

EasyKin: a flexible and user-friendly online tool for forensic
kinship testing and missing person identification

User guide

Content

1. Introduction	1
1.1 Object	1
1.2 Terms and abbreviations.....	1
1.3 references.....	1
2. General description.....	2
2.1 Main functions.....	2
2.2 Performance.....	2
3. Usage.....	3
3.1 Input.....	3
3.2 Output	6
4. Examples	10
1) Allele frequency and kits/markers.....	10
2) Standard duo paternity testing cases	11
3) Standard trio paternity testing cases.....	12
4) Both parents are suspected	13
5) Multi-individual testing.....	14
6) User defined pedigrees	15
7) LR distribution and system power	15
8) Pedigree pruning	16
9) LR calculation for single case	17

1. Introduction

1.1 Object

The determination of genetic relatedness is frequently adopted in several forensic applications, such as kinship confirmation after separation, inheritance disputes between illegitimate children, immigration cases, and personal identification of missing persons, unknown bodies, and disaster victims (DVI)[1–3]. Generally, two questions need to be answered before testing: (i) how many markers are needed and (ii) how many reference relatives and who should be genotyped? Choosing the most informative references and marker sets may not only increase the efficacy and accuracy, but also reduce cost at the same time. Despite that several useful tools, such as *EasyDNA*[4], *Bonaparte*[5, 6], *Converge Software*[7], *Familias*[8], and *forrel*[9] have been developed, they haven't addressed these issues satisfactorily. Therefore, we develop *Easykin*, a flexible and user-friendly online tool, to provide a fast, easy, and powerful tool for kinship analysis and missing person identification for enforcement officers, forensic practitioner, and researchers.

1.2 Terms and abbreviations

person of interest	POI
short tandem repeat	STR
likelihood ratio	LR
sensitivity	Sen
specificity	Spe
positive predictive value	PPV
negative predictive value	NPV
false positive rate	FPR
false negative rate	FNR

1.3 references

1. Ge J, Budowle B, Chakraborty R (2011) Choosing Relatives for DNA Identification of Missing

-
- Persons. *J Forensic Sci* 56: . doi: 10.1111/j.1556-4029.2010.01631.x
 2. Kling D, Tillmar AO, Egeland T (2014) Familias 3 - Extensions and new functionality. *Forensic Sci Int Genet* 13: . doi: 10.1016/j.fsigen.2014.07.004
 3. Pinto N, Simões R, Amorim A, Conde-Sousa E (2019) Optimizing the information increase through the addition of relatives and genetic markers in identification and kinship cases. *Forensic Sci Int Genet* 40:210–218 . doi: 10.1016/j.fsigen.2019.02.019
 4. Fung WK (2003) User-friendly programs for easy calculations in paternity testing and kinship determinations. *Forensic Sci Int* 136:22–34 . doi: 10.1016/S0379-0738(03)00218-4
 5. Bruijning-van Dongen CJ, Slooten K, Burgers W, Wiegerinck W (2009) Bayesian networks for victim identification on the basis of DNA profiles. *Forensic Sci Int Genet Suppl Ser* 2:466–468 . doi: 10.1016/j.fsigss.2009.08.024
 6. Slooten K (2011) Validation of DNA-based identification software by computation of pedigree likelihood ratios. *Forensic Sci Int Genet* 5:308–315 . doi: 10.1016/j.fsigen.2010.06.005
 7. Converge Forensic Analysis Software - CN.
//www.thermofisher.cn/cn/zh/home/industrial/forensics/human-identification/forensic-dna-analysis/forensic-dna-data-interpretation/converge-forensic-analysis-software.html. Accessed 25 May 2021
 8. Egeland T, Mostad PF, Mervang B, Stenersen M (2000) Beyond traditional paternity and identification cases. Selecting the most probable pedigree. *Forensic Sci Int* 110:47–59 . doi: 10.1016/S0379-0738(00)00147-X
 9. Vigeland MD, Marsico FL, Herrera Piñero M, Egeland T (2020) Prioritising family members for genotyping in missing person cases: A general approach combining the statistical power of exclusion and inclusion. *Forensic Sci Int Genet* 49:102376 . doi: 10.1016/j.fsigen.2020.102376

2. General description

2.1 Main functions

1) Quick construction of pedigrees based on two competitive hypotheses (H1 and H2); 2) LR calculation assuming independent between markers; 3) System power estimation (Sen, Spe, PPV, NPV, FPR, FNR, inconclusive, and effectiveness) under either single or double thresholds based on simulated pedigrees; and 4) Pedigree pruning (thus generating a series of subsets of the original pedigree), which is useful for choosing the most informative subsets of reference relatives.

2.2 Performance

- 1) The type and number of reference relatives

Reference relatives include father/mother, 0-6 children (0-3 sons and 0-3 daughters), paternal grandparent(s), 0-3 paternal uncles, 0-3 paternal aunts, maternal grandparent(s), 0-3 maternal

uncles, 0-3 maternal aunts, 0-6 full siblings, 0-6 paternal half siblings, 0-6 maternal half siblings, 0-6 grandchildren (the children of son), 0-6 grandchildren (the children of daughter), 0-6 nephews/nieces (the children of brother) and 0-6 nephews/nieces (the children of sister). Several genetically unrelated individuals can also be included, e.g., spouse, the mother of paternal half sibling, the father of maternal half sibling, daughter-in-law, son-in-law, brother-in-law, and sister-in-law. Theoretically, more than one billion scenarios can be constructed, covering the majority of common cases. For more complex scenarios, say involving incest, users are encouraged to upload their own pedigrees under the instruction in the user guide. At least, one reference should be included, i.e. pairwise kinship testing.

