

## About

Mitochondrial mixture database and interpretation tool (MMDIT) is an open-source, interactive software for the probabilistic genotyping of mitochondrial DNA mixtures based on complete mitochondrial genomes (mtGenomes). MMDIT can perform both 'Mixture Deconvolution' and 'Mixture Analysis'. This user guide is intended to help get started with MMDIT with sample data provided within the application. This guide will be updated to include new features as they are integrated in the application.

This tool can be accessed at <https://www.unthsc.edu/mmdit/> or at [https://github.com/SammedMandape/MMDIT\\_UI](https://github.com/SammedMandape/MMDIT_UI).

## Github repository set up

This setup assumes you have R (version > 3.8) and Rstudio (version > 1.3) installed on your computer. Clone or download the repository from github. Open files ui.R and server.R. Make sure all the libraries mentioned in the beginning of ui.R are installed in R. The library “MMDIT” (required to run this tool) can be installed directly from github repository (“ahhgust/MMDIT”). However, the library “devtools” is required to install a package directly from gitbhu.

```
> library(devtools)
> install_github("ahhgust/MMDIT")|
```

## Quick start using sample data

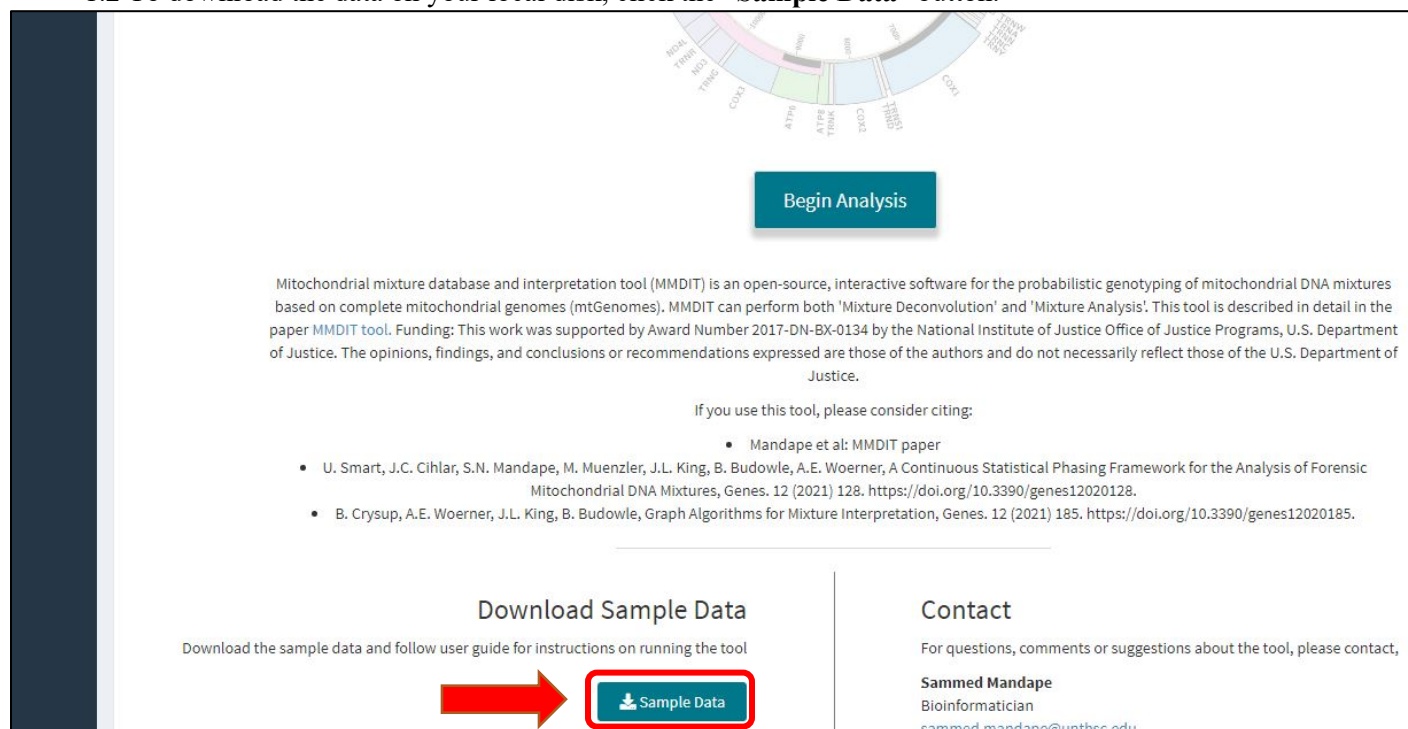
1. Download the sample data
2. Import the sample data
3. Mixture deconvolution
4. Mixture analysis

