

RelSearch version 1.0.0 user manual

Contents

1	Initial setup	3
2	Quick guide	4
3	Files	9
3.1	STR: Victim database	9
3.2	STR: Reference database	9
3.3	STR: Allele frequencies	10
3.4	Y-STR: Victim database	10
3.5	Y-STR: Reference database	10
3.6	mtDNA: Victim database	11
3.7	mtDNA: Reference database	11
4	Setting	12
4.1	Criteria	12
4.1.1	STR	12
4.1.2	Y-STR	12
4.1.3	mtDNA	12
4.2	Relationship	12
4.2.1	Edit	13
4.2.2	Add	13
4.2.3	Delete	13
4.2.4	Reset	13
4.2.5	Family tree	14
4.3	Mutation rate	14
4.3.1	Edit	14
4.3.2	Add	14
4.3.3	Delete	14
4.3.4	Reset	14
4.4	Other parameters	14
5	Result	15
5.1	Summary	15
5.1.1	Default display	15
5.1.2	Identified pairs	15

5.1.3	Multiple candidates	15
5.1.4	Not support lineage	15
5.1.5	Minimum LR displayed	15
5.1.6	Download summary data	15
5.2	Selected data in detail	16
5.3	Analysis conditions	16
6	Other functions	17
6.1	Project	17
6.1.1	New project	17
6.1.2	Load project	17
6.1.3	Save project	17
6.2	View database	17
6.3	Example files	17
7	Computational principle	18
7.1	STR	18
7.1.1	Likelihood ratio	18
7.1.2	Allelic drop-out	18
7.1.3	Mutation	18
7.2	Y-STR	18
7.3	mtDNA	18
	Reference	19

1 Initial setup

1. Ensure that R ($\geq 4.4.0$) is installed. It is available from the [R Development Core Team website](#).
2. Begin an R session.
3. Execute the following command in R to install required packages.

```
install.packages(c("shiny", "data.table", "dplyr", "DT", "magrittr",  
  "pandoc", "pedtools", "Rcpp", "ribd", "rmarkdown", "shinyFeedback",  
  "shinyjs", "shinythemes", "waiter"))
```

4. Download "relsearch_1.0.0.zip" from the [GitHub repository page](#).
5. Install "relsearch_1.0.0.zip" from "Install package(s) from local files..."

2 Quick guide

1. Execute the following commands in R to start GUI.

```
library(relsearch)
relsearch()
```

2. Load files from each “**Browse...**” button.

The screenshot shows the relsearch ver. 0.11.0 GUI. The interface is organized into three main columns: STR, Y-STR, and mtDNA. Each column contains a 'Victim database' section with a 'Browse...' button (highlighted with an orange box) and a 'Reference database' section with a 'Browse...' button (also highlighted with an orange box). Below these, there is an 'Allele frequencies' section with a 'Browse...' button (highlighted with an orange box). At the bottom left of the main content area is a blue 'Analyze' button. The top of the window features a blue navigation bar with the following menu items: Load, Result, Project (with a dropdown arrow), Database (with a dropdown arrow), Settings (with a dropdown arrow), Example files, and Manual.

Note See section [Files](#) for information on each file.

3. Click the “**Analysis**” button.

relsearch ver. 0.11.0

LoadResultProject ▾Database ▾Settings ▾Example filesManual

STR

Victim database

Browse...

str_victim_example.csv

Upload complete

Reference database

Browse...

str_ref_example.csv

Upload complete

Allele frequencies

Browse...

str_af_example.csv

Upload complete

Analyze

Y-STR

Victim database

Browse...

y_victim_example.csv

Upload complete

Reference database

Browse...

y_ref_example.csv

Upload complete

mtDNA

Victim database

Browse...

mt_victim_example.csv

Upload complete

Reference database

Browse...

mt_ref_example.csv

Upload complete

4. After finishing the analysis, the result window is automatically opened.

relesearch ver. 0.11.0
Load
Result
Project
Database
Settings
Example files
Manual