2) Time cost

Only several seconds are needed to calculate the LRs of 100 pedigrees under “Equal” mutation model. It increases linearly with the number of STR markers. If “stepwise” is specified, the run time varies greatly depending on the numbers and types of reference relatives. Cases with full siblings but without their parents need relatively more time and the time cost increases exponentially with one more sibling. The same goes for cases with uncles/aunts but without grandparents.

3. Usage

3.1 Input

Settings specify the markers/kits, corresponding frequency data, STR mutation model, and the number of simulations.

- 1) **STR Frequency data.** 62 STRs from Chinese Han population are provided by default, i.e., D3S1358, CSF1PO, D2S441, D21S11, PentaE, D8S1179, D5S818, D19S433, D16S539, PentaD, vWA, D2S1338, D18S51, D22S1045, TH01, D12S391, TPOX, FGA, D13S317, D1S1656, D10S1248, D6S1043, D7S820, D6S477, D18S535, D19S253, D15S659, D11S2368, D20S470, D22.GATA198B05, D7S3048, D8S1132, D4S2366, D21S1270, D13S325, D9S925, D3S3045, D14S608, D10S1435, D17S1290, D5S2500, D7S1517, D3S1744, D2S1360, D6S474, D21S2055, D10S2325, SE33, D12ATA63, D1S1677, D11S4463, D1S1627, D3S4529, D6S1017, D4S2408, D17S1301, D1GATA113, D18S853, D20S482, D14S1434, D9S1122, D2S1776. Users can upload

your own data as long as you follow the required formats (Table 1). Users may first prepare the frequency data in Excel and save as another TAB delimited text file, which can then be imported into EasyKin.

Table 1. An example of STR frequency data format.

locus	alleles	frequency
D3S1358	14	0.0379
D3S1358	15	0.3295
D3S1358	16	0.3023
CSF1PO	10	0.2308
CSF1PO	11	0.2506
CSF1PO	12	0.3725

- 2) **Kits/Markers** can be a subset of STRs in the inputted frequency data. If so, tick the box *Restrict markers in following selected kit(s)*. If more than one kit is selected, markers overlapped in different kits are eliminated automatically. If the kits you are going to use are not listed in the selection box, you may type the markers one by one in the box *Input marker(s)*. Once selected, the number and detailed STRs are shown below the box. If all the STRs in the input file will be genotyped, just cancel the tick (*Restrict markers in following selected kit(s)*).
- 3) **STR mutation model:**
 - **Model:** the allowed values are " Stepwise " and " Equal " (the default);
 - **Rate1:** i.e., integer mutation rate (u_1), ranging from 0 to 1. Default is 0.002;
 - **Rate2:** i.e., non-integer mutation rate (u_2), ranging from 0 to 1, Default is 0.00000001;
 - **Range:** ratios for two-step mutations relative to one-step mutations, ranging from 0 to 1. Default is 0. 1.
 - **Ratio (male/female):** ratios of male-specific mutation rate to female-specific mutation rate. Default is 3.
 - **Modeling STR individually:** this may allow users to define the mutation metrics for each STR differently. The format is presented in Table 2. Again, save it as a TAB delimited text file.

Table 2. Format of mutation model for each STR.

Locus	Model	Rate1	Rate2	Range	Ratio
D1S1656	Stepwise	0.002	0.00000001	0.1	3
TPOX	Equal	0.001	0.00000001	0.05	3
D2S441	Stepwise	0.002	0.00000001	0.1	3
D2S1338	Stepwise	0.002	0.00000001	0.1	3
D3S1358	Stepwise	0.002	0.00000001	0.1	3
D4S2408	Stepwise	0.002	0.00000001	0.1	3
FGA	Stepwise	0.003	0.00000001	0.1	3
D5S818	Stepwise	0.002	0.00000001	0.1	3
CSF1PO	Stepwise	0.002	0.00000001	0.1	3
D6S1043	Stepwise	0.002	0.00000001	0.1	3
D8S1179	Stepwise	0.002	0.00000001	0.1	3
D10S1248	Stepwise	0.002	0.00000001	0.1	3

- 4) **Set seeds:** Default is 123.
- 5) **Simulations:** the more number of simulated families, the more smooth LR distribution will be (also more time cost). Default is 100.

Pedigrees specify the sex of POI, his/her known parent, and reference relatives. Users just need to click the selection box to include a specific reference relative.

- 1) **Sex** of person of interest (POI): the allowed values are " male" (the default) and " female ";
- 2) **known parent** of POI: the allowed values are “none” (the default) , “father” and “mother” ;
- 3) **Reference samples:** for the purpose of simplicity, 1st and 2nd degree relatives as well as several genetically unrelated individuals can be included for pedigree construction.
- 4) **User defined pedigrees:** For more complex scenarios, say involving incest, users are encouraged to upload their own pedigrees, just following the format in Table 3. The file is similar to a *.ped file of PLINK, where columns 1-6 are the hypothesis/family ID, the individual ID, the paternal ID, the maternal ID, the sex (“1” for “male” and “2” for “female”) and the status (“1” for “present” and “0” for “absent”), respectively. All the family members should be specified. If one’s parents are not included in the testing, name them “NA” in the corresponding cells in columns 3 and 4. It is recommended that put the person of interest (POI) in the first column and name him/her as “POI”. Finally, save it as a TAB delimited text file.

