$relsearch_0.1.0_manual$

Sho Manabe (manabesh@hirakata.kmu.ac.jp)

2022-11-11

Contents

Information on relsearch	3
Getting started	4
Autosomal STR	5
Load query database	. 5
Load reference database	. 7
Load allele frequencies	. 9
Set conditions of calculation	. 10
Set mutation rates	. 12
Edit mutation rates	. 12
Add a locus	. 14
Delete a locus	. 16
Set IBD probabilities	. 17
Edit IBD probabilities	. 17
Add a relationship	. 20
Delete a relationship	. 22
Perform screening	. 23
Check the result of screening	. 24
Change the displayed information	. 25
Show the result of a query-reference pair in detail	. 27
Export data	. 29
Y-chromosomal STR	31
Load query database	. 31
Load reference database	. 33
Perform screening	. 35
Check the result of screening	. 36

Change the displayed information	37
Show the result of a query-reference pair in detail	39
Export data	41
Mitochondrial DNA	42
Load query database	42
Load reference database	43
Perform screening	44
Check the result of screening	45
Change the displayed information	46
Show the result of a query-reference pair in detail	48
Export data	50

Information on relsearch

relsearch is open-source software for screening relatives between query and reference databases. The software is graphical-user-interface written in R language (>= 4.2.0) with GNU General Public License v3.0. It can be applied to autosomal short tandem repeat (STR) markers, Y-STR markers, and mitochondrial DNA sequences commonly used in forensic genetics. For autosomal STR markers, likelihood ratios (LRs) of each query-reference pair are calculated considering mutation and drop-out. For Y-STR markers and mitochondrial DNA sequences, the software investigates the number of inconsistencies between query and reference haplotypes.

Getting started

- 1. Ensure that R (>= 4.2.0) is installed. It is available from the R Development Core Team website (http://www.R-project.org).
- 2. Begin an R session.
- 3. Execute the following command in R to install required packages.

```
install.packages(c("Rcpp", "readr", "tcltk2"))
```

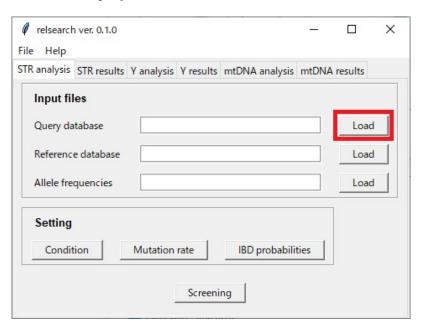
- 4. Go to https://github.com/manabe0322/relsearch/releases.
- 5. Download "relsearch_0.1.0.zip".
- 6. Execute the following commands in R to start GUI.

library(relsearch)
relsearch()

Autosomal STR

Load query database

1. Press the 'Load' button for query database.



2. Select a .csv file for query database. Format of query database is shown in Fig. 1.

Sample Name	D3S1358	D3S1358	vWA	vWA	D16S539	D16S539	CSF1P0	CSF1PO	TPOX	TPOX
Q11	17	17	17		11		12			
Q12	17		18	18			10	11	8	
Q13	15	17	17	18	10	10	10	12	8	11
Q14	15	18	15	18	9	12	11	12	8	8
Q15	15	15	18		9	9	10		9	11
Q16	16	16	17				10	13	8	8
Q17	14	15	17	18	12				8	8
Q18	16	18	17	18	11		9	12	8	
Q19	15	15	16	18	11	11	12	12	8	11
Q20	15	16	18	19	9	10	12	12	8	11

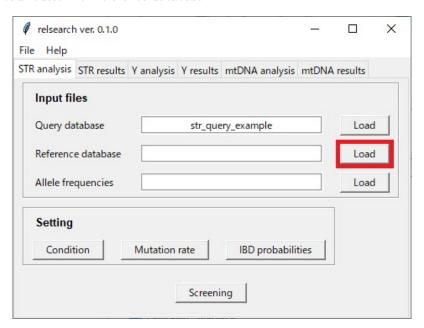
Figure 1: Format of query database for autosomal STR

Note: format of query database

- This file must include information regarding 'Sample Name' and each marker.
- There are two columns in each marker.
- The marker with two empty cells (e.g., D16S539 of sample 'Q12' in Fig. 1) is ignored when calculating the LR.
- The marker with one empty cell (e.g., D3S1358 of sample 'Q12' in Fig. 1) can be regarded as allelic drop-out depending on the setting 'Drop-out of query genotypes' (see section 3.4).
- $\bullet\,$ An example file named "str_query_example.csv" is located at extdata > examples.

