

Manual for *HierarchyMix*

1. Introduction

HierarchyMix is a software that can be used to reconstruct the 4-way admixture history considering the complex admixture structure. It could automatically select the optimal admixture model between the hierarchical admixture model and the sequential admixture model. *HierarchyMix* is an extension of the *MultiWaver* series software.

2. Input Files

The format of input ancestral tracts file is similar to *MultiWaver* series software but with two additional columns. Each line represents an ancestral tract with the genetic distance of start point, the genetic distance of end point, from which ancestry the tract originates, the index of admixed haplotype, the haplotype label, and the chromosome label. The unit of first two columns are both in Morgan. Here takes the first few rows of “Hier.seg” as an example:

| | | | | |
|------------|------------|-------|-------|-------|
| 0 | 0.01113699 | Anc3 | hap0 | 1 |
| 0.01113699 | 0.02255031 | Anc2 | hap0 | 1 |
| 0.02255031 | 0.02763909 | Anc3 | hap0 | 1 |
| 0.02763909 | 0.03211435 | Anc4 | hap0 | 1 |
| 0.03211435 | 0.08753307 | Anc3 | hap0 | 1 |
| 0.08753307 | 0.09612572 | Anc4 | hap0 | 1 |
| | | | | |

3. Output Files

The format of output model file is same as *MultiWaver* series software. Here takes the output file “Hier.txt” as an example, which is inferred based on the input file “Hier.seg”:

```
Best Model:      Hierarchical Admixture Model
Anc1      143(G)      0.208525
Anc2      143(G)      0.791475
Anc3      147(G)      0.700884
Anc4      147(G)      0.299116
Anc1_Anc2  50(G)      0.398402
Anc3_Anc4  50(G)      0.601598
-----
Bootstrapping details
Bootstrapping support ratio:100.0% (100/100)
Anc1      [137, 147](G) [0.20302, 0.214607]
Anc2      [137, 147](G) [0.785393, 0.79698]
Anc3      [144, 149](G) [0.694946, 0.706197]
Anc4      [144, 149](G) [0.293803, 0.305054]
Anc1_Anc2 [48, 51](G)   [0.389099, 0.40447]
Anc3_Anc4 [48, 51](G)   [0.59553, 0.610901]
```

4. Arguments and Options

| Arguments | Type | Description | Note |
|-------------|--------|---|-----------------------------|
| --input | string | Input file name | required |
| --lower | float | Lower bound to discard short tracts | optional, default = 0 |
| --bootstrap | int | Number of bootstrapping | optional, default = 0 |
| --ci | float | Confidence level of bootstrapping confidence interval | optional, default = 0.95 |
| --output | string | Prefix of output file | optional, default="out" |
| -h | / | Print help message | / |

5. Notes

A special hierarchical admixture model needs special attention, within which the admixture proportions of two populations of left-side are same, and the admixture proportions of two populations of right-side are same too. In this case, the admixture time derived from the length distribution of ancestral tracts doesn't have the unique solution and the estimated time of *HierarchyMix* can be unreliable.

6. Questions and Trouble-shooting

Questions and suggestions are welcome, feel free to contact Shi Zhang 19121640@bjtu.edu.cn or Rui Zhang ruizhang2023@gmail.com.