Table 3. An example format of user defined pedigrees (first cousin relationship testing).

#hypothesis	individual	father	mother	sex	status
H1	POI	fa	mo	2	1
H1	fa	gf	gm	1	0
H1	mo	NA	NA	2	0
H1	gf	NA	NA	1	0
H1	gm	NA	NA	2	0
H1	uncle	gf	gm	1	0
H1	aunt	NA	NA	2	0
H1	fc	uncle	aunt	1	1
H2	POI	random1	random2	2	1
H2	random1	NA	NA	1	0
H2	random2	NA	NA	2	0
H2	uncle	NA	NA	1	0
H2	aunt	NA	NA	2	0
H2	fc	uncle	aunt	1	1

For **Real case**, just input the genotypes of family members. The required format is shown in Table 4. The first row is the names of these samples (name them accordingly to the pedigrees/hypotheses). Sample absent from the pedigrees are not included for LR calculation. Rows 2 – n are the genotypes at each locus. Alleles are separated with comma. Note that both alleles of homozygotes should be specified. Finally, save it as a TAB delimited text file.

Table 4. Format of input genotypes

Locus	S1	S2	S5	POI
D1S1656	12,12	13,17	15,16	16,17
TPOX	8,8	8,11	8,8	8,8
D2S441	10,12	10,10	10,11	11,11
D2S1338	20,24	23,24	20,25	19,20
D3S1358	15,17	16,17	14,15	18,18
D4S2408	9,10	10,10	10,12	10,11

3.2 Output

There are four outputs for simulated families, namely hypotheses/pedigrees, LR distribution,

system power metrics, and system power metrics for pruned pedigrees. For real cases, LR values at each locus and the combined LR for the defined pedigree, as well as LR values for all pairs of references are displayed.

- 1) **Hypotheses:** Pedigrees based on two alternative hypotheses are constructed once reference relatives are specified (see an example in Fig. 1). Note that singleton relatives may not be printed in the pedigrees but they are included in the LR calculation.

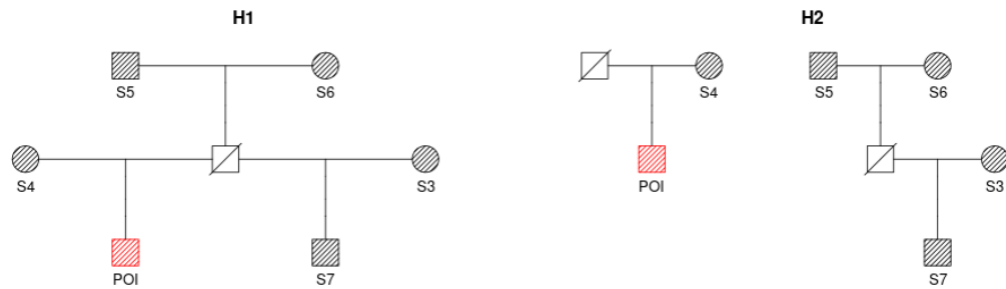


Fig. 1. An example of pedigrees based on two alternative hypotheses. Circles are females and squares are males. Shadows represent references who are available in the testing while blank represent unavailable references. POI is usually colored red.

- 2) **LR distribution** is a histogram of $\log_{10}(\text{LR})$ under H1 and H2, respectively. Along with the histogram is a curve of probability density after normal fitting (Fig. 2).

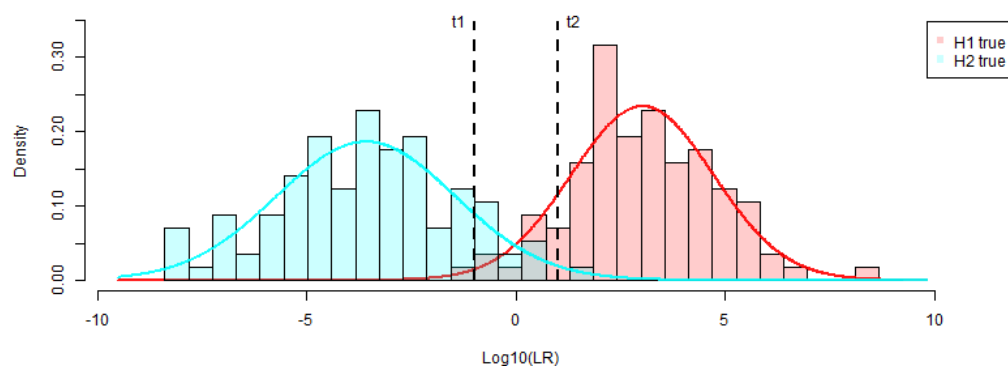


Fig. 2. An example LR distribution output.

