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.

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

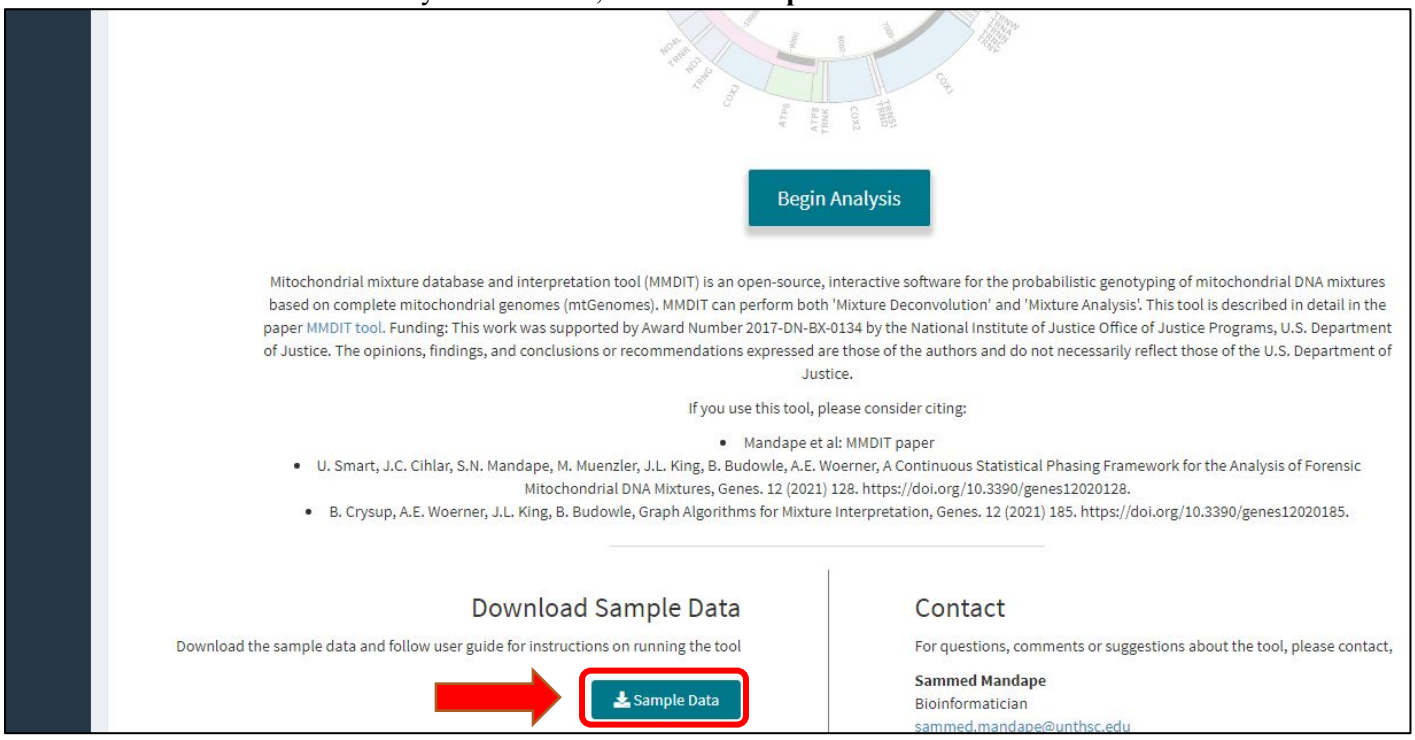
1. Download the sample data
2. Import the sample data
3. Mixture deconvolution
4. Mixture analysis

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:

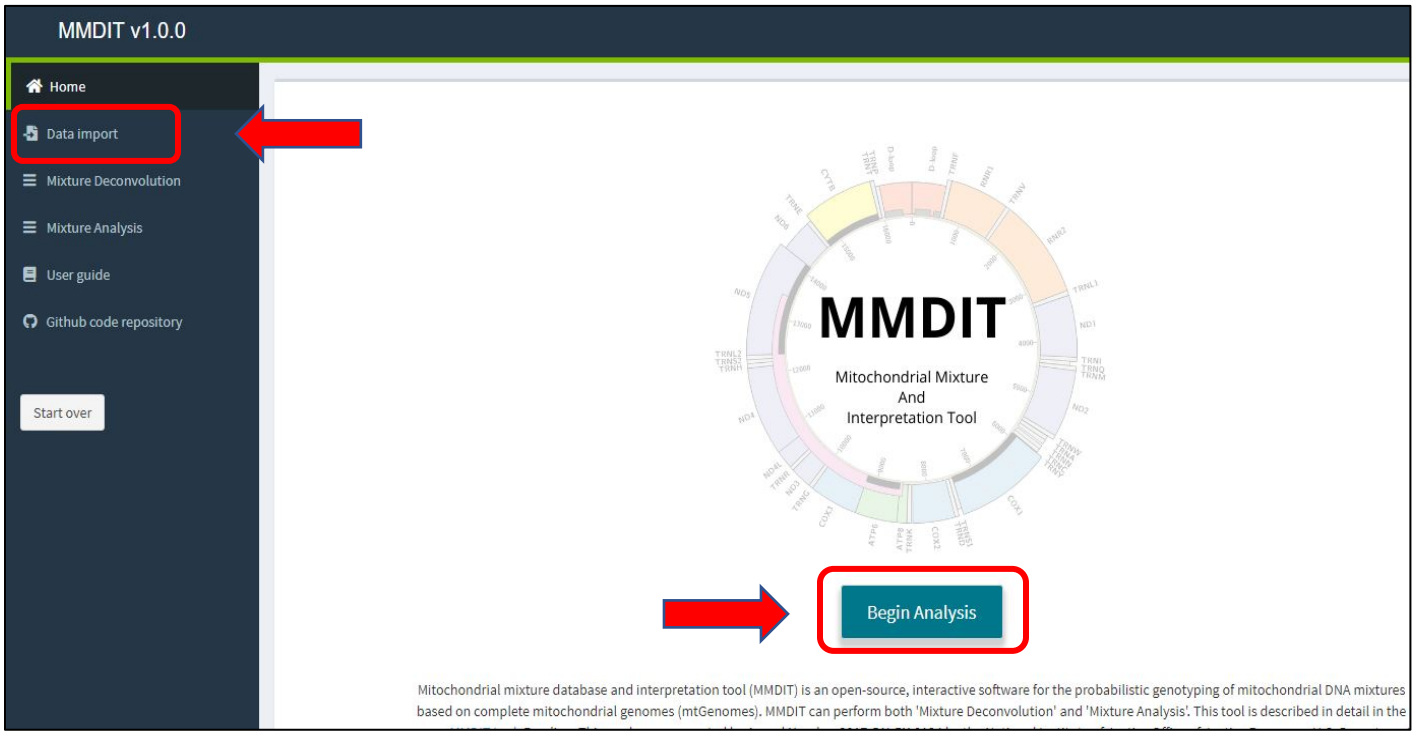
a. Two-persons mixture from European descend at a ratio of 1:4 (Mixture_Profile1_profile2_1_4.txt)
b. Two single-source profiles comprising the mixture (Profile1.txt and Profile2.txt)
c. Excel file from converge containing variants and read depth information (Mixture_quantitative_file.xlsx)
d. Sites to include and exclude (Sites_to_inlude_exclude.txt)

1.2 To download the data on your local disk, click the "**Sample Data**" button.

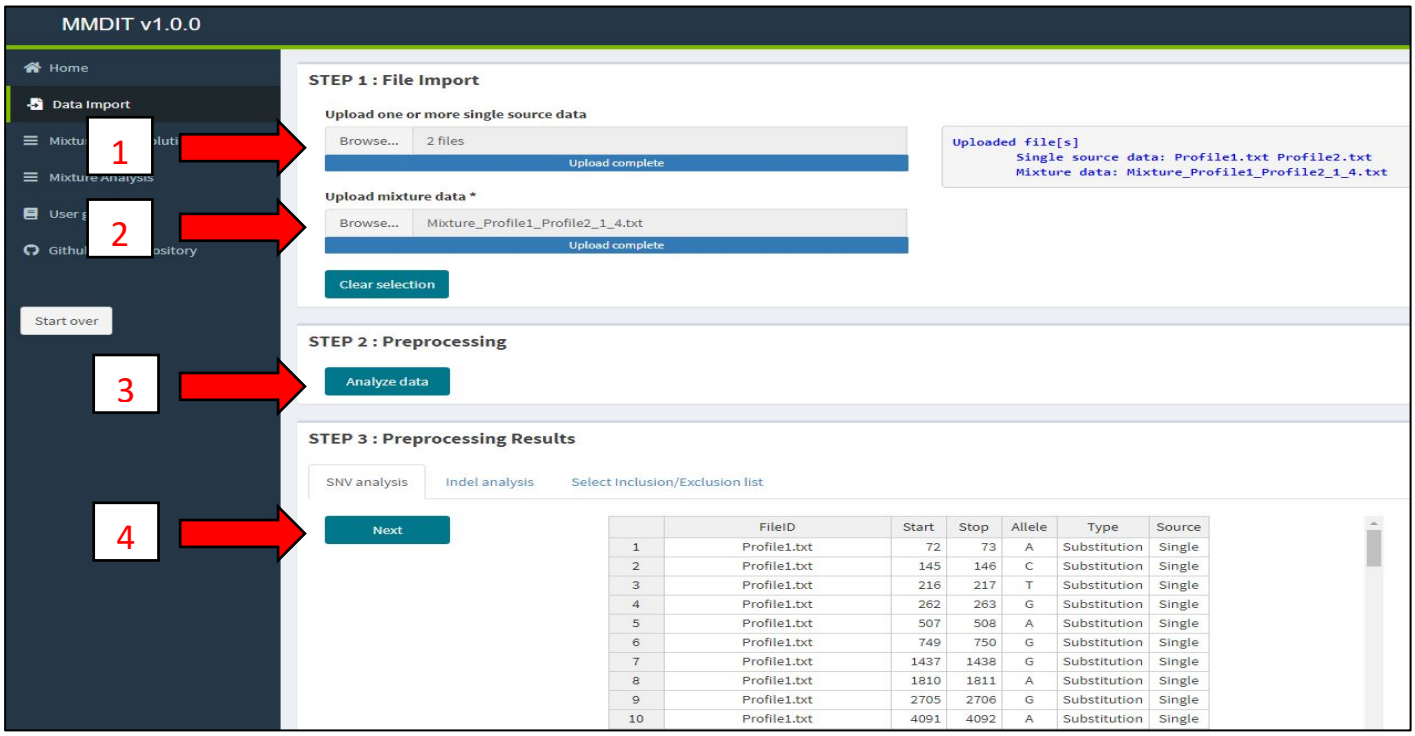


2. Import the sample data

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.