### 1. Download the sample data

1.1 The sample data (SampleData.zip) is provided on the homepage at <https://www.unthsc.edu/mmdit/>. The data consists of the following:

- a. Two-persons mixture from European descend at a ratio of 1:4 (Mixture\_Profile1\_profile2\_1\_4.txt)
- b. Two single-source profiles comprising the mixture (Profile1.txt and Profile2.txt)
- c. Excel file from converge containing variants and read depth information (Mixture\_quantitative\_file.xlsx)
- d. Sites to include and exclude (Sites\_to\_include\_exclude.txt)

1.2 To download the data on your local disk, click the “Sample Data” button.



The screenshot shows the MMDIT homepage. At the top, there is a circular diagram of a mitochondrial genome with various genes labeled. Below it is a blue button labeled "Begin Analysis". The main text describes MMDIT as an open-source, interactive software for probabilistic genotyping of mitochondrial DNA mixtures. It mentions funding from the National Institute of Justice. Below this, there is a section for citing the tool, listing two papers. At the bottom, there is a "Download Sample Data" section with a red arrow pointing to a blue button labeled "Sample Data". To the right of this is a "Contact" section with the name "Sammed Mandape", title "Bioinformatician", and email "sammed.mandape@unthsc.edu".

Begin Analysis

Mitochondrial mixture database and interpretation tool (MMDIT) is an open-source, interactive software for the probabilistic genotyping of mitochondrial DNA mixtures based on complete mitochondrial genomes (mtGenomes). MMDIT can perform both 'Mixture Deconvolution' and 'Mixture Analysis'. This tool is described in detail in the paper [MMDIT tool](#). Funding: This work was supported by Award Number 2017-DN-BX-0134 by the National Institute of Justice Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect those of the U.S. Department of Justice.

If you use this tool, please consider citing:

- Mandape et al: MMDIT paper
- U. Smart, J.C. Cihlar, S.N. Mandape, M. Muenzler, J.L. King, B. Budowle, A.E. Woerner, A Continuous Statistical Phasing Framework for the Analysis of Forensic Mitochondrial DNA Mixtures, *Genes*. 12 (2021) 128. <https://doi.org/10.3390/genes12020128>.
- B. Crysap, A.E. Woerner, J.L. King, B. Budowle, Graph Algorithms for Mixture Interpretation, *Genes*. 12 (2021) 185. <https://doi.org/10.3390/genes12020185>.

Download Sample Data

Download the sample data and follow user guide for instructions on running the tool