Result

Summary
Selected data in detail
Analysis conditions

Display setting

Default display

Identified pairs

Multiple candidates

Not support lineage

Minimum LR displayed

100

Apply

Download

Show
10
entries

Search:

Victim	Reference	Assumed relationship	LR	Estimated relationship	Paternal lineage	Maternal lineage
All	All	All	All	All	All	All
V7	R7	Father_Son	9.23e+8	Father_Son	Support	Support
V8	R8	Father_Son	6.48e+8	Father_Son	Support	Not support
V14	R14	Brother_Brother	2.85e+7	Brother_Brother	Support	Support
V9	R9	Father_Son	1.34e+7	Father_Son	Support	Not support
V11	R11	Mother_Daughter	1.25e+7			Not support
V16	R16	Brother_Sister	2.01e+6	Brother_Sister		Support
V15	R15	Brother_Sister	1.57e+6	Brother_Sister		Support
V7	R30	Brother_Brother	1.02e+5	Brother_Brother	Support	Support
V10	R10	Mother_Daughter	4.49e+4	Mother_Daughter		Support
V12	R12	Mother_Daughter	3.09e+3	Mother_Daughter		Support

Showing 1 to 10 of 12 entries

Previous
1
2
Next

Note See section [Result](#) for navigation of the result window.

5. Select **Project** ► **Save project**.

research ver. 0.11.0

Load

Result

Project ▾

Database ▾

Settings ▾

Example files

Manual

New project

Load project

Save project

Result

Summary

Selected data in detail

Analysis conditions

Display setting

Default display

Identified pairs

Multiple candidates

Not support lineage

Minimum LR displayed

100

Apply

Download

Show 10 entries

Search:

Victim	Reference	Assumed relationship	LR	Estimated relationship	Paternal lineage	Maternal lineage
All	All	All	All	All	All	All
V7	R7	Father_Son	9.23e+8	Father_Son	Support	Support
V8	R8	Father_Son	6.48e+8	Father_Son	Support	Not support
V14	R14	Brother_Brother	2.85e+7	Brother_Brother	Support	Support
V9	R9	Father_Son	1.34e+7	Father_Son	Support	Not support
V11	R11	Mother_Daughter	1.25e+7			Not support
V16	R16	Brother_Sister	2.01e+6	Brother_Sister		Support
V15	R15	Brother_Sister	1.57e+6	Brother_Sister		Support
V7	R30	Brother_Brother	1.02e+5	Brother_Brother	Support	Support
V10	R10	Mother_Daughter	4.49e+4	Mother_Daughter		Support
V12	R12	Mother_Daughter	3.09e+3	Mother_Daughter		Support

Showing 1 to 10 of 12 entries

Previous

1

2

Next

127.0.0.1:4409/#tab-1131-3

6. Enter the project name and click the “**Save as**” button.

research ver. 0.11.0 Load Result Project Database Settings Example files Manual

Save project

Enter the project name

Save as

Note The saved project can be loaded from **Project ► Load project**. How to handle projects is described in section [Other functions](#).

3 Files

3.1 STR: Victim database

SampleName	D3S1358	D3S1358.1	vWA	vWA.1	D16S539	D16S539.1	CSF1PO	CSF1PO.1	TPOX	TPOX.1
Victim1	17	17	17		11		12			
Victim2	17		18	18			10	11	8	
Victim3	15	17	17	18	10	10	10	12	8	11
Victim4	15	18	15	18	9	12	11	12	8	8
Victim5	15	15	18		9	9	10		9	11
Victim6	16	16	17				10	13	8	8
Victim7	14	15	17	18	12				8	8
Victim8	16	18	17	18	11		9	12	8	
Victim9	15	15	16	18	11	11	12	12	8	11
Victim10	15	16	18	19	9	10	12	12	8	11

Note

- File type: .csv
- This file requires the column “SampleName” and columns for each marker (two columns in each).
- The marker with two empty cells (e.g., TPOX of sample “Victim1” in the above table) is ignored when calculating the likelihood ratio.
- The marker with one empty cell (e.g., vWA of sample “Victim1” in the above table) can be regarded as both homozygote (i.e., no drop-out) and heterozygote with drop-out of one allele when calculating the likelihood ratio.