Load reference database

1. Press the 'Load' button for reference database.



2. Select a .csv file for reference database. Format of reference database is shown in Fig. 2.

Sample Name	Relationship	D3S1358	D3S1358	vWA	vWA	D16S539	D16S539	CSF1PO	CSF1PO	TPOX	TPOX
R11	parent-child	17	17	14	17	10	10	12	12	8	8
R12	parent-child	17	18	17	18	12	13	10	11	11	11
R13	sibling	15	16	17	18	9	10	11	12	8	11
R14		15	15	15	19	11	12	11	12	8	8
R15	sibling	15	16	15	18	9	9	10	12	9	11
R16	sibling	16	17	17	17	10	12	13	13	8	11
R17	sibling	16	17	14	18	9	9	12	15	8	8
R18	sibling	15	16	17	18	9	12	11	12	8	9
R19	2nd-degree	15	15	14	16	10	11	11	12	8	9
R20		15	16	17	18	9	12	12	12	8	12

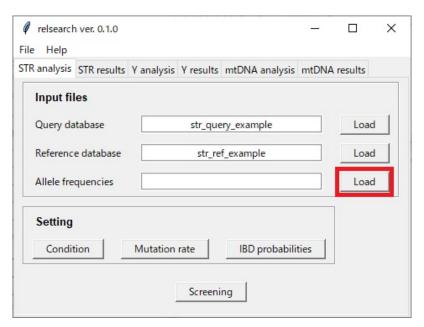
Figure 2: Format of reference database for autosomal STR

Note: format of reference database

- This file must include information regarding 'Sample Name', 'Relationship', and each marker.
- There are two columns in each marker.
- Names of the column 'Relationship' relate to those of relationships in the IBD probabilities (see section 3.6).
- If the relationship is not identified, leave the cell blank (e.g., samples 'R14' and 'R20'). When calculating LRs, all relationships for which the IBD probabilities is defined are considered.
- The marker with two empty cells is ignored when calculating the LR.
- The marker with one empty cell or with two same alleles (e.g., D3S1358 of sample 'R11' in Fig. 2) is regarded as the homozygotes.
- An example file named "str_ref_example.csv" is located at extdata > examples.

Load allele frequencies

1. Press the 'Load' button for allele frequencies.



2. Select a .csv file for allele frequencies. Format of allele frequencies is shown in Fig. 3.

Allele	D3S1358	vWA	D16S539	CSF1P0	TPOX	D8S1179
10			0.200266	0.215471	0.033566	0.128775
10.1						
10.2						
10.3						
11			0.186981	0.206175	0.356597	0.106206
11.1						
11.2						
11.3						
12	0.002324		0.172036	0.420983	0.038219	0.123133
12.2						
13	0.001328	0.000664	0.069744	0.069389	0.001329	0.225357
13.2						
14	0.02656	0.194887	0.008967	0.017596	0.000997	0.207766
14.1						

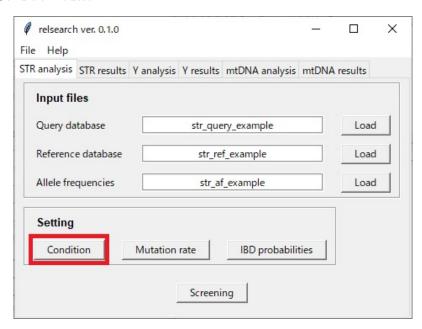
Figure 3: Format of allele frequencies for autosomal STR

Note: format of allele frequencies

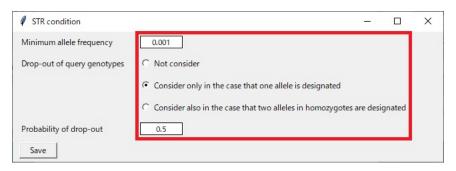
- This file must include information regarding 'Allele' and each marker.
- An example file named "str_af_example.csv" is located at extdata > examples.

Set conditions of calculation

1. Press the 'Condition' button.

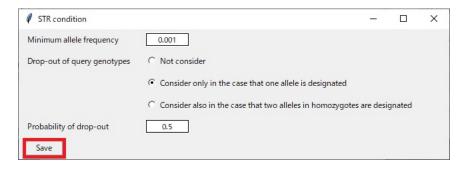


2. Set conditions of 'Minimum allele frequency', 'Drop-out of query genotypes', and 'Probability of drop-out'.



Note: Drop-out of query genotypes

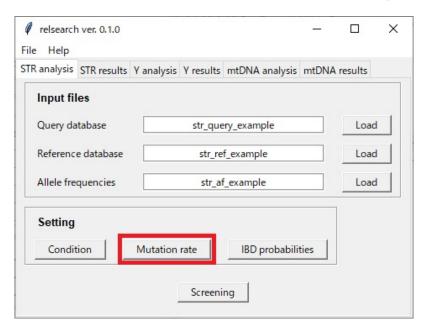
- Not consider: When calculating the LR, the marker with one empty cell in a query or reference database is regarded as the homozygote. The marker with two empty cells is ignored (i.e., LR = 1).
- Consider only in the case that one allele is designated: When calculating the LR, the marker with one empty cell in a query or reference database is regarded as both the homozygote (without drop-out) and the heterozygote (with drop-out). The marker with two empty cells is ignored (i.e., LR = 1).
- Consider also in the case that two alleles in homozygotes are designated: When calculating the LR, the marker with two same alleles in a query or reference database is also regarded as both the homozygote (without drop-out) and the heterozygote (with drop-out).
- 3. Press the 'Save' button.