- 3) **System power metrics** are evaluated under pre-defined thresholds. Users may drag the slider to change the thresholds and their positions (black dot lines) will also change accordingly (Fig. 2). Both single and double thresholds (default) are allowed. Metrics for system power include

sensitivity (Sen), specificity (Spe), positive predictive value (PPV), negative predictive value (NPV), false positive rate (FPR), false negative rate (FNR), inconclusive and effectiveness. They are defined as:

- Sen: proportion of pedigrees under H1 judged as H1 true;
- Spe: proportion of pedigrees under H2 judged as H2 true;
- PPV: proportion of pedigrees correctly judged as H1 true;
- NPV: proportion of pedigrees correctly judged as H2 true;
- FPR: proportion of pedigrees under H2 judged as H1 true;
- FNR: proportion of pedigrees under H1 judged as H2 true;
- Inconclusive: proportion of pedigrees that cannot be judged as either H1 true or H1 true;
- Effectiveness: proportion of pedigrees judged as H1 true or H2 true.

- 4) **System power metrics for pruned pedigrees:** As shown in Table 5, the first column displays the combinations of possible subsets, and the remaining eight columns are corresponding system power metrics, which are the same as those in 3). If H1 and H2 are equivalent, cells are filled with blank, e.g. POI+S7 and POI+S6 in Table 5.

Table 5. An example of system power metrics for pruned pedigrees. Sample labeling is consistent with that in Fig. 1.

Subset	Sen	Spe	PPV	NPV	FPR	FNR	Inconclusive	Effectiveness
POI+S4+S5+S6+S3+S7	0.9982	0.9995	>0.9999	>0.9999	<0.0001	<0.0001	0.0011	0.9989
POI+S4+S5+S6+S3	0.9854	0.9986	>0.9999	>0.9999	<0.0001	<0.0001	0.0080	0.9920
POI+S4+S5+S6+S7	0.9961	0.9993	>0.9999	>0.9999	<0.0001	<0.0001	0.0023	0.9977
POI+S4+S5+S3+S7	0.8658	0.9374	0.9998	>0.9999	0.0002	<0.0001	0.0983	0.9017
POI+S4+S6+S3+S7	0.8637	0.9281	0.9995	>0.9999	0.0005	<0.0001	0.1039	0.8961
POI+S5+S6+S3+S7	0.8704	0.9841	0.9997	>0.9999	0.0002	<0.0001	0.0726	0.9274
POI+S4+S5+S6	0.9854	0.9986	>0.9999	>0.9999	<0.0001	<0.0001	0.0080	0.9920
POI+S4+S5+S3	0.2479	0.1086	>0.9999	0.9999	<0.0001	<0.0001	0.8217	0.1783
POI+S4+S5+S7	0.7204	0.8386	0.9995	>0.9999	0.0004	<0.0001	0.2203	0.7797
POI+S4+S6+S3	0.2599	0.1038	>0.9999	0.9998	<0.0001	<0.0001	0.8181	0.1819
POI+S4+S6+S7	0.7348	0.8429	0.9991	>0.9999	0.0007	<0.0001	0.2108	0.7892
POI+S4+S3+S7	0.5629	0.2635	>0.9999	0.9998	<0.0001	0.0001	0.5868	0.4132
POI+S5+S6+S3	0.6484	0.9568	0.9989	>0.9999	0.0007	<0.0001	0.1970	0.8030
POI+S5+S6+S7	0.8238	0.9791	0.9997	>0.9999	0.0003	<0.0001	0.0984	0.9016
POI+S5+S3+S7	0.5239	0.7784	0.9987	>0.9999	0.0007	<0.0001	0.3485	0.6515
POI+S6+S3+S7	0.5329	0.7946	0.9986	>0.9999	0.0008	<0.0001	0.3358	0.6642
POI+S4+S5	0.2479	0.1086	>0.9999	0.9999	<0.0001	<0.0001	0.8217	0.1783

POI+S4+S6	0.2599	0.1038	>0.9999	0.9998	<0.0001	<0.0001	0.8181	0.1819
POI+S4+S3	-	-	-	-	-	-	-	-
POI+S4+S7	0.2692	0.0960	>0.9999	0.9997	<0.0001	<0.0001	0.8174	0.1826
POI+S5+S6	0.6484	0.9568	0.9989	>0.9999	0.0007	<0.0001	0.1970	0.8030
POI+S5+S3	0.0520	0.0185	>0.9999	0.9996	<0.0001	<0.0001	0.9647	0.0353
POI+S5+S7	0.3252	0.5963	0.9990	>0.9999	0.0003	<0.0001	0.5391	0.4609
POI+S6+S3	0.0526	0.0156	>0.9999	0.9996	<0.0001	<0.0001	0.9659	0.0341
POI+S6+S7	0.3302	0.6126	0.9983	>0.9999	0.0006	<0.0001	0.5283	0.4717
POI+S3+S7	0.2618	0.0963	>0.9999	0.9997	<0.0001	<0.0001	0.8209	0.1791
POI+S4	-	-	-	-	-	-	-	-
POI+S5	0.0520	0.0185	>0.9999	0.9996	<0.0001	<0.0001	0.9647	0.0353
POI+S6	0.0526	0.0156	>0.9999	0.9996	<0.0001	<0.0001	0.9659	0.0341
POI+S3	-	-	-	-	-	-	-	-
POI+S7	0.0557	0.0121	>0.9999	0.9993	<0.0001	<0.0001	0.9661	0.0339

- 5) An example of output for a single real case is shown in Table 6. Posterior probability is estimated under equal prior probability of H1 and H2.

Table 6. An example of output for a single real case.

