.population(0).individual(0)  
>>> ind1 = simu.population(0).individual(1)  
>>> print ind.info('father_idx'), ind.info('mother_idx')  
59.0 50.0  
>>> print ind1.info('father_idx'), ind1.info('mother_idx')  
59.0 50.0  
>>>  
>>>
```

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

3 simuPOP components

- Population
- Individual
- Operator
- Mating scheme
- Simulator

Life cycle of a generation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

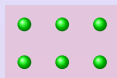
Population

Individual

Operator

Mating scheme

Simulator



Parental
generation

Every operator has a default **stage**, and a **stage** parameter to change it.

Life cycle of a generation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

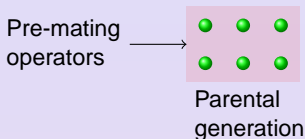
Population

Individual

Operator

Mating scheme

Simulator



Every operator has a default **stage**, and a **stage** parameter to change it.

Life cycle of a generation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

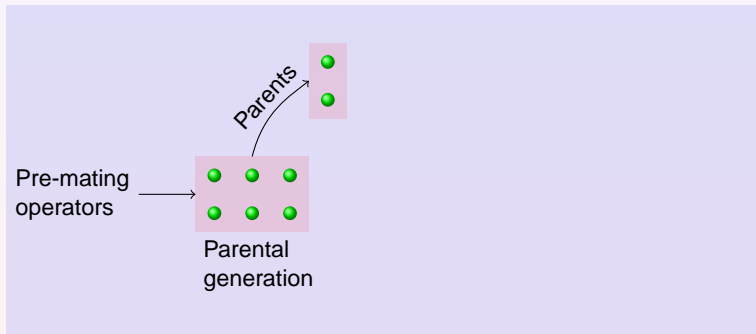
Population

Individual

Operator

Mating scheme

Simulator



Every operator has a default **stage**, and a **stage** parameter to change it.

Life cycle of a generation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

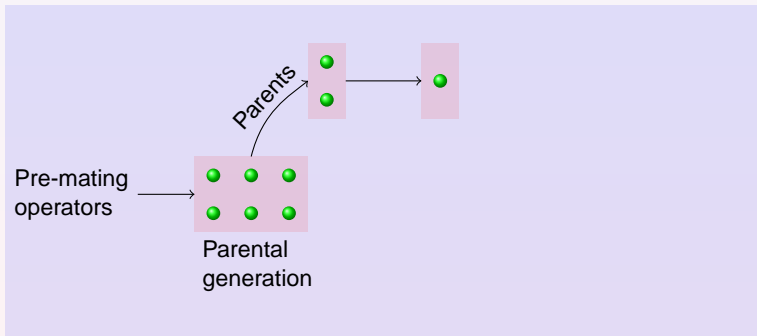
Population

Individual

Operator

Mating scheme

Simulator



Every operator has a default **stage**, and a **stage** parameter to change it.

Life cycle of a generation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

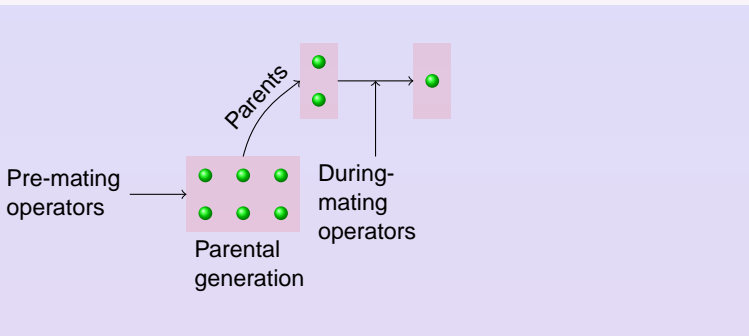
Population

Individual

Operator

Mating scheme

Simulator



Every operator has a default **stage**, and a **stage** parameter to change it.

Life cycle of a generation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

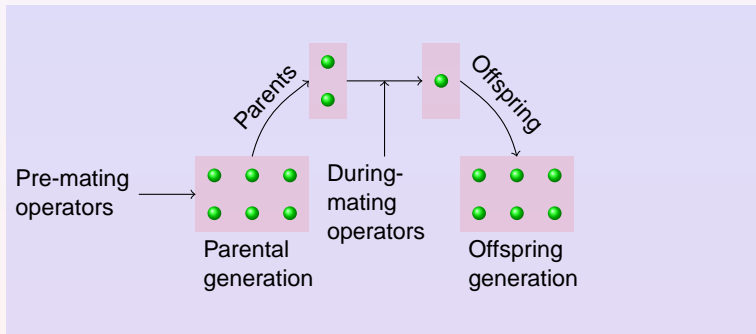
Population

Individual

Operator

Mating scheme

Simulator



Every operator has a default **stage**, and a **stage** parameter to change it.

Life cycle of a generation

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

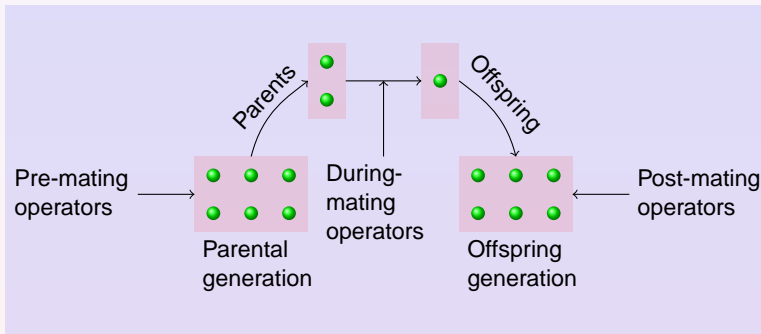
Population

Individual

Operator

Mating scheme

Simulator



Every operator has a default **stage**, and a **stage** parameter to change it.

Pre-, During- and Post- operators

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> simu = simulator(
...     population(subPop=[20, 80], loci=[3]),
...     randomMating())
>>> simu.evolve(
...     preOps = [initByFreq([0.2, 0.8])],
...     ops = [
...         kamMutator(maxAllele=10, rate=0.00005, atLoci=[0,2]),
...         recombinator(rate=0.001),
...         dumper(stage=PrePostMating),
...         stat(alleleFreq=[1]),
...     ],
...     dryrun=True
... )
```

