

Manual for MultiWaver 2.0

MultiWaver series software was designed to infer population admixture history in case of various and complex scenarios. The earlier version of *MultiWaver* considered only the discrete admixture models. In the newly developed version *MultiWaver 2.0*, we implemented a more flexible framework to automatically select an optimal admixture model among discrete models and continuous models.

Four candidate admixture models now available for *MultiWaver 2.0*, which are general discrete admixture models, HI (Hybrid Isolation) model, CGF (Continuous Gene Flow) model and GA (Gradient Admixture) model. The software can automatically select a best-fit model and estimate the corresponding model parameters. We also applied a bootstrapping technique to provide a supporting rate of the chosen model and the confidence interval (CI) of estimations of parameters.

Function of *MultiWaver 1.0* and *AdmixInfer* were also implemented in the new version, which can be invoked with specific parameters.

1 Getting Started

1.1 Library dependency

MultiWaver 2.0 depend on the *OpenMP* Library (<https://www.openmp.org/>) to perform the bootstrap procedure parallelly.

Users can also compile the software without *OpenMP*. (See 1.2)

1.2 Compile

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

```
$tar -zxvf MultiWaver-2.0.tar.gz
$cd MultiWaver-2.0/src
$make all
```

To compile the software without *OpenMP* Library, just type:

```
$make FLAGS=-DNOOMP all
```

Then the executable software will be in the bin file. You may type the following command for help:

```
$MultiWaver-2.0/bin/MultiWaver2 -h
```

2 Input Files

Same as *MultiWaver-1.0* and *AdmixInfer*, only one input file required, in which each line represents an 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  0.22071382  1
0.22071382  0.23472265  2
0.23472265  0.27476200  1
...
2.41681269  2.49731957  1
```

3 Output Files

3.1 Results summary file

Best Model: CGF ("1" as donor)
Bootstrapping support ratio: 96% (96/100)
1: 96(G) 0.782809
2: 96(G) 0.217191

Bootstrapping details

CGF ("1" as donor)
Bootstrapping supporting ratio: 96% (96/100)
1: 94~98(G) 0.77383~0.792846
2: 94~98(G) 0.207154~0.22617

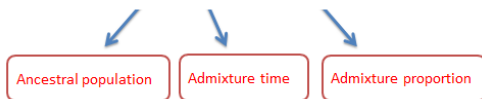
Multi 2-1
Bootstrapping supporting ratio: 3% (3/100)
1: 30~32(G) 0.505214~0.554851
1: 78~80(G) 0.483831~0.549552
2: 78~80(G) 0.450448~0.516169

Multi 2-2
Bootstrapping supporting ratio: 1% (1/100)
1: 28~28(G) 0.47409~0.47409
1: 93~93(G) 0.746509~0.746509
2: 52~52(G) 0.227058~0.227058
2: 93~93(G) 0.253491~0.253491

The example results summary file separated by “---“ line. The first part shows the best model (CGF) and its corresponding bootstrapping support ratio (96%). The following parts show bootstrapping summary for all possible models.

An explanation of numbers in the results

Best Model: CGF ("1" as donor)
Bootstrapping support ratio: 96% (96/100)
1: 96(G) 0.782809
2: 96(G) 0.217191



3.2 Log file

The log file contains “N+1” lines. The first line is the detail result for input data. The following N line is the detail of each bootstrapping. N is number of bootstrapping.

4 Arguments & options

-i/--input <string>

This argument is required, in which user specifies the filename of input ancestral tracks (See 2 Input Files).

-o/--output <string>

This argument is required, in which user specifies the filename of output result summary file (See 3.1 Results summary file).

`-g/--log <string>`

This argument is optional, in which user specifies the filename of output log file (See 3.2 Log file).

`-l/--lower [double]`

This argument is optional, in which user specifies 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. The unit of this argument is Morgan.

`-a/--alpha [double]`

This argument is optional, in which user specifies the significance level to reject null hypothesis in likelihood ratio test (LRT). Default is 0.001. This argument only affects the model selection of discrete models.

`-e/--epsilon [double]`

This argument is optional, in which user specifies epsilon to check whether parameters converge or not. Default is 1.0e-6. This argument only affects the discrete models.

`-p/--minProp [double]`

This argument is optional, in which user specifies the minimum survival proportion for a wave at the final generation. Default is 0.05. This argument only affects the discrete models.

`-m/--maxIter [integer]`

This argument is optional, in which user specifies the maximum number of iterations to scan for waves of admixture events. Default is 10000. This argument only affects the discrete models.

`-w/--waves [integer]`

This argument is optional, in which user specifies the maximum number of waves from one population. Default is 2. This argument only affects the discrete models.

`-c/--ci [double]`

This argument is optional, in which user specifies the confidence level of bootstrapping confidence interval. Default is 0.95.

`-b/--bootstrap [integer]`

This argument is optional, in which user specifies the number of bootstrapping repeats. The default is 1, which does not perform bootstrapping.

`-t/--thread [integer]`

This argument is optional, in which user specifies the number of thread in parallel computation. The default is 1, which does not perform parallel computation. This argument required software compiled with *OpenMP*.

`-M/--mode [DMode/CMode/MixMode]`

This argument is optional, in which user specifies the software running mode. DMode stands for running only consider the discrete model. (Same as MultiWaver 1.0) CMode stands for running only consider the continuous model. (Same as AdmixInfer) MixMode stands for considering all those models together.

-s/--sample

Run in simple mode, default is OFF. This option only affects the discrete models.

-h/--help

Print help message.

5. Questions and suggestions

Questions and suggestions are welcomed, feel free to contact
Yuan Kai
yorkklause@gmail.com