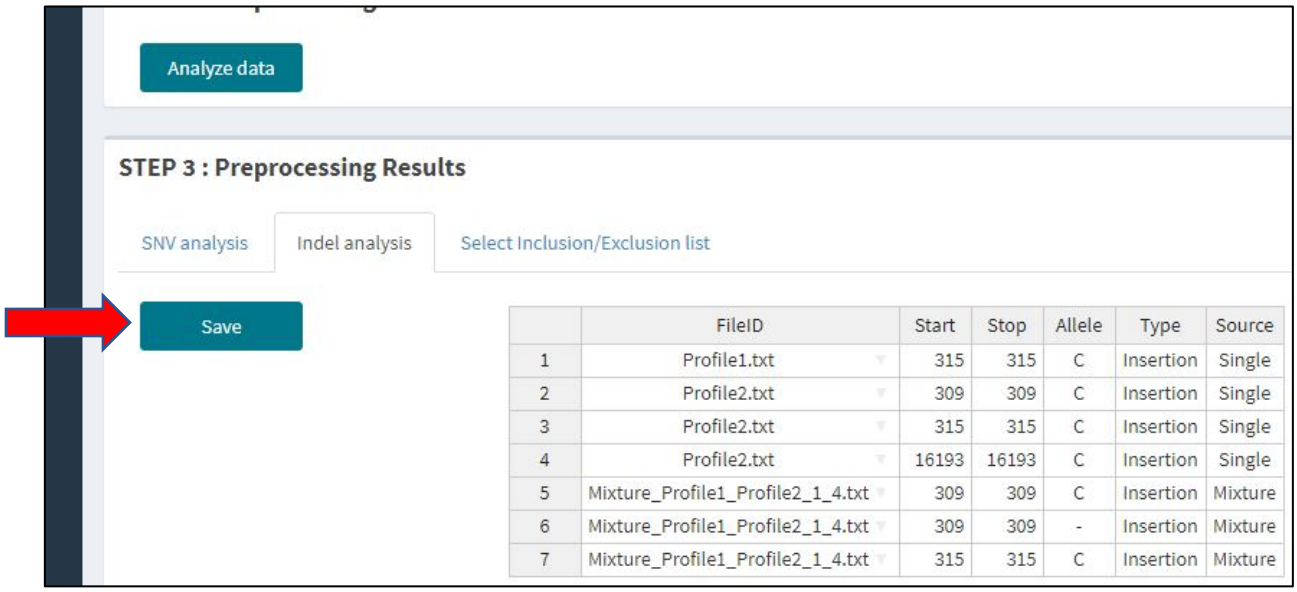


2.2 The above step brings you to the “Data Import” page. Use the first “**Browse**” (arrow 1 in figure) button to upload single-source profiles from local disk (Profile1.txt and Profile2.txt).
2.3 Use the second “**Browse**” (arrow 2 in figure) button to upload mixture data from local disk (Mixture_Profile1_profile2_1_4.txt).
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.



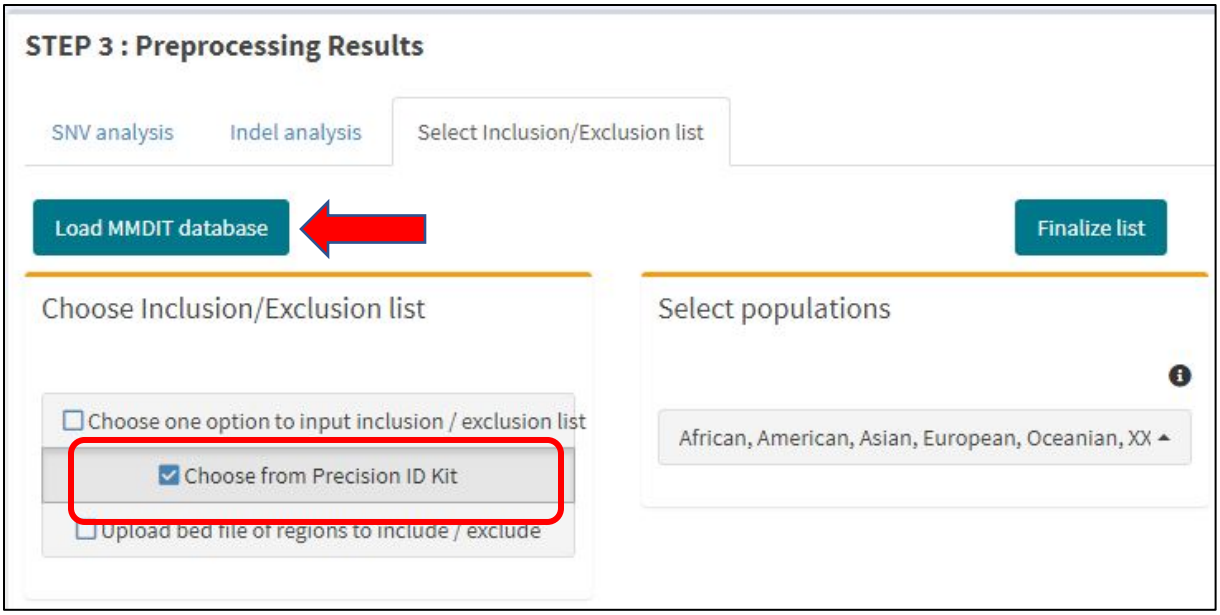
	FileID	Start	Stop	Allele	Type	Source
1	Profile1.txt	72	73	A	Substitution	Single
2	Profile1.txt	145	146	C	Substitution	Single
3	Profile1.txt	216	217	T	Substitution	Single
4	Profile1.txt	262	263	G	Substitution	Single
5	Profile1.txt	507	508	A	Substitution	Single
6	Profile1.txt	749	750	G	Substitution	Single
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.

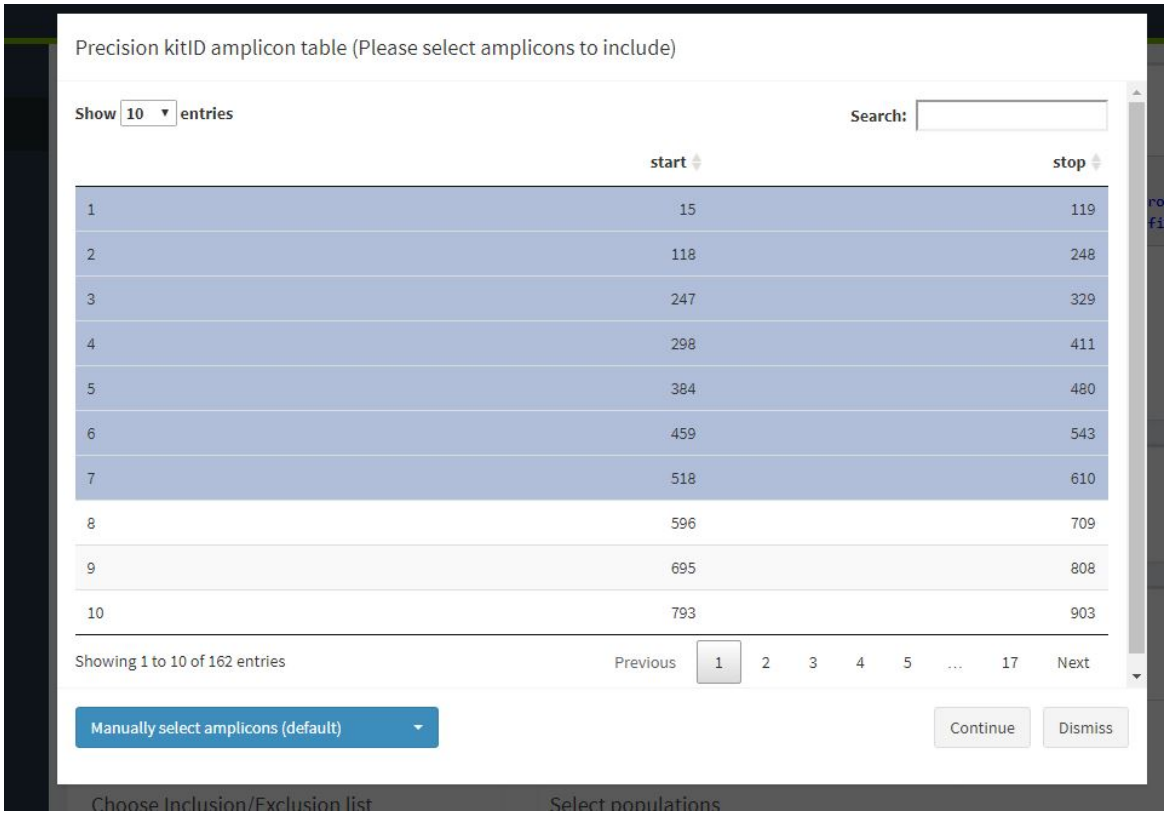


	FileID	Start	Stop	Allele	Type	Source
1	Profile1.txt	315	315	C	Insertion	Single
2	Profile2.txt	309	309	C	Insertion	Single
3	Profile2.txt	315	315	C	Insertion	Single
4	Profile2.txt	16193	16193	C	Insertion	Single
5	Mixture_Profile1_Profile2_1_4.txt	309	309	C	Insertion	Mixture
6	Mixture_Profile1_Profile2_1_4.txt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

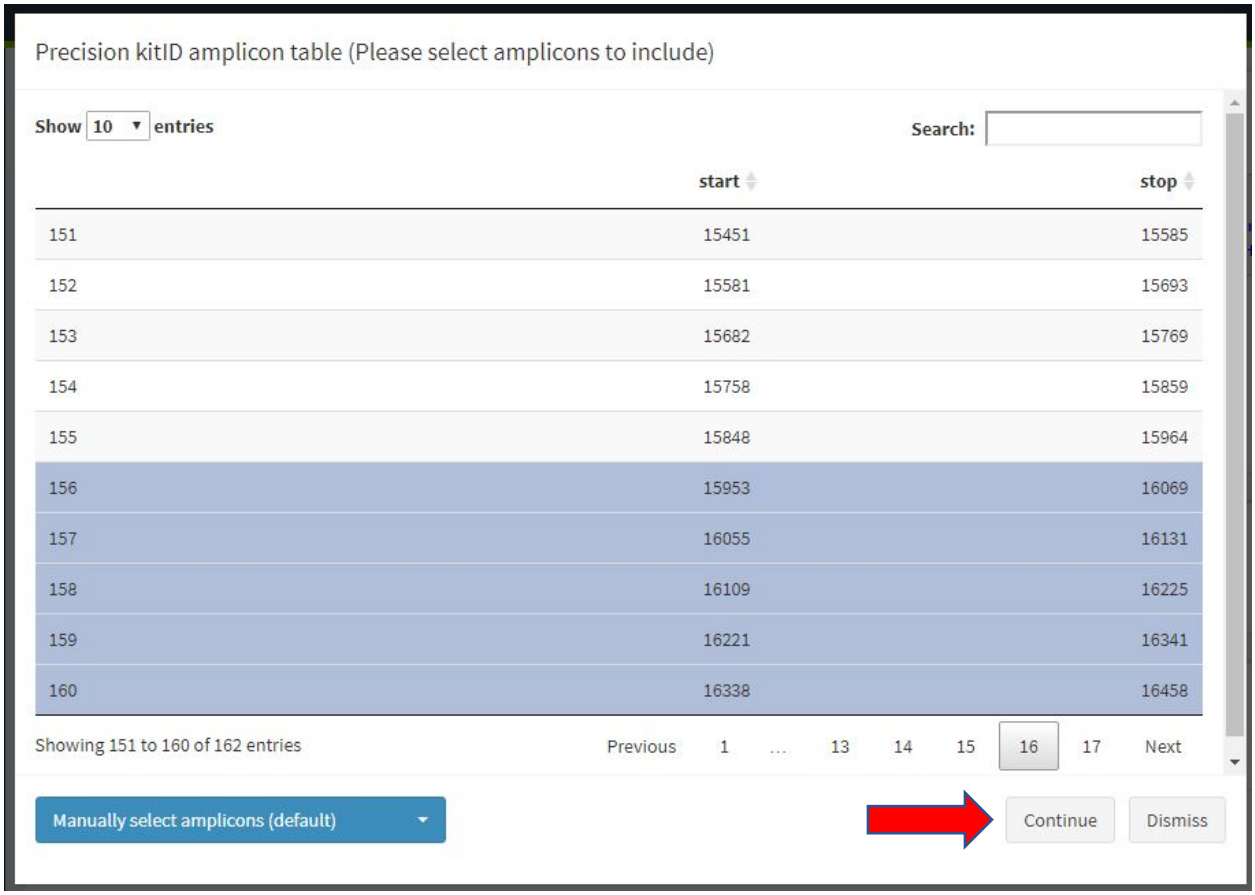
2.7 On “Select Inclusion/Exclusion list” tab click “**Load MMDIT database**” button followed by selecting the option “**Choose from Precision ID Kit**” as shown in the figure.



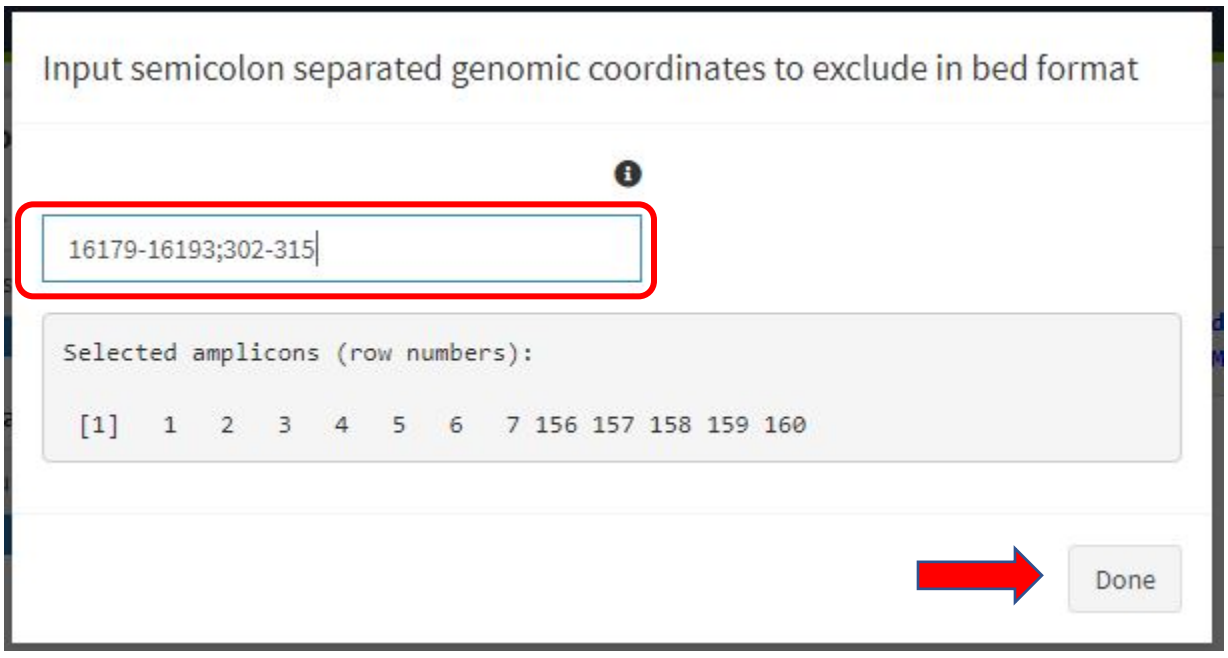
2.8 Select amplicons 1 to 7 (as mentioned in the file Sites_to_inlude_exclude.txt) from page 1 and go to page 16



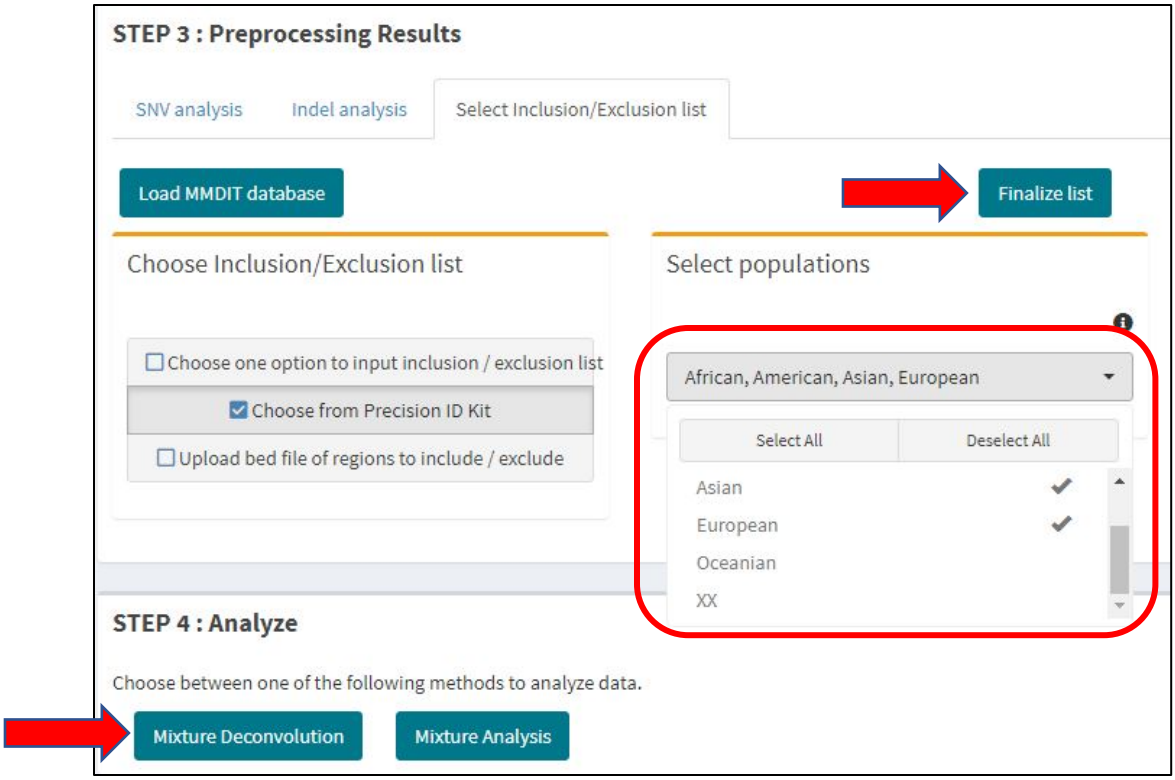
2.9 Select amplicons 156 to 160 and click "**Continue**" button.



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

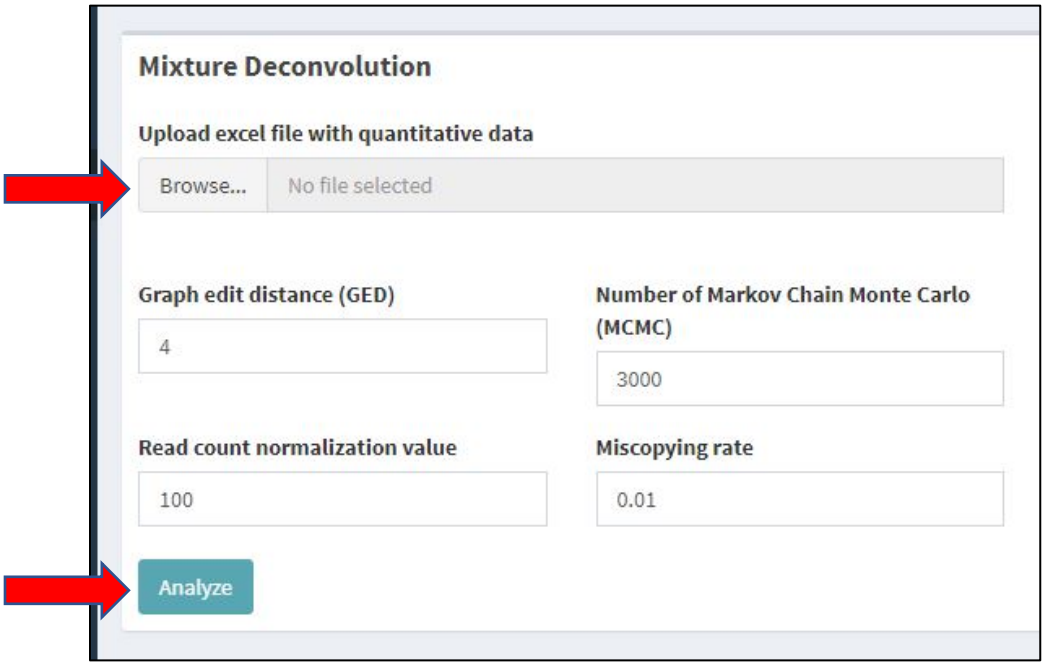


2.11 In “Select populations” uncheck “Oceanian” and “XX” and click “**Finalize list**” button to save all the parameters selected. Click “**Mixture Deconvolution**” button.

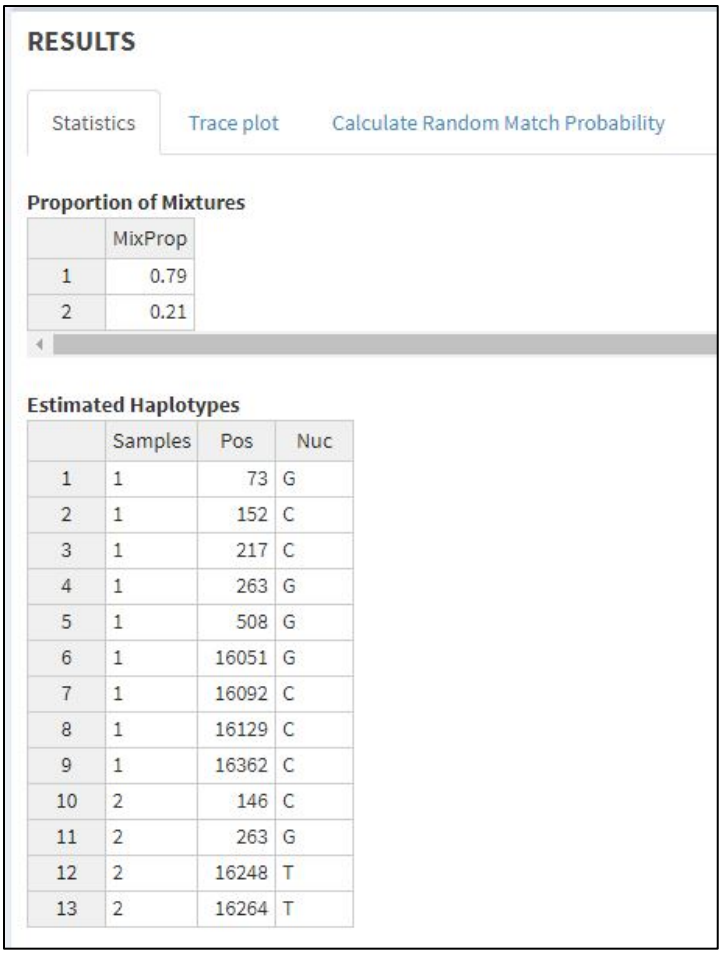


3. Mixture deconvolution

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.



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.



RESULTS

Statistics | Trace plot | Calculate Random Match Probability

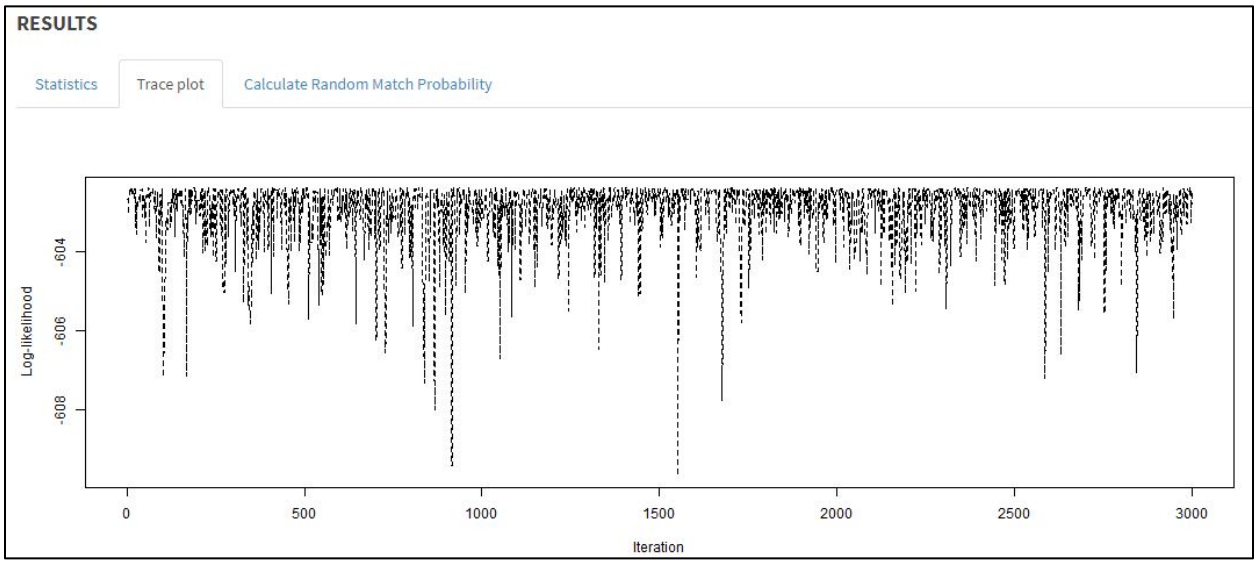
Proportion of Mixtures

	MixProp
1	0.79
2	0.21

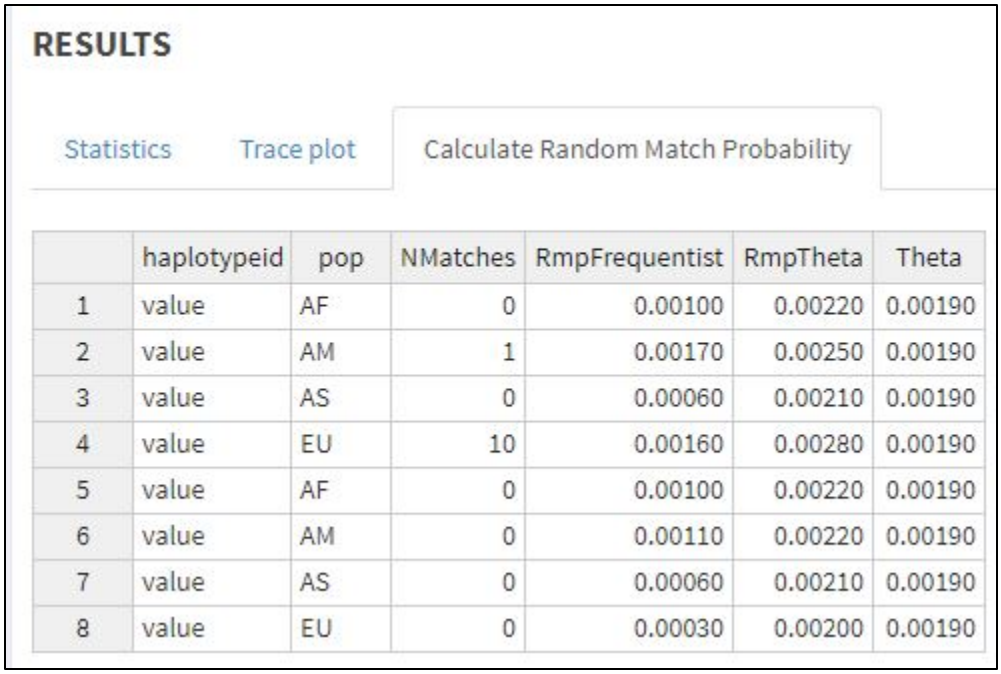
Estimated Haplotypes

	Samples	Pos	Nuc
1	1	73	G
2	1	152	C
3	1	217	C
4	1	263	G
5	1	508	G
6	1	16051	G
7	1	16092	C
8	1	16129	C
9	1	16362	C
10	2	146	C
11	2	263	G
12	2	16248	T
13	2	16264	T

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



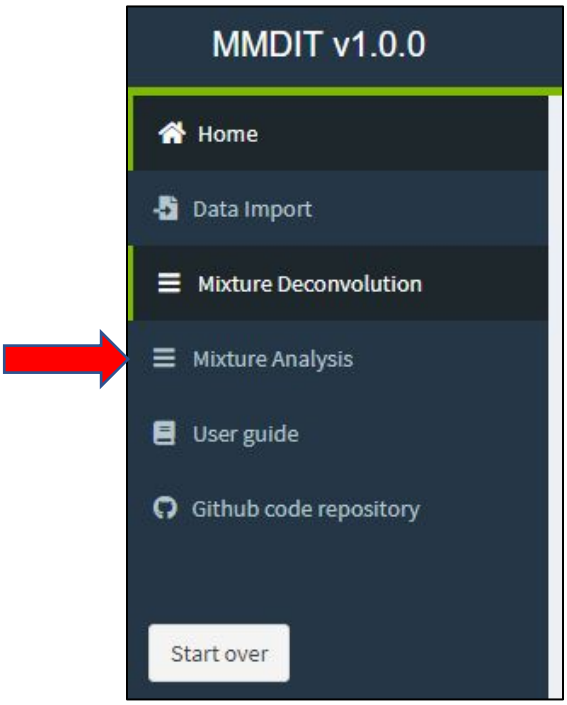
3.4 “Calculate Random Match Probability” tab displays the random match probability of estimated haplotypes for every population considered in the analysis.



	haplotypeid	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.

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