Tutorial for MultiWaveInfer v1.0.1

==

Short Description:

MultiWaveInfer is designed to estimate the parameters of multi-waves, multi-ancestral populations admixture models by the ancestral tracks. The program proceed in two steps: Firstly, use EM-algorithm to identify the waves for each ancestral population. Secondly, use the theoretical length distribution of ancestral tracks to estimate the parameters.

1.Compile

1.1 Compile from source code

It's very easy to compile from the source code by the following commands:

bash$ *tar -zvxf MultiWaveInfer.tar.gz*

bash$ *cd MultiWaveInfer/src*

bash$ *make*

After compiling, you will get the executable MultiWaveInfer, just typing the command below to get help information:

bash$ *./MultiWaveInfer -h or bash$ ./MultiWaveInfer --help*

2. Test with the toy data

2.1 two simple examples

bash$ *./MultiWaveInfer --input ../example/two.se*g

Example explanation:

MultiWaveInfer will read the ancestral tracks from two.seg, after a while, the optimal model and corresponding generation and proportion will print to screen. The format will explained later.

The following is output of the toy data:

```
// COMMAND ./MultiWaveInfer -i ../example/two.seg

Reading data from ../example/two.seg...
Start scan for admixture waves...
Perform EM scan for waves of population 2...
```

Perform EM scan for waves of population 1...
Finished scanning for admixture waves.

There is(are) 2 waves of admixture event(s) detected

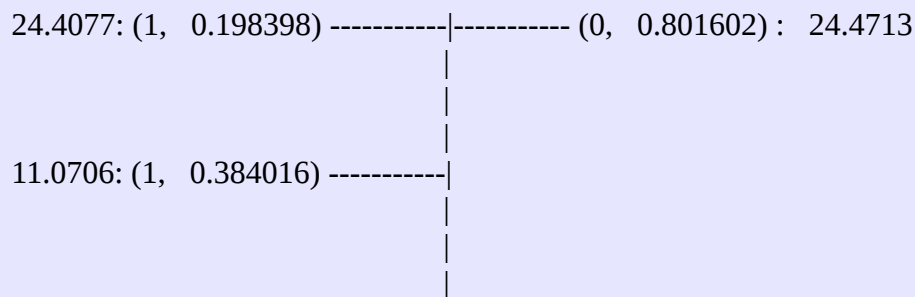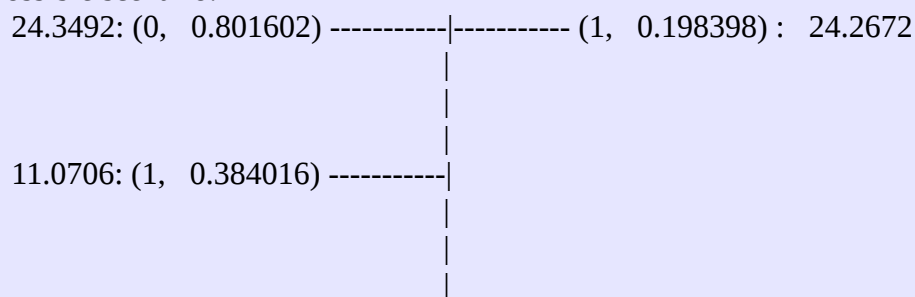   Results summary

            Parental population   Admixture proportion
                     1   0.506226
                     2   0.493774

Possible scenario: #1
  24.4077: (1,   0.198398) -----------|----------- (0,   0.801602) :   24.4713
                                      |
                                      |
                                      |
  11.0706: (1,   0.384016) -----------|
                                      |
                                      |
                                      |

Possible scenario: #2
  24.3492: (0,   0.801602) -----------|----------- (1,   0.198398) :   24.2672
                                      |
                                      |
                                      |
  11.0706: (1,   0.384016) -----------|
                                      |
                                      |
                                      |

Hint:
0: population-2; 1: population-1;

We use a tree to present the results. The simulated admixed population has two ancestral populations (population 1 and 2). There are 2 waves of admixture events. The first admixture event was happened in 24.4715 generations ago. The ancestral populations are pop2 and pop1 and corresponding mixture proportions are 0.198394 and 0.801606. The second admixture event was happend in 11.0706 generations ago. The ancestral population and corresponding mixture proportions is pop1 and 0.384019.

User can redirect the output to a file, such as:

```
bash$ ./MultiWaveInfer --input ../example/sim1.seg > sim1_opt.log
```

3. File format

3.1 Input file format

MultiWaveInfer is easy to use, only need one file, in which each line represents a ancestral track with the start point, end points, from which ancestry the track originates. The start and end points units are in Morgan.

For example:

```
0.00000000    0.34602058    Yoruba
0.34602058    0.34614778    French
......
0.40759031    0.41517938    Yoruba
```

4. Arguments

-i/--input   <string>

      This argument is required, in which user specify the filename of input ancestral tracks, format described above.

-l/--lower   [double]

      This argument is optional, in which user specify the lower bound to discard short tracks. The default is 0, which does not discard any short tracks. However, due to method limitation in local ancestry inference, very short tracks are generally not reliable.

-m/--maxIt   [integer]

      Max number of iterations to perform EM [optional, default 10000]

6. License

GNU GENERAL PUBLIC LICENSE Version 3

http://www.gnu.org/licenses/gpl-3.0.html

===

7. Questions and Suggestions

Questions and Suggestions are welcomed, feel free to contact

Shawn xyang619@gmail.com