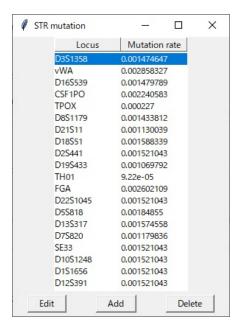
Set mutation rates

Edit mutation rates

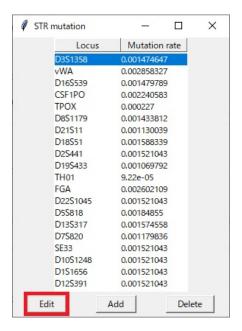
1. Press the 'Mutation rate' button. Then the window 'STR mutation' will be open.



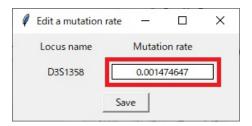
2. Select a locus to change the mutation rate.



3. Press the 'Edit' button. Then the window 'Edit a mutation rate' will be open.



4. Enter an arbitrary mutation rate.

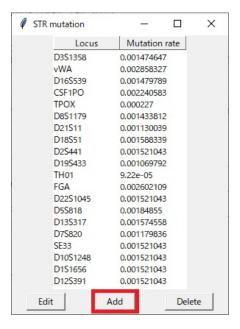


5. Press the 'Save' button.

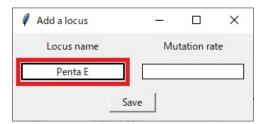


Add a locus

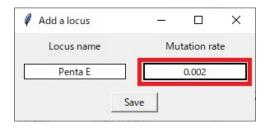
1. Press the 'Add' button in the window 'STR mutation'. Then the window 'Add a locus' will be open.



2. Enter an arbitrary locus name.



3. Enter an arbitrary mutation rate.

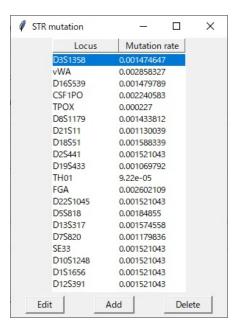


4. Press the 'Save' button.

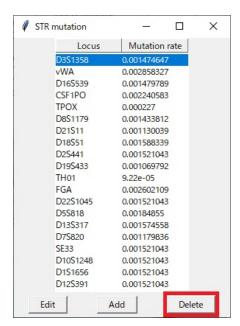


Delete a locus

1. Select a locus in the window 'STR mutation'.



2. Press the 'Delete' button. Then the selected locus will be deleted.



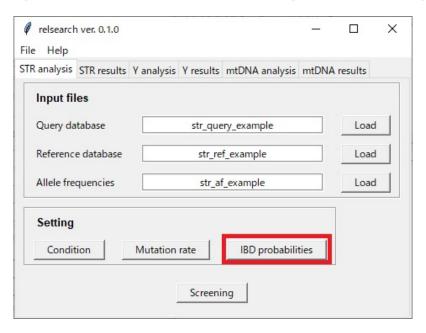
Set IBD probabilities

Note:

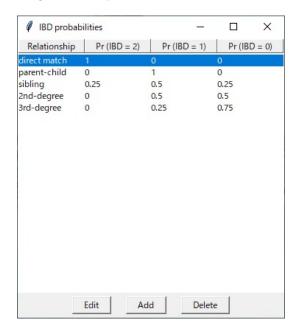
• When the relationship of each reference profile is not identified, all relationships for which the IBD probabilities is defined are considered when calculating the LR.

Edit IBD probabilities

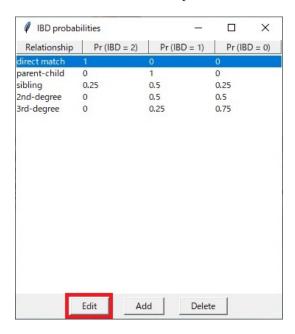
1. Press the 'IBD probabilities' button. Then the window 'IBD probabilities' will be open.



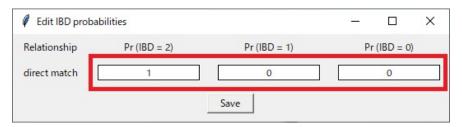
2. Select a relationship to change the IBD probabilities.



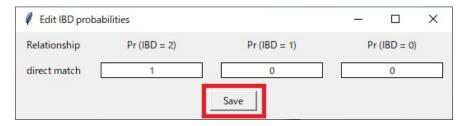
3. Press the 'Edit' button. Then the window 'Edit IBD probabilities' will be open.



4. Enter arbitrary probabilities of IBD = 2, IBD = 1, and IBD = 0.

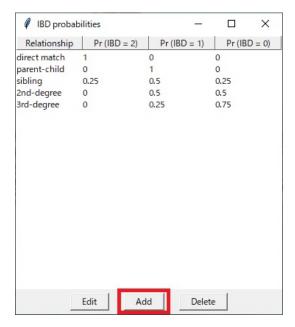


5. Press the 'Save' button.

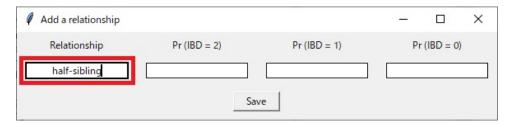


Add a relationship

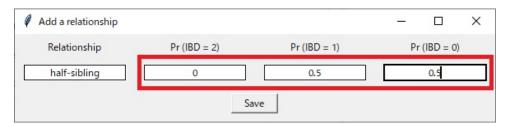
1. Press the 'Add' button in the window 'IBD probabilities'. Then the window 'Add a relationship' will be open.