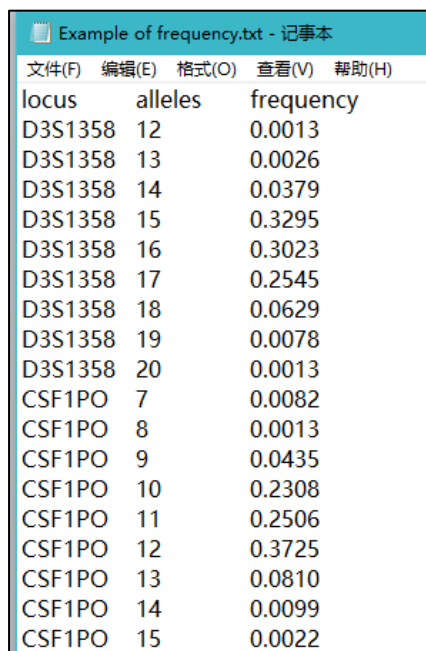
	S5	S1	S2	POI	LR
D3S1358	16,16	15,16	16,17	15,16	0.760156
CSF1PO	13,13	10,13	10,13	9,13	0.413812
D2S441	11,12	11,14	11,12	11,11	1.032526
D21S11	30,32.2	32,32.2	30,30.3	30,32	8.878741
PentaE	13,19	12,13	13,16	12,13	2.086573
D8S1179	13,15	11,15	13,13	11,13	3.228246
D5S818	11,13	11,13	11,11	11,13	1.50924
D19S433	14,16.2	14,16.2	13,14	14,14	1.779435
D16S539	9,11	9,11	9,9	9,9	2.703354
PentaD	10,13	9,13	12,13	9,10	0.659986
vWA	18,19	17,18	16,18	16,18	1.512296
D2S1338	21,23	19,21	19,21	19,21	1.837489
D18S51	15,19	13,15	14,15	14,19	1.312565
D22S1045	17,17	15,17	15,17	15,17	1.26377
TH01	6,9	9,9	6,6	9,9	1.105452
D12S391	22,24	19,24	19,24	19,24	1.84463
TPOX	8,8	8,8	8,8	8,8	1.561218
FGA	23.2,26	23.2,25	25,26	25,26	3.505724
D13S317	8,11	11,13	8,10	10,11	1.668279
D1S1656	13,17	13,15	15,17	13,15	1.514367
D10S1248	15,15	15,16	15,16	13,15	0.45706
D6S1043	11,18	13,18	11,11	11,18	1.967816

D7S820	12,12	12,12	12,12	12,12	3.124062
D6S477	12,15	14,15	12,14	12,15	frequency data not available
D18S535	9,12	12,14	9,12	9,12	frequency data not available
Combined LR (CLR)					33990.31
Posterior probability					0.999971

4. Examples

1) Allele frequency and kits/markers

- Click *Kinship* → *Settings* → *Browse* (Frequency data file) successively and upload a pre-prepared frequency data file (only two STR, D3S1358 and CSF1PO, are shown).



locus	alleles	frequency
D3S1358	12	0.0013
D3S1358	13	0.0026
D3S1358	14	0.0379
D3S1358	15	0.3295
D3S1358	16	0.3023
D3S1358	17	0.2545
D3S1358	18	0.0629
D3S1358	19	0.0078
D3S1358	20	0.0013
CSF1PO	7	0.0082
CSF1PO	8	0.0013
CSF1PO	9	0.0435
CSF1PO	10	0.2308
CSF1PO	11	0.2506
CSF1PO	12	0.3725
CSF1PO	13	0.0810
CSF1PO	14	0.0099
CSF1PO	15	0.0022

- Kits/markers:** Assuming that a subset of inputted STRs (frequency data file) will be genotyped, tick the box *Restrict markers in following selected kit(s)* (red arrow). Then, select the kit(s) to be used among the candidate kits. Here we chose two kits, AmpFI STR Identifiler and AmpFI STR GlobalFiler (green arrow). If the kit you are going to use is not presented in the candidate kits, you may input corresponding STR markers in the box *Input marker(s)* (blue arrow). Finally, we can see that 24 STR loci have been selected, i.e., CSF1PO, D10S1248, D12ATA63, D12S391, D13S317, D16S539, D18S51, D19S433, D1S1656, D21S11, D22S1045, D2S1338, D2S1776,

D2S441, D3S1358, D5S2500, D5S818, D7S820, D8S1179, FGA, SE33, TH01, TPOX, and vWA (purple arrow).

Kits and markers

Frequency data file (S2 markers are available by default)

Browse... No file selected

☒ Restrict markers in following selected kit(s)

Choose the kit(s)

AmpFISTR Identifier AmpFISTR GlobalFiler

Input marker(s)

D2S1776 D5S2500 D12AT463

You will type 24 loci, including CSF1PO, D10S1248, D12AT463, D12S391, D13S317, D16S539, D16S51, D19S433, D13S856, D21S11, D22S1045, D25S138, D251776, D25441, D3S1358, D5S2500, D5S818, D7S820, D8S1179, FGA, SE33, TH01, TPOX, vWA

Mutation

Model: Stepwise Rate1: 0.002 Rate2(non_integer): 0.00000001

Range: 0.1 Ratio(male/female): 3 Modeling STR individually: Browse... No file selected

subpopulation effects: 0 Simulations: 100

2) Standard duo paternity testing cases

Method 1: Click *Kinship* → *Pedigrees*. Then, specify sex: male, known parent: none, and reference parent(s): father (red squares). The pedigrees under H1 true and H2 true are plotted on the right. The father is not plotted under H2 true as he is a singleton.