Dryrun mode: display calling sequence

Apply pre-evolution operators

Replicate 0

- <simuPOP::initByFreq> end at 1

Start evolution

Replicate 0

Pre-mating operators

- <simuPOP::dumper> at all generations

Start mating

- <simuPOP::recombination> at all generations

Apply post-mating operators

- <simuPOP::k-allele model mutator K=10> at all generations

- <simuPOP::dumper> at all generations

- <simuPOP::statistics> at all generations

True

>>>

Applicable generations

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> simu = simulator(
...     population(10000, loci=[3]),
...     randomMating())
>>> eval1 = r"'Gen: %3d  Freq: %f\n' % (gen, alleleFreq[1][0])"
>>> eval2 = r"'Last Gen: %3d  Freq: %s\n' % (gen, alleleFreq[1])"
>>> simu.evolve(
...     preOps = [initByFreq([0.3, 0.7])],
...     ops = [
...         recombinator(rate=0.01, begin=10, end=30),
...         stat(alleleFreq=[1], step=10),
...         pyEval(eval1, step=10),
...         pyEval(eval2, at=[-1])
...     ],
...     end = 50
... )
Gen:    0  Freq: 0.311250
Gen:   10  Freq: 0.306550
Gen:   20  Freq: 0.316050
Gen:   30  Freq: 0.311650
Gen:   40  Freq: 0.301850
Gen:   50  Freq: 0.303200
Last Gen:  50  Freq: [0.30320000000000003, 0.69679999999999997]
True
```

Applicable replicates

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

```
>>> simu = simulator(
...     population(100, loci=[3]),
...     randomMating(),
...     rep=5, grp=[1,1,2,2,2])
>>> simu.evolve(
...     preOps = [initByFreq([0.5, 0.5])],
...     ops = [
...         stat(alleleFreq=[1]),
...         recombinator(rate=0.01, grp=1),
...         recombinator(rate=0.01, grp=2),
...         pyEval(r"'%.2f ' % alleleFreq[1][0]", grp=1),
...         pyEval(r"'\\n'", rep=REP_LAST),
...     ],
...     end=5
... )
0.54 0.52
0.62 0.51
0.61 0.54
0.65 0.56
0.62 0.57
0.59 0.64
True
>>>
```

Output

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

Table-like output

simuPOP
tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

3 simuPOP components

- Population
- Individual
- Operator
- Mating scheme
- Simulator

Mating schemes

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator

Outline

simuPOP tutorial

Bo Peng,
Ph.D.

What is
simuPOP

An example

simuPOP
components

Population

Individual

Operator

Mating scheme

Simulator

3 simuPOP components

- Population
- Individual
- Operator
- Mating scheme
- Simulator

Simulator

simuPOP tutorial

Bo Peng,
Ph.D.

What is simuPOP

An example

simuPOP components

Population

Individual

Operator

Mating scheme

Simulator