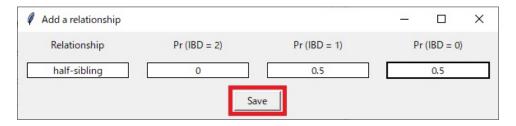
 $2. \ \, {\rm Enter}$ an arbitrary relationship.



3. Enter arbitrary probabilities of IBD = 2, IBD = 1, and IBD = 0.

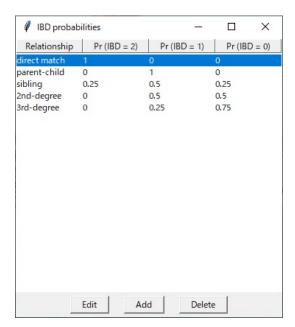


4. Press the 'Save' button.

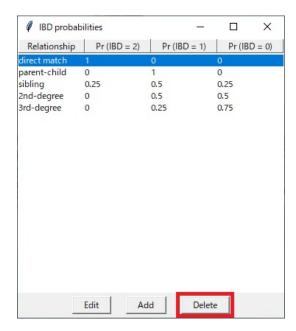


${\bf Delete}~{\bf a}~{\bf relationship}$

 $1. \ \, \text{Select}$ a relationship in the window 'IBD probabilities'.

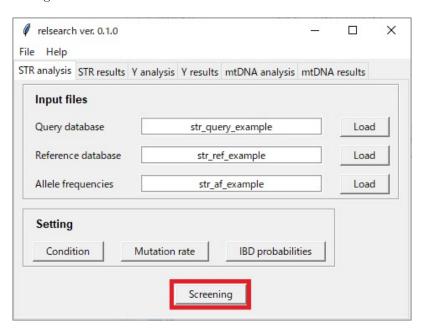


2. Press the 'Delete' button. Then the selected relationship will be deleted.

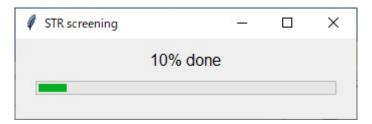


Perform screening

1. Press the 'Screening' button.

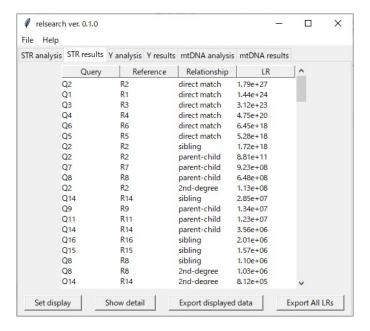


2. Wait until the screening finishes.



Check the result of screening

1. After finishing the screening, the result of the screening will be displayed.

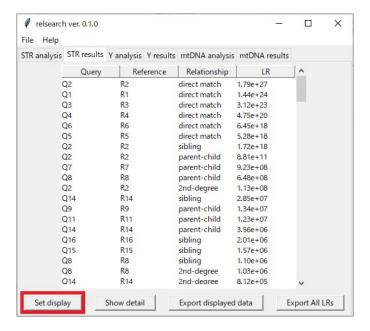


Note: Default display

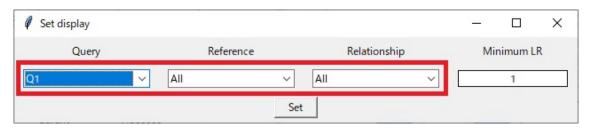
- All query names
- All reference names
- All relationships
- Likelihood ratio (LR) > 1
- Descending order of LR

Change the displayed information

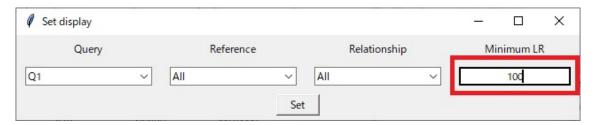
1. Press the 'Set display' button. Then the window 'Set display' will be open.



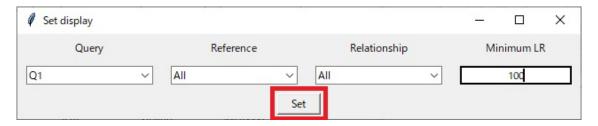
2. Select a query name, a reference name, and a relationship.