Sample Data

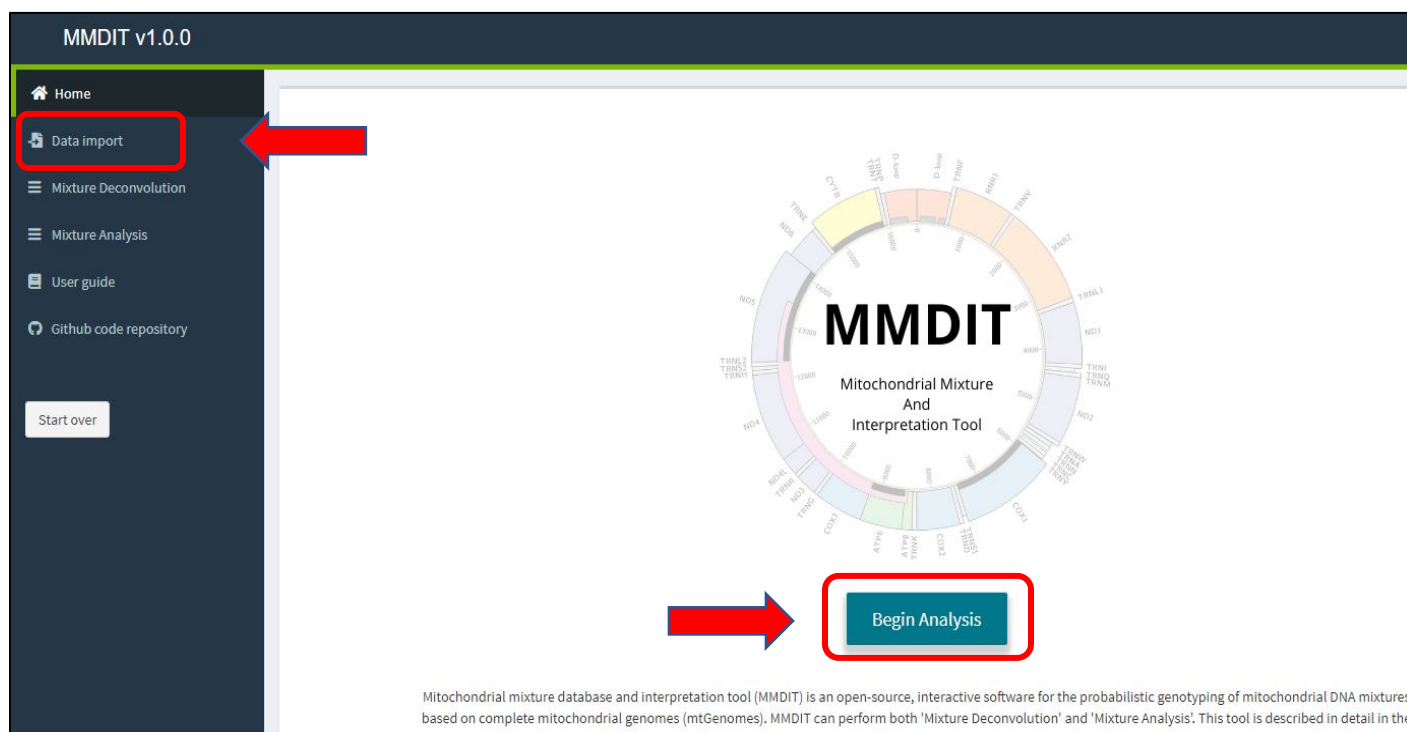
Contact

For questions, comments or suggestions about the tool, please contact,

**Sammed Mandape**  
Bioinformatician  
[sammed.mandape@unthsc.edu](mailto:sammed.mandape@unthsc.edu)

## 2. Import the sample data

2.1 There are two ways to import data from homepage. One is by clicking the “Begin Analysis” button and the other is by clicking the “Data Import” tab in the left panel of the homepage.



The screenshot shows the MMDIT v1.0.0 interface. On the left is a dark sidebar with a menu. The "Data import" option is highlighted with a red box and a red arrow points to it. The main area features a circular diagram of a mitochondrial genome with the text "MMDIT Mitochondrial Mixture And Interpretation Tool" in the center. Below this is a blue button labeled "Begin Analysis", which is also highlighted with a red box and a red arrow points to it. At the bottom, there is a paragraph of text describing MMDIT as an open-source, interactive software for probabilistic genotyping of mitochondrial DNA mixtures based on complete mitochondrial genomes (mtGenomes).