3.2 STR: Reference database

Note

- File type: .csv
- This file requires the columns “SampleName”, “Relationship”, and columns for each marker (two columns in each).
- The names of the relationship need to be pre-defined in **Settings Relationships** (see section [Setting](#)).
- If a reference has multiple missing family members, add rows for each relationship of the members (e.g., sibling and parent-child of the sample “Reference7” in the above table).

- The marker with two empty cells (e.g., CSF1PO of sample “Reference10” in the above table) is ignored when calculating the likelihood ratio.
- The marker with one empty cell (e.g., vWA of sample “Reference10” in the above table) can be regarded as both homozygote (i.e., no drop-out) and heterozygote with drop-out of one allele when calculating the likelihood ratio.

3.3 STR: Allele frequencies

Note

- File type: .csv
- This file requires the columns “Allele” and columns for each marker (one column in each).

3.4 Y-STR: Victim database

Note

- File type: .csv
- This file requires the column “SampleName” and columns for each marker (one column in each).
- In the marker with more than one allele, each allele must be separated by a comma without any spaces (e.g., DYS385).
- The marker with an empty cell (e.g., DYS456 of sample ‘Victim6’) is ignored for analysis.

3.5 Y-STR: Reference database

Note

- File type: .csv
- This file requires the column “SampleName”, “Relationship”, and columns for each marker (one column in each).
- The relationship of a missing family member should be designated in the column “Relationship”. The name of the relationship should be defined in Settings > Relationships.
- If a reference has multiple missing family members, add rows for each relationship of the members (e.g., sibling and parent-child of the sample “Reference7” in the above table).

- In the marker with more than one allele, each allele must be separated by a comma without any spaces (e.g., DYS385).
- The marker with an empty cell (e.g., DYS518 of sample ‘Reference10’) is ignored for analysis.

3.6 mtDNA: Victim database

Note

- File type: .csv
- This file requires the column “SampleName”, “Range”, and “Haplotype”.

3.7 mtDNA: Reference database

Note

- File type: .csv
- This file requires the column “SampleName”, “Relationship”, “Range”, and “Haplotype”.
- The relationship of a missing family member should be designated in the column “Relationship”. The name of the relationship should be defined in Settings > Relationships.
- If a reference has multiple missing family members, add rows for each relationship of the members (e.g., sibling and parent-child of the sample “Reference7” in the above table).

4 Setting

4.1 Criteria

You can set the following criteria to support the assumed relationship. Press the **Save** button to reflect the changes. Press the **Reset** button if you want to return to default values.

4.1.1 STR

- **Minimum LR:** The minimum likelihood ratio that supports the assumed relationship.

4.1.2 Y-STR

- **Maximum number of mismatched loci:** The upper limit of the number of mismatched loci between victim and reference profiles to support paternal lineage.
- **Maximum total mutational steps:** The upper limit of total mutational steps (i.e., sum of mutational steps of all loci) between victim and reference profiles to support paternal lineage.

4.1.3 mtDNA

- **Maximum number of inconsistency:** The upper limit of inconsistency between victim and reference profiles to support maternal lineage.

4.2 Relationship

The names of the relationship in each reference profile need to be pre-defined. There are 44 pre-defined victim-reference relationships as the default.

Information on the table is as follows.

- **$\Pr(\text{IBD} = 2)$:** The probability that two alleles are identity by descent (IBD).
- **$\Pr(\text{IBD} = 1)$:** The probability that one allele is identity by descent (IBD).
- **$\Pr(\text{IBD} = 0)$:** The probability that zero allele is identity by descent (IBD).
- **Paternal lineage:** Information on whether each relationship is paternal lineage or not.
- **Maternal lineage:** Information on whether each relationship is maternal lineage or not.

You can edit, add, and delete the relationships from the left sidebar.

4.2.1 Edit

You can edit the defined relationship name from the **Edit** button.