3. Enter the minimum LR.

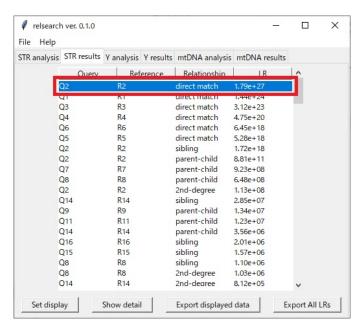


4. Press the 'Set' button.

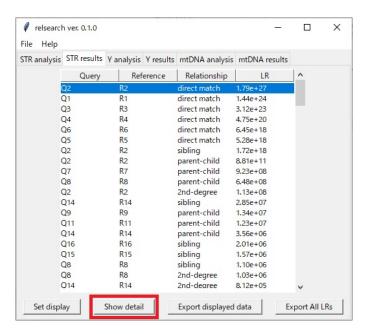


Show the result of a query-reference pair in detail

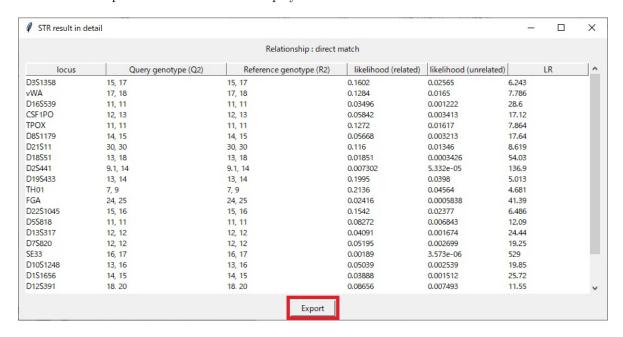
1. Select a query-reference pair.



2. Press the 'Show detail' button. Then the window 'STR result in detail' will be open.

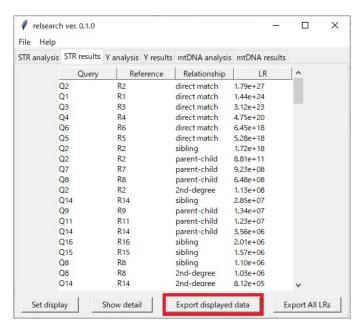


3. Press the 'Export' button to save the displayed data.

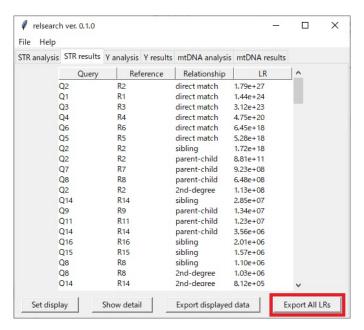


Export data

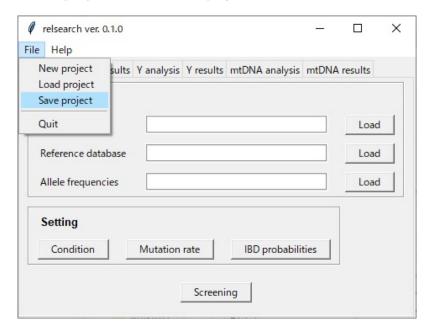
1. Press the 'Export displayed data' to save the displayed data in the tab 'STR results'.



2. Press the 'Export All LRs' to save all LR values.



3. Select $File \rightarrow Save project$ to save current project data.



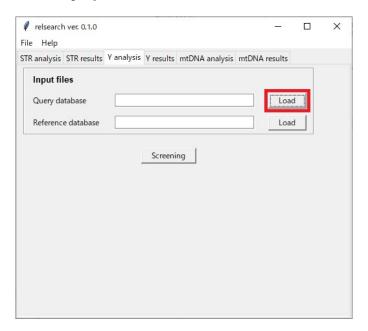
Note

- Users can load the saved project data. Select File -> Load project.
- Users can start a new project. Select File -> New project.

Y-chromosomal STR

Load query database

1. Press the 'Load' button for query database.



2. Select a .csv file for query database. Format of query database is shown in Fig. 4.

Sample Name	DYS456	DYS390	DYS438	DYS392	DYS518	DYS570	DYS437	DYS385
Q1	15	24	10	11	37	17	14	13,17
Q2	15	22	13	13	38	19	14	10,20
Q3	15	25	10	11	37	17	14	14
Q4	15	26	11	11	37	16	14	13
Q5	15	23	11			18		13,17
Q6				11	40		14	14
Q7	15	23	10	11	38	16	14	11,19
Q8	16	24	10	11	38	19	14	13,15
Q9	14	24	11	14	37	18	14	13,18
Q10	16	24	11	14		16	15	13,17

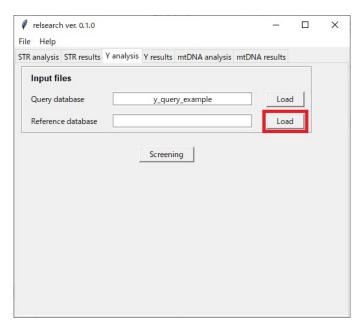
Figure 4: Format of query database for Y-STR

Note: format of query database

- This file must include information regarding 'Sample Name' and each marker.
- There is one column in each marker.
- In the marker with more than one allele, each allele must be separated by a comma (e.g., DYS385).
- The marker with an empty cell (e.g., DYS456 of sample 'Q6' in Fig. 4) is regarded as the 'ignored loci'.
- An example file named "y_query_example.csv" is located at extdata > examples.

Load reference database

1. Press the 'Load' button for reference database.



2. Select a .csv file for reference database. Format of reference database is shown in Fig. 5.

Sample Name	DYS456	DYS390	DYS438	DYS392	DYS518	DYS570	DYS437	DYS385
R1	15	24	10	11	37	17	14	13,17
R2	15	22	13	13	38	19	14	10,20
R3	15	25	10	11	37	17	14	14,17
R4	15	26	11	11	37	16	14	13,16
R5	15	23	11	12	42	18	14	13,17
R6	16	26	10	11	40	20	14	14
R7	15	23	10	11	38	16	14	11,19
R8	16	24	10	11	38	19	14	13,15
R9	14	24	11	14	37	18	14	13,18
R10	16	24	11	14	34	16	15	13,17

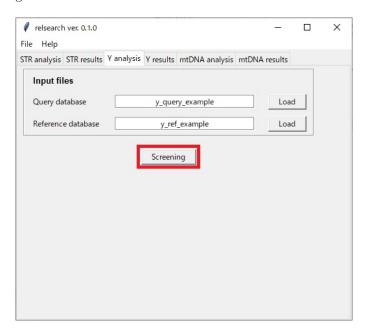
Figure 5: Format of reference database for Y-STR

Note: format of reference database

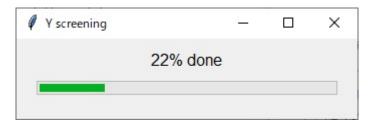
- This file must include information regarding 'Sample Name' and each marker.
- There is one column in each marker.
- In the marker with more than one allele, each allele must be separated by a comma (e.g., DYS385).
- An example file named "y_ref_example.csv" is located at extdata > examples.

Perform screening

1. Press the 'Screening' button.

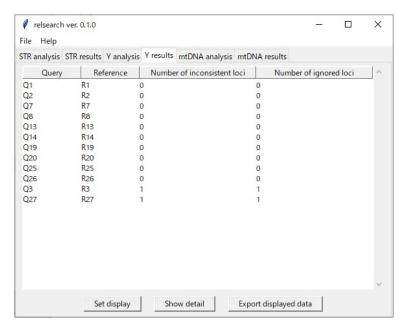


2. Wait until the screening finishes.



Check the result of screening

1. After finishing the screening, the result of the screening will be displayed.

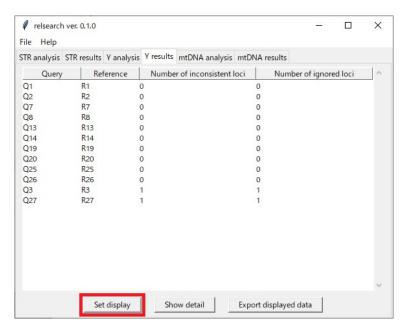


Note: Default display

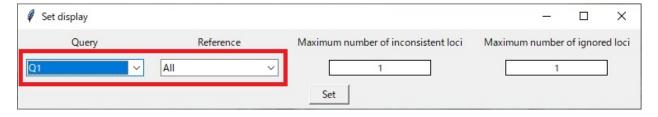
- All query names
- All reference names
- Number of inconsistent loci : 0 or 1
- Ascending order of the number of inconsistent loci and the number of ignored loci

Change the displayed information

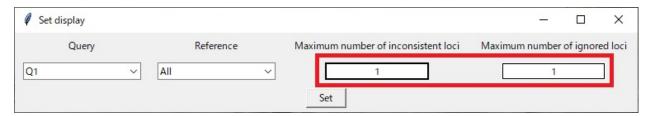
1. Press the 'Set display' button. Then the window 'Set display' will be open.



2. Select a query name and a reference name.



3. Enter the minimum number of inconsistent loci and the maximum number of loci explained by dropouts.

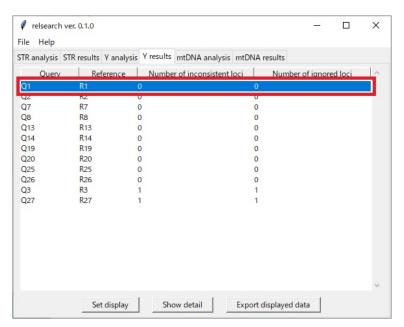


4. Press the 'Set' button.

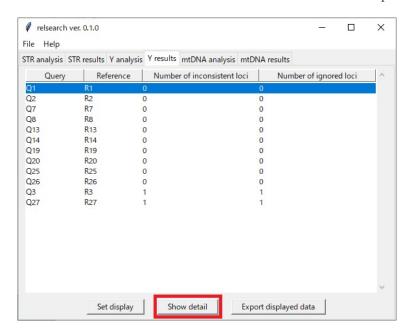


Show the result of a query-reference pair in detail

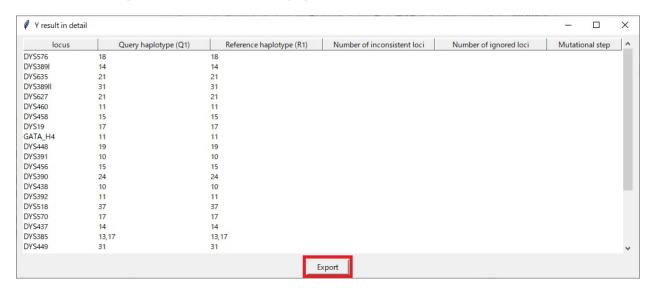
1. Select a query-reference pair.



2. Press the 'Show detail' button. Then the window 'Y result in detail' will be open.

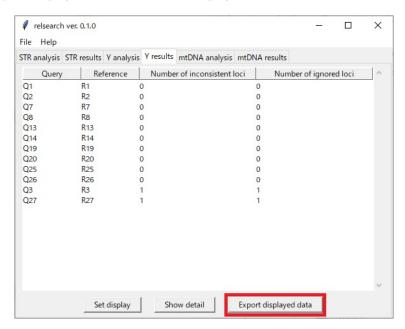


3. Press the 'Export' button to save the displayed data.

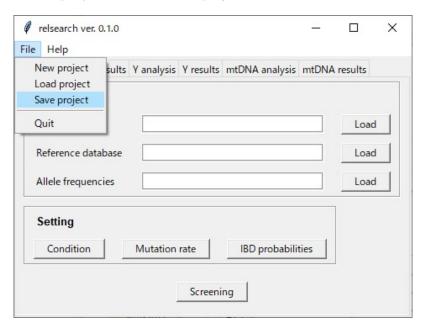


Export data

1. Press the 'Export displayed data' to save the displayed data in the tab 'Y results'.



2. Select File -> Save project to save current project data.



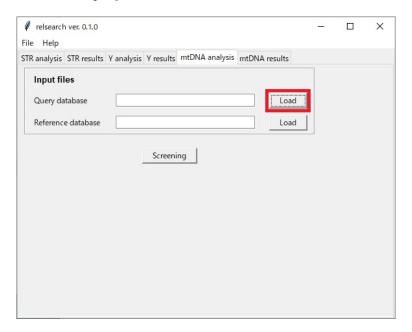
Note

- Users can load the saved project data. Select File -> Load project.
- Users can start a new project. Select ${\bf File}$ -> ${\bf New}$ ${\bf project}$.

Mitochondrial DNA

Load query database

1. Press the 'Load' button for query database.



2. Select a .csv file for query database. Format of query database is shown in Fig. 6.



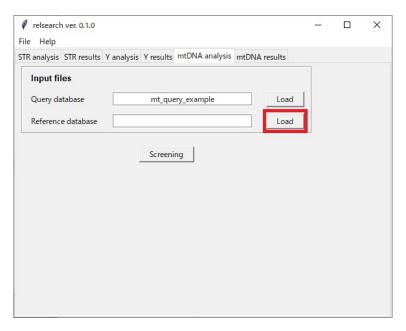
Figure 6: Format of query database for mtDNA

Note: format of query database

- This file must include information regarding 'Sample Name', 'Range', and 'Haplotype'.
- An example file named "mt_query_example.csv" is located at extdata > examples.

Load reference database

1. Press the 'Load' button for reference database.



2. Select a .csv file for reference database. Format of reference database is shown in Fig. 7.

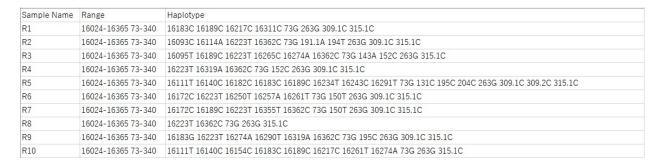


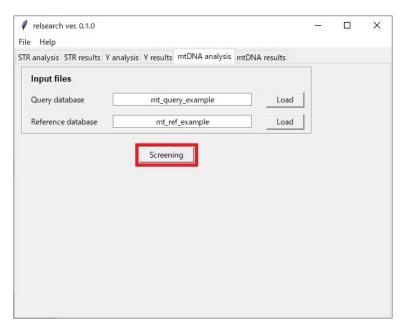
Figure 7: Format of reference database for mtDNA

Note: format of reference database

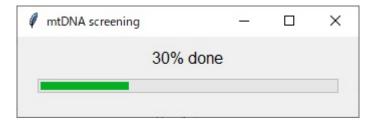
- This file must include information regarding 'Sample Name', 'Range', and 'Haplotype'.
- An example file named "mt_ref_example.csv" is located at extdata > examples.

