

Project n°1 : Genetic of Populations (Wright-Fisher) ReadMe

[Team n°1]

I. What the program does

The main goal of this program is to simulate the temporal evolution of a population represented by N individuals with a genome. In order to do that, we will follow the model of Wright-Fisher. There is a fixed number A of alleles in the population and each individual has a type of allele. At each generation, we will sort a new population with a size N among the individuals of the last generation according to the multinomial distribution. The alleles of this population can mutate during the simulation (with mutation coefficients) or can be more or less selected (with fitness coefficients).

II. The software you need to install

- CMake (compilation)
- Doxygen (documentation)
- Gtest (tests)
- TCLAP (command on terminal)

III. How to execute the program

First, after cloning the project, you need to create a folder “build” and compile the program in this folder with the commands : “cmake ..” and then “make” before starting the program.

You have two possible ways to perform the simulation : with or without a fasta file. If you do not want to work with a fasta file you need to enter all the factors the simulation needs. :

- The time of simulation or the number of generations with **-T** (default value : 10)
- The number of repetitions of the simulation with **-R** (default value : 2)
- The size of the population with **-N** (default value : 100)

- The number of alleles of the population with **-A** (default value : 2)
- The frequencies of the alleles with **-f** (no default value) -
- The fitness coefficient of the alleles with **-S** (default value : 0 for each allele)

If you to use a fasta file to perform the simulation, you need to enter these factors :

- The time of simulation or the number of generations with **-T** (default value : 10)
- The number of repetitions of the simulation with **-R** (default value : 2)
- The name of the fasta file with **-F** (default value : ../tests/test_for_retrieveData.fasta)
- The marks of the 3 nucleotides of the alleles with **-m** (no default value)
- The mutation rate of the alleles with **-M** (default value : 0 for each allele)
- The mutations sites with **-m** (default value : 0 for each allele)
- The fitness coefficient of the alleles with **-S** (default value : 0 for each allele)
- The delta coefficient of the mutation with **-d** (default value : 1/3)

Examples of command :

- *Without a fasta file and with selection :*

```
./PopulationGenetic -R 4 -T 100 -A 4 -f 0.1 -f 0.1 -f 0.3 -f 0.5 -N 100 -S -1 -S 0 -S -1 -S 0.4
```

- *With a fasta file and mutations and without selection :*

```
./PopulationGenetic -F ../test.fa -m 2 -m 4 -m 8 -T 100 -R 2 -M 0.01 -M 0.02 -M 0.5 -m 2 -  
m 4 -m 8 -d 0.5
```