MMDIT v1.0.0

Home

Data import

Mixture Deconvolution

Mixture Analysis

User guide

Github code repository

Start over

MMDIT

Mitochondrial Mixture And Interpretation Tool

Begin Analysis

Mitochondrial mixture database and interpretation tool (MMDIT) is an open-source, interactive software for the probabilistic genotyping of mitochondrial DNA mixtures based on complete mitochondrial genomes (mtGenomes). MMDIT can perform both 'Mixture Deconvolution' and 'Mixture Analysis'. This tool is described in detail in the

2.2 The above step brings you to the “Data Import” page. Use the first “**Browse**” (arrow 1 in figure) button to upload single-source profiles from local disk (Profile1.txt and Profile2.txt).

2.3 Use the second “**Browse**” (arrow 2 in figure) button to upload mixture data from local disk (Mixture\_Profile1\_profile2\_1\_4.txt).

2.4 Click “**Analyze data**” (arrow 3 in figure) button to preprocess the uploaded data.

2.5 Click “**Next**” (arrow 4 in figure) button to go to the “**Indel analysis**” tab.

**STEP 1 : File Import**

Upload one or more single source data

Browse... 2 files Upload complete

Upload mixture data \*

Browse... Mixture\_Profile1\_Profile2\_1\_4.txt Upload complete

Clear selection

**STEP 2 : Preprocessing**

Analyze data

**STEP 3 : Preprocessing Results**

SNV analysis Indel analysis Select Inclusion/Exclusion list

Next

	FileID	Start	Stop	Allele	Type	Source
1	Profile1.txt	72	73	A	Substitution	Single
2	Profile1.txt	145	146	C	Substitution	Single
3	Profile1.txt	216	217	T	Substitution	Single
4	Profile1.txt	262	263	G	Substitution	Single
5	Profile1.txt	507	508	A	Substitution	Single
6	Profile1.txt	749	750	G	Substitution	Single
7	Profile1.txt	1437	1438	G	Substitution	Single
8	Profile1.txt	1810	1811	A	Substitution	Single
9	Profile1.txt	2705	2706	G	Substitution	Single
10	Profile1.txt	4091	4092	A	Substitution	Single

2.6 Click “**Save**” button to save the table for use in downstream analysis.

**STEP 3 : Preprocessing Results**

SNV analysis Indel analysis Select Inclusion/Exclusion list

Save

	FileID	Start	Stop	Allele	Type	Source
1	Profile1.txt	315	315	C	Insertion	Single
2	Profile2.txt	309	309	C	Insertion	Single
3	Profile2.txt	315	315	C	Insertion	Single
4	Profile2.txt	16193	16193	C	Insertion	Single
5	Mixture_Profile1_Profile2_1_4.txt	309	309	C	Insertion	Mixture
6	Mixture_Profile1_Profile2_1_4.txt	309	309	-	Insertion	Mixture
7	Mixture_Profile1_Profile2_1_4.txt	315	315	C	Insertion	Mixture

2.7 On “Select Inclusion/Exclusion list” tab click “**Load MMDIT database**” button followed by selecting the option “**Choose from Precision ID Kit**” as shown in the figure.

**STEP 3 : Preprocessing Results**

SNV analysis   Indel analysis   **Select Inclusion/Exclusion list**

**Load MMDIT database**   **Finalize list**

**Choose Inclusion/Exclusion list**

☐ Choose one option to input inclusion / exclusion list

☒ **Choose from Precision ID Kit**

☐ Upload bed file of regions to include / exclude

**Select populations**

African, American, Asian, European, Oceanian, XX

2.8 Select amplicons 1 to 7 (as mentioned in the file Sites\_to\_include\_exclude.txt) from page 1 and go to page 16

Precision kitID amplicon table (Please select amplicons to include)

Show **10** entries   Search:

	start	stop
1	15	119
2	118	248
3	247	329
4	298	411
5	384	480
6	459	543
7	518	610
8	596	709
9	695	808
10	793	903

Showing 1 to 10 of 162 entries   Previous   **1**   2   3   4   5   ...   17   Next