In the pop-up window, select a relationship from the defined relationships, and enter a new name of the selected relationship. Press the **Save** button to reflect the new name.

4.2.2 Add

You can define a new relationship from the **Add** button.

In the pop-up window, you have to set the relationship name and make a family tree to define the relationship between a victim and a reference.

The following is a list of functions to make a family tree.

- **Add a person:** The function to add an unknown person (UK) to the list of persons.
- **Delete a person:** The function to delete an unknown person (UK) from the list of person.
- **Sex:** The function to select a biological sex (male: M or female: F).
- **Father:** The function to select a person who is the father.
- **Mother:** The function to select a person who is the mother.
- **Founder:** The function to select whether the person is a founder of the family tree or not.

After defining a family tree, press the **View family tree** button to check whether the family tree is correctly set up or not.

If there is no problem in the family tree, press the **Save** button.

Note Information on the IBD probabilities, paternal lineage, and maternal lineage in the new relationship is automatically determined based on your-defined family tree.

4.2.3 Delete

You can delete a defined relationship name from the **Delete** button.

In the pop-up window, select a relationship from the defined relationships. Press the **Save** button to delete the selected relationship.

4.2.4 Reset

If you want to return to default settings, press the **Reset** button. In the pop-up window, press the **Restore default** button.

4.2.5 Family tree

You can check the family tree of each defined relationship. Select a row from the relationship table and press the **Family tree** button.

4.3 Mutation rate

Mutation rates of each autosomal STR marker need to be set to calculate the likelihood ratio (LR). The default mutation rates considering the difference between paternal and maternal and mutational steps (-1, +1, -2, and +2) are reported by Morimoto et al [1]. You can edit, add, and delete these mutation rates from the left sidebar.

4.3.1 Edit

You can edit the mutation rates from the **Edit** button.

In the pop-up window, select a locus and change the mutation rates. Press the **Save** button to reflect the changes.

4.3.2 Add

You can add a locus from the **Add** button.

In the pop-up window, enter a locus name and enter the mutation rates. Press the **Save** button to add the locus.

4.3.3 Delete

You can delete a locus from the **Delete** button.

In the pop-up window, select a locus. Press the **Save** button to delete the selected locus.

4.3.4 Reset

If you want to return to default settings, press the **Reset** button. In the pop-up window, press the **Restore default** button.

4.4 Other parameters

You can set the minimum allele frequency which is used for the unobserved alleles. Press the **Save** button to reflect the changes. Press the **Reset** button if you want to return to the default value.

5 Result

5.1 Summary

After finishing the analysis, summary data is automatically displayed. You can change the displayed data from the left sidebar.

5.1.1 Default display

Default display depends on whether the input data includes autosomal STR data or not.

If the input data includes autosomal STR data, the data where the LR exceeds the minimum LR (see section [Setting](#)) is displayed.

If the input data is composed of only Y-STR and/or mtDNA data, the data that satisfies the criteria to support paternal/maternal lineage (see section [Setting](#)) is displayed.

If the number of data that satisfies the criteria is more than 10,000, the top 10,000 data is displayed.

5.1.2 Identified pairs

If the data supports that a victim (or a reference) has the assumed relationship with only one reference (or victim), the victim-reference pair is categorized as the identified pair.

5.1.3 Multiple candidates

If the data supports that a victim (or a reference) has the assumed relationship with multiple references (or victims), the victim-reference pair is categorized as the multiple candidates.

5.1.4 Not support lineage

If the LR exceeds the minimum LR but the data does not satisfy the criteria to support paternal/maternal lineage when assuming the paternal/maternal lineage relative, the victim-reference pair is categorized as not support lineage.

5.1.5 Minimum LR displayed

You can change the minimum LR displayed. The lower limit is $LR = 1$. If the number of data where the LR exceeds the minimum LR is more than 10,000, the top 10,000 data is displayed.

5.1.6 Download summary data

You can download summary data as a .csv file. Press the **Downlaod** button in Summary tab.