Person of interest (POI)

sex: male known parent: none

Reference samples

parent(s): father children:

full siblings: genetically unrelated

paternal grandparents: maternal grandparents

paternal avuncular: maternal avuncular

paternal half siblings: maternal half siblings

grandchild (son): grandchild (daughter)

nephew niece (bro): nephew niece (sister)

User defined pedigrees

Browse... No file selected

Hypotheses

H1: S1 (father) and unknown mother have child POI.

H2: Unknown mother has child POI. S1 (father) is a singleton.

Likelihood ratio(LR) distribution

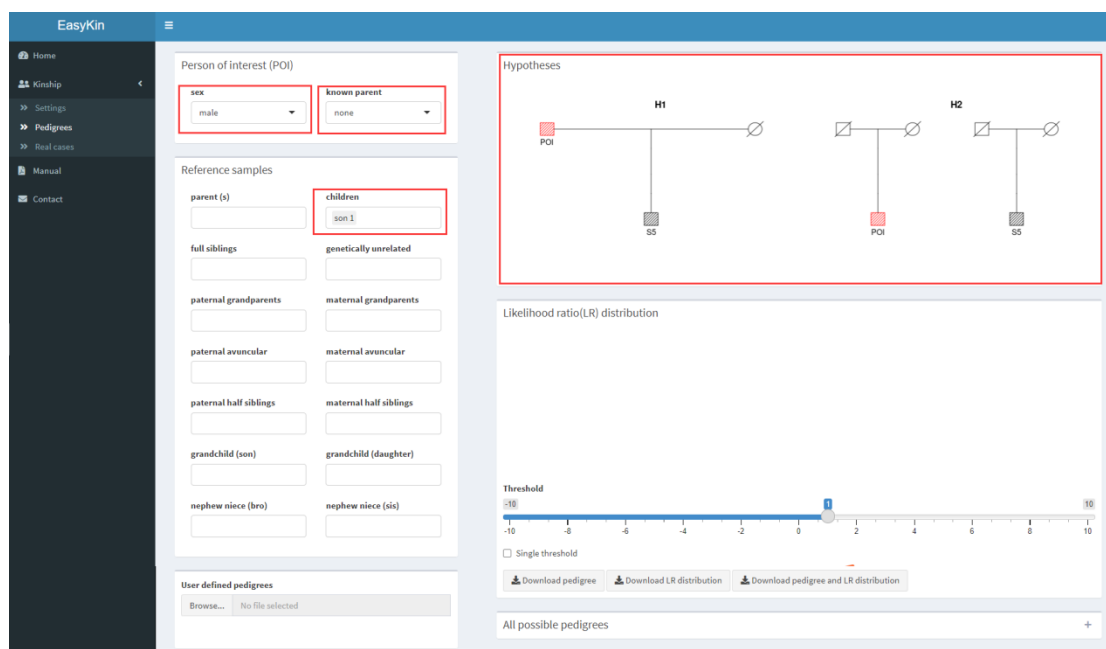
Threshold: -10 0 10

☐ Single threshold

Download pedigree Download LR distribution Download pedigree and LR distribution

All possible pedigrees +

Method 2: Click *Kinship* → *Pedigrees*. Then, specify *sex*: male, *known parent*: none and reference *children*: son 1 (red squares). The pedigrees under H1 true and H2 true are plotted on the right.



Note: Although the father is not plotted under H2 true using Method 1, it is included in the LR calculation. Therefore, both hypotheses under H2 true using Method 1 and 2 are equivalent.

3) Standard trio paternity testing cases

Method 1: Click *Kinship* → *Pedigrees*. Then, specify *sex*: male, *known parent*: mother, and reference *parent(s)*: father (red squares). The pedigrees under H1 true and H2 true are plotted on the right. The father is not plotted under H2 true as he is a singleton

EasyKin

Home
Kinship
Settings
Pedigrees
Real cases
Manual
Contact

Person of Interest (POI)

sex: male known parent: mother

Reference samples

parent (s): father children: son 1

full siblings: genetically unrelated

paternal grandparents: maternal grandparents

paternal avuncular: maternal avuncular

paternal half siblings: maternal half siblings

grandchild (son): grandchild (daughter)

nephew niece (bro): nephew niece (sis)

User defined pedigrees

Browse... No file selected

Hypotheses

H1 H2

Likelihood ratio(LR) distribution

Threshold: 10

Single threshold

Download pedigree Download LR distribution Download pedigree and LR distribution

All possible pedigrees

Method 2: Click *Kinship* → *Pedigrees*. Then, specify *sex*: male, *known parent*: none, *children*: son 1, and *genetically unrelated*: spouse (red squares). The pedigrees under H1 true and H2 true are plotted on the right.

EasyKin

Home
Kinship
Settings
Pedigrees
Real cases
Manual
Contact

Person of Interest (POI)

sex: male known parent: none

Reference samples

parent (s): children: son 1

full siblings: genetically unrelated

paternal grandparents: maternal grandparents

paternal avuncular: maternal avuncular

paternal half siblings: maternal half siblings

grandchild (son): grandchild (daughter)

nephew niece (bro): nephew niece (sis)

User defined pedigrees

Browse... No file selected

Hypotheses

H1 H2

Likelihood ratio(LR) distribution

Threshold: 10

Single threshold

Download pedigree Download LR distribution Download pedigree and LR distribution

All possible pedigrees

4) Both parents are suspected

Click *Kinship* → *Pedigrees*. Then, specify *sex*: male, *known parent*: none, and parent(s): father and mother (red squares). The pedigrees under H1 true and H2 true are plotted on the right. The father and

mother are not plotted under H2 as they are both singleton.