**Manually select amplicons (default)**   **Continue**   **Dismiss**

Choose Inclusion/Exclusion list   Select populations


2.9 Select amplicons 156 to 160 and click “Continue” button.

Precision kitID amplicon table (Please select amplicons to include)

Show  entries Search:

	start	stop
151	15451	15585
152	15581	15693
153	15682	15769
154	15758	15859
155	15848	15964
156	15953	16069
157	16055	16131
158	16109	16225
159	16221	16341
160	16338	16458

Showing 151 to 160 of 162 entries Previous 1 ... 13 14 15 **16** 17 Next


Manually select amplicons (default)  Continue Dismiss

2.10 Input regions from mitochondrial genome to exclude 16179-16193;302-315. Click “Done” button.

Input semicolon separated genomic coordinates to exclude in bed format

Selected amplicons (row numbers):

[1] 1 2 3 4 5 6 7 156 157 158 159 160

 Done

2.11 In “Select populations” uncheck “Oceanian” and “XX” and click “**Finalize list**” button to save all the parameters selected. Click “**Mixture Deconvolution**” button.

**STEP 3 : Preprocessing Results**

SNV analysis   Indel analysis   Select Inclusion/Exclusion list

Load MMDIT database   **Finalize list**

Choose Inclusion/Exclusion list

☐ Choose one option to input inclusion / exclusion list

☒ Choose from Precision ID Kit

☐ Upload bed file of regions to include / exclude

Select populations

African, American, Asian, European

Select All   Deselect All

Asian   ✓

European   ✓

Oceanian

XX

**STEP 4 : Analyze**

Choose between one of the following methods to analyze data.

**Mixture Deconvolution**   Mixture Analysis

### 3. Mixture deconvolution

3.1 On “Mixture Deconvolution” page click “**Browse**” button to upload quantitative data from local disk (Mixture\_quantitative\_file.xlsx). Click “**Analyze**” button to start the deconvolution analysis.

**Mixture Deconvolution**

Upload excel file with quantitative data

Browse...   No file selected

Graph edit distance (GED)

4

Number of Markov Chain Monte Carlo (MCMC)