5.2 Selected data in detail

The data of a selected victim-reference pair is displayed in **Selected data in detail** tab.

1. Select a row in the summary table.
2. Go to **Selected data in detail** tab.

In **Selected data in detail** tab, you can check and download the analyzed data of STR, Y-STR, and mtDNA in the selected victim-reference pair.

5.3 Analysis conditions

You can check the following analysis conditions in **Analysis conditions** tab.

- **Database:** Input .csv file names of each database are displayed.
- **Allele probability:** Unobserved alleles in the population database are displayed, if any. The frequency of these alleles is set to the user-defined minimum allele frequency and the sum of the frequencies in each locus is corrected to 1.
- **Criteria:** The user-defined criteria to support the assumed relationship (see section [Setting](#)) is displayed.
- **Assumed relationship:** Information on the user-defined relationships (see section [Setting](#)) is displayed.
- **Mutation rate:** The user-defined mutation rates (see section [Setting](#)) are displayed.
- **Parameter:** The user-defined minimum allele frequency (see section [Setting](#)) is displayed.

6 Other functions

6.1 Project

6.1.1 New project

To start a new project, go to **Project ► New project** and press the **New project** button.

Important: The current project should be saved before starting a new project because the current project will be deleted.

6.1.2 Load project

To load a previous project, go to **Project ► Load project**. Select a project file from the **Browse...** button and press the **Load project** button.

Important: The current project should be saved before loading another project because the current project will be deleted.

6.1.3 Save project

To save the current project, go to **Project ► Save project**. Enter the project name and press the **Save as** button.

6.2 View database

Each loaded database can be checked in **Database** tab.

6.3 Example files

Example files for each database can be downloaded from the **Example files** tab.

7 Computational principle

7.1 STR

7.1.1 Likelihood ratio

The likelihood ratio (LR) is calculated for each victim-reference pair by assuming the following two hypotheses:

H_1 : The victim and the reference are a certain relationship,

H_2 : The victim and the reference are unrelated.

The equation of the LR is as follows:

$$LR = \frac{Pr(G_v, G_r | H_1)}{Pr(G_v, G_r | H_2)}$$

where G_v and G_r denote the genotypes of the victim and the reference, respectively.

In RelSearch, users can set the minimum threshold of the LR to support H_1 . The default threshold is 100, which is determined based on the verbal scale “Moderate Support” written in the Scientific Working Group on DNA Analysis Methods (SWGDAM) guideline [2].

The method of calculating the LR uses the probabilities that two, one and zero alleles are identity by descent (IBD). If allelic drop-out and mutations are not considered, the LR is calculated according to the method of Wenk et al [3].

$$Pr(G_v, G_r | H) = Pr(G_v, G_r | IBD = 2)Pr(IBD = 2) + Pr(G_v, G_r | IBD = 1)Pr(IBD = 1) + Pr(G_v, G_r | IBD = 0)Pr(IBD = 0)$$

7.1.2 Allelic drop-out

7.1.3 Mutation

7.2 Y-STR

RelSearch investigates whether the Y-STR profiles of each victim-reference pair support paternal lineage. There are two criteria to support paternal lineage as follows:

- Maximum number of mismatched loci: The upper limit of the number of mismatched loci between victim and reference profiles to support paternal lineage.
- Maximum total mutational steps: The upper limit of total mutational steps (i.e., sum of mutational steps of all loci) between victim and reference profiles to support paternal lineage.

7.3 mtDNA

Reference

- [1] C. Morimoto, H. Tsujii, S. Manabe, S. Fujimoto, E. Hirai, Y. Hamano, K. Tamaki, Development of a software for kinship analysis considering linkage and mutation based on a bayesian network, *Forensic Science International: Genetics* 47 (2020) 102279.
- [2] Scientific Working Group on DNA Analysis Methods, Recommendations of the SWGDAM ad hoc working group on genotyping results reported as likelihood ratios., (2018).
- [3] R.E. Wenk, M. Traver, F.A. Chiafari, Determination of sibship in any two persons, *Transfusion* 36 (1996) 259–262.