The screenshot shows the EasyKin web application interface. On the left is a navigation menu with options: Home, Kinship, Settings, Pedigrees, Real Cases, Manual, and Contact. The main area is divided into several sections:

- Person of interest (POI):** Includes dropdowns for 'sex' (set to 'male') and 'known parent' (set to 'none').
- Reference samples:** A grid of input fields for various relationships:
 - parent (s): father, mother
 - children
 - full siblings
 - genetically unrelated
 - paternal grandparents
 - maternal grandparents
 - paternal avuncular
 - maternal avuncular
 - paternal half siblings
 - maternal half siblings
 - grandchild (son)
 - grandchild (daughter)
 - nephew niece (bro)
 - nephew niece (sis)
- User defined pedigrees:** A section with a 'Browse...' button and a note 'No file selected'.
- Hypotheses:** Two pedigree charts labeled H1 and H2. H1 shows a male (S1) and a female (S3) as parents of a male (POI). H2 shows a male and a female as parents of a male (POI).
- Likelihood ratio (LR) distribution:** A horizontal axis ranging from -10 to 10, with a threshold line at 0. Below the axis are buttons for 'Download pedigree', 'Download LR distribution', and 'Download pedigree and LR distribution'.
- All possible pedigrees:** A section with a '+' button.

5) Multi-individual testing

Assume that a boy (POI) claims to be the child of a deceased man, whose DNA is not available anymore. The genetic data of the two putative paternal half-brothers (as full siblings) and their mother, can be available. Click *Kinship* → *Pedigrees*. Then, specify *sex*: male, *known parent*: none, *paternal half siblings*: brother 1 and brother 2, and *genetically unrelated*: the mother of paternal half sibling (red squares).

This screenshot shows the EasyKin web application interface configured for multi-individual testing. The 'Person of interest (POI)' section has 'sex' set to 'male' and 'known parent' set to 'none'. The 'Reference samples' section has 'paternal half siblings' set to 'brother 1' and 'brother 2', and 'genetically unrelated' set to 'mother of paternal half sibling'. The 'Hypotheses' section shows two pedigree charts, H1 and H2, illustrating the relationships between the POI and the reference samples. H1 shows a male (POI) and a female (S6) as parents of a male (S7). H2 shows a male (POI) and a female (S6) as parents of a male (S7). The 'Likelihood ratio (LR) distribution' section shows a horizontal axis ranging from -10 to 10, with a threshold line at 0. Below the axis are buttons for 'Download pedigree', 'Download LR distribution', and 'Download pedigree and LR distribution'. The 'All possible pedigrees' section has a '+' button.

6) User defined pedigrees

Click *Kinship* → *Pedigrees* → *Browse* (User defined pedigrees) and upload following text file (Example of ped.txt). The pedigrees under H1 true and H2 true are shown below.

Example of ped.txt - 记事本

#hypothesis	individual	father	mother	sex	status
H1	POI	fa	mo	2	1
H1	fa	gf	gm	1	0
H1	mo	NA	NA	2	0
H1	gf	NA	NA	1	0
H1	gm	NA	NA	2	0
H1	uncle	gf	gm	1	0
H1	aunt	NA	NA	2	0
H1	fc	uncle	aunt	1	1
H2	POI	random1	random2	2	1
H2	random1	NA	NA	1	0
H2	random2	NA	NA	2	0
H2	uncle	NA	NA	1	0
H2	aunt	NA	NA	2	0
H2	fc	uncle	aunt	1	1

第 3 行, 第 6 列 100% Windows (CRLF) UTF-8

EasyKin

Person of interest (POI)

sex: male known parent: none

Reference samples

parent (s): children:

full siblings: genetically unrelated:

paternal grandparents: maternal grandparents:

paternal avuncular: maternal avuncular:

paternal half siblings: maternal half siblings:

grandchild (son): grandchild (daughter):

nephew niece (bro): nephew niece (sis):

User defined pedigrees

Browse... Upload complete

Hypotheses

H1: Pedigree diagram showing a family with a POI (red circle) and a full sibling (fc, grey square).

H2: Pedigree diagram showing a family with a POI (red circle) and a full sibling (fc, grey square).