3000

Read count normalization value

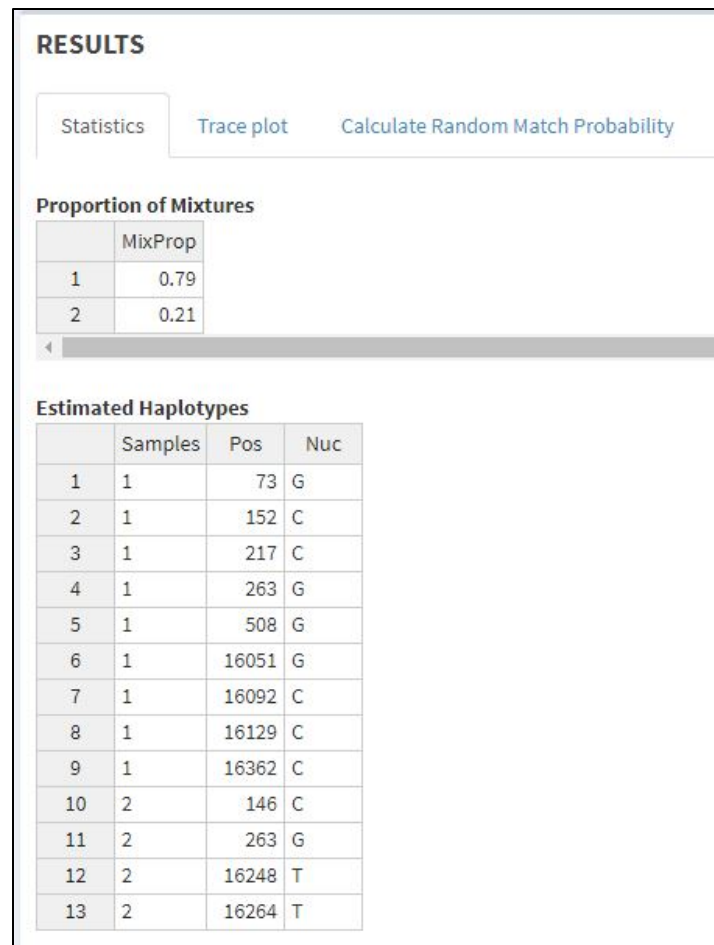
100

Miscopying rate

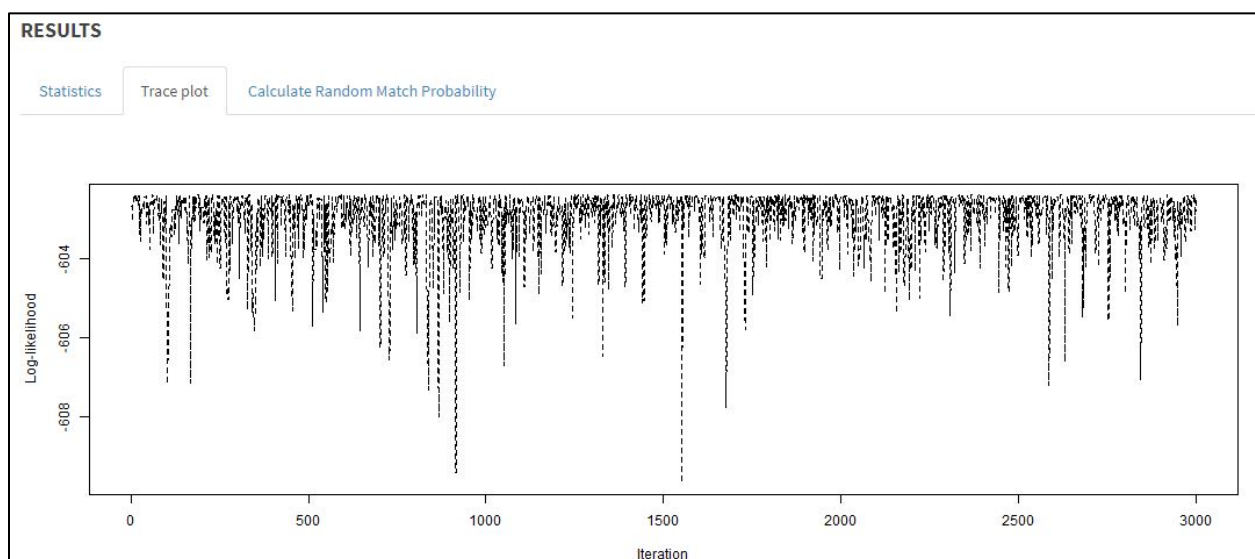
0.01

**Analyze**

3.2 Results are displayed in the “Results” pane on “Mixture Deconvolution” page. “Statistics” tab shows the proportion of two-persons in the mixture and estimated haplotypes of those two-persons.



3.3 “Trace plot” tab displays how well converged is the mixture deconvolution analysis.



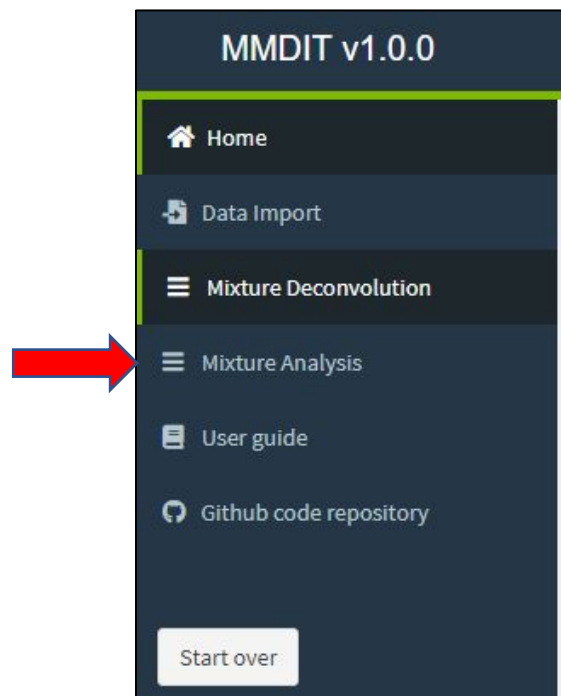


3.4 “Calculate Random Match Probability” tab displays the random match probability of estimated haplotypes for every population considered in the analysis.