Perform screening

1. Press the 'Screening' button.

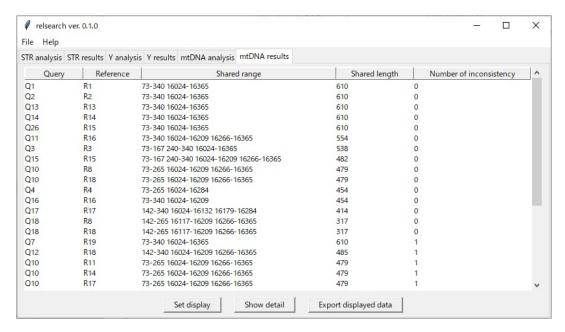


2. Wait until the screening finishes.



Check the result of screening

1. After finishing the screening, the result of the screening will be displayed.

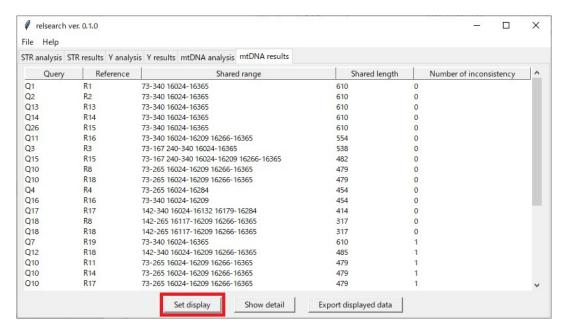


Note: Default display

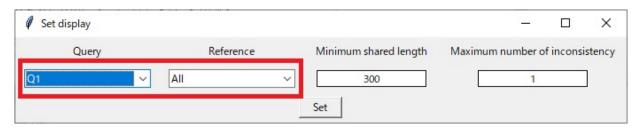
- All query names
- All reference names
- Shared length > 300
- Number of inconsistency: 0 or 1
- Descending order of the shared length
- Ascending order of the number of inconsistency

Change the displayed information

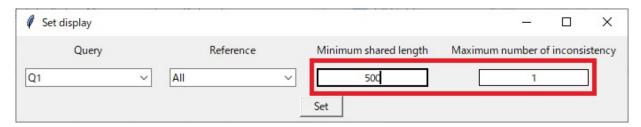
1. Press the 'Set display' button. Then the window 'Set display' will be open.



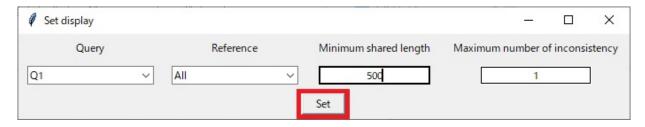
2. Select a query name and a reference name.



3. Enter the minimum shared length and the maximum number of inconsistency.

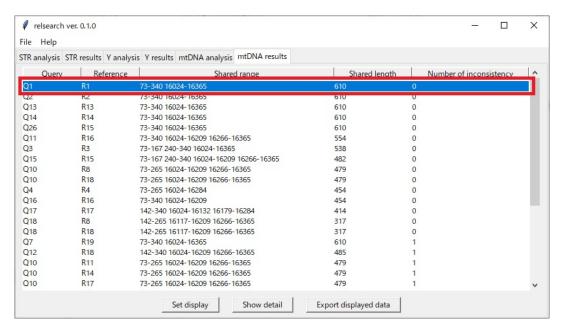


4. Press the 'Set' button.

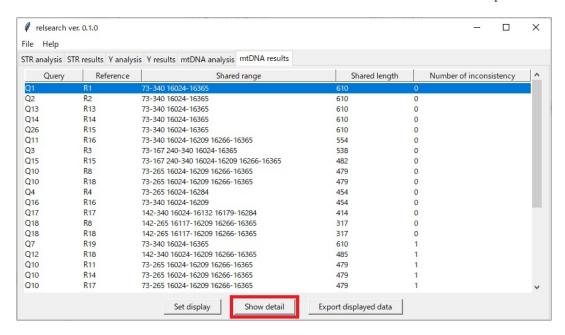


Show the result of a query-reference pair in detail

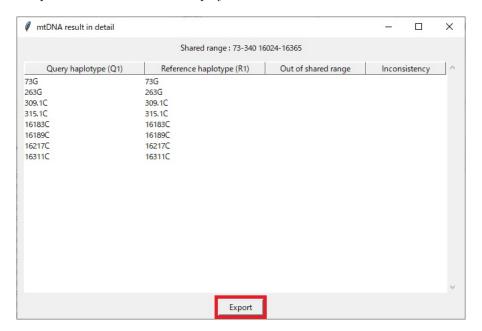
1. Select a query-reference pair.



2. Press the 'Show detail' button. Then the window 'mtDNA result in detail' will be open.

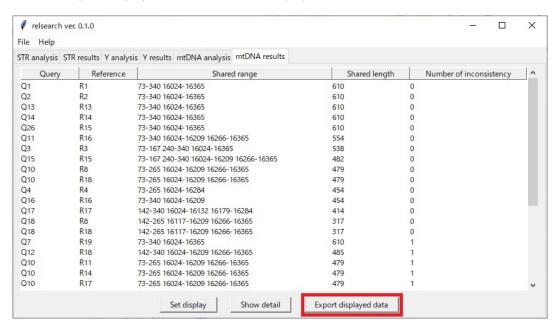


3. Press the 'Export' button to save the displayed data.

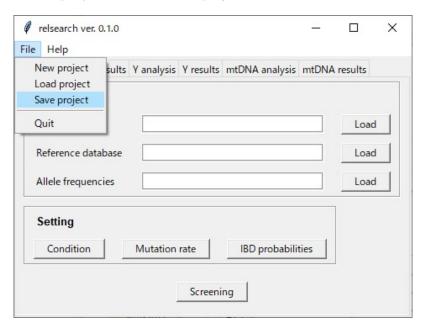


Export data

1. Press the 'Export displayed data' to save the displayed data in the tab 'mtDNA results'.



2. Select File -> Save project to save current project data.



Note

- Users can load the saved project data. Select File -> Load project.
- Users can start a new project. Select File -> New project.