Likelihood ratio (LR) distribution

Threshold: -10 to 10

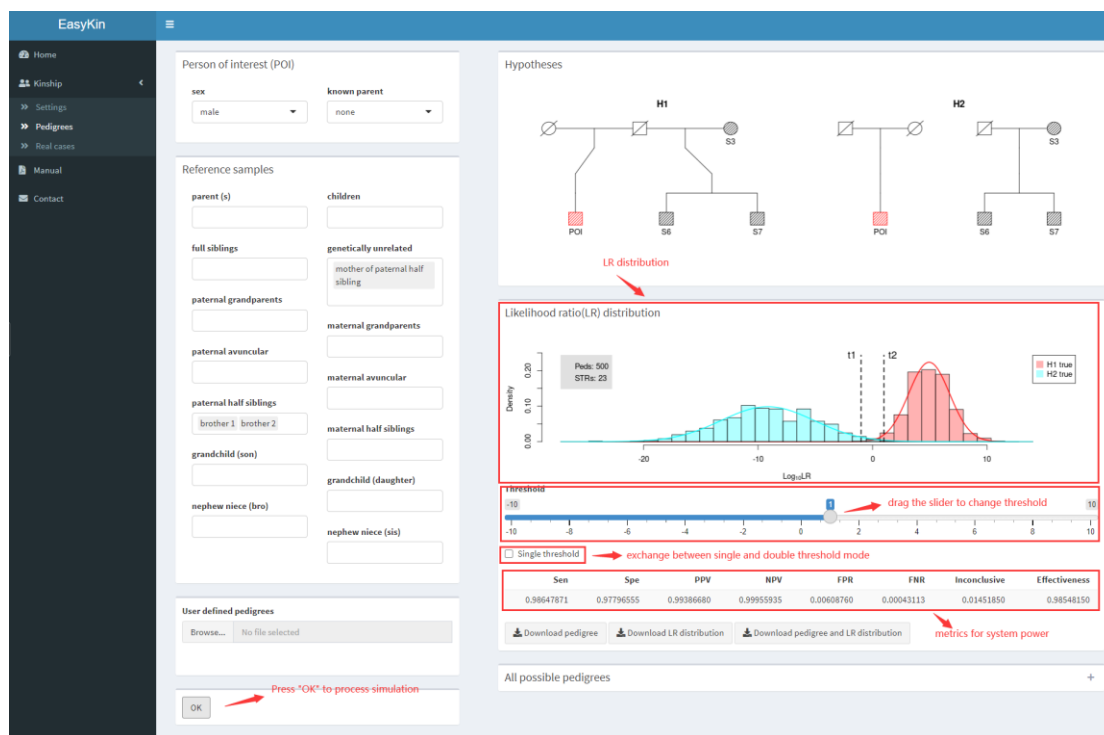
Single threshold: ☐

Download pedigree Download LR distribution Download pedigree and LR distribution

All possible pedigrees

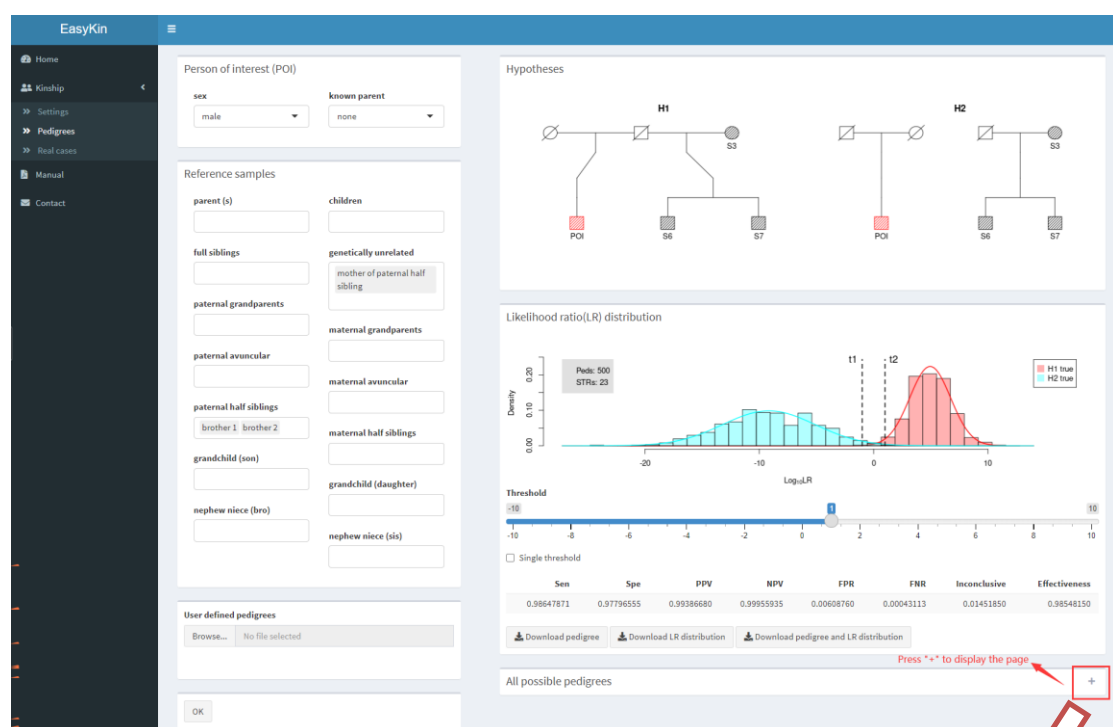
7) LR distribution and system power

Take the case in 5) as an example. After reference selection and pedigree construction, press the button “OK” (left bottom) and then simulation will be processed. LR distribution and metrics for system power are shown on the right bottom. Users may drag the slider to change the threshold and metrics for system power will be updated immediately.




8) Pedigree pruning


Take the case in 5) as an example. After reference selection and pedigree construction, press the "+" to display the page (right bottom). Then, click the button *prune pedigree* to process the pruning. Results are shown below.



All possible pedigrees

Prune pedigree  Press to prune the pedigree

This process will cost much more time than single pedigree.

CSV  Result

Search:

	Sen	Spe	PPV	NPV	FPR	FNR	Inconclusive	Effectiveness
POI+S3