RESULTS						
Statistics   Trace plot   Calculate Random Match Probability						
	haploypeid	pop	NMatches	RmpFrequentist	RmpTheta	Theta
1	value	AF	0	0.00100	0.00220	0.00190
2	value	AM	1	0.00170	0.00250	0.00190
3	value	AS	0	0.00060	0.00210	0.00190
4	value	EU	10	0.00160	0.00280	0.00190
5	value	AF	0	0.00100	0.00220	0.00190
6	value	AM	0	0.00110	0.00220	0.00190
7	value	AS	0	0.00060	0.00210	0.00190
8	value	EU	0	0.00030	0.00200	0.00190

## 4. Mixture analysis

4.1 Click “Mixture Analysis” button in the panel on the left side of MMDIT page.





4.2 On “Mixture Analysis” page, in “Enter the number of contributors in mixture” select “**2-persons mixture**” and click “**Generate Mixture Statistics**” button.

### Mixture Analysis

Enter the number of contributors in mixture

2-persons mixture



Generate Mixture Statistics

4.3 The “Results” pane on “Mixture Analysis” displays the Random Man Not Excluded (RMNE) statistics as well as Likelihood Ratios (LR) statistics for the input data.

RESULTS

Random Man Not Excluded (RMNE) Statistics

	pop	Count	CountExcludingKnown	Tot	LogRMNEUB
1	Known haplotype 1	1	0.00	1.00	0.00
2	Known haplotype 2	1	0.00	1.00	0.00
3	AF	23	23.00	3701.00	-2.03
4	AM	63	62.00	3285.00	-1.61
5	AS	30	30.00	6397.00	-2.17
6	EU	981	971.00	11822.00	-1.05

Likelihood Ratios (LR) Statistics

	pop	NCombinationsThatExplain	NKnown	NExplain	Divisor	WhichKnown	Feasibility	LogLikelihood_GBC	NinMix
1	AF	0.00	0.00	0.00	6846850.00		Feasible	-6.27	2.00
2	AF	0.00	1.00	0.00	3701.00	2	Feasible	-3.00	2.00
3	AF	1.00	1.00	0.00	3701.00	1	Feasible	-3.00	2.00
4	AF	1.00	2.00	1.00	1.00	1,2	Feasible	0.00	2.00
5	AM	0.00	0.00	0.00	5393970.00		Feasible	-6.16	2.00
6	AM	0.00	1.00	0.00	3285.00	2	Feasible	-2.95	2.00
7	AM	1.00	1.00	1.00	3285.00	1	Feasible	-2.77	2.00
8	AM	1.00	2.00	1.00	1.00	1,2	Feasible	0.00	2.00
9	AS	0.00	0.00	0.00	20457606.00		Feasible	-6.74	2.00
10	AS	0.00	1.00	0.00	6397.00	2	Feasible	-3.24	2.00
11	AS	1.00	1.00	0.00	6397.00	1	Feasible	-3.24	2.00
12	AS	1.00	2.00	1.00	1.00	1,2	Feasible	0.00	2.00
13	EU	0.00	0.00	0.00	69873931.00		Feasible	-7.28	2.00
14	EU	0.00	1.00	0.00	11822.00	2	Feasible	-3.51	2.00
15	EU	1.00	1.00	10.00	11822.00	1	Feasible	-2.81	2.00
16	EU	1.00	2.00	1.00	1.00	1,2	Feasible	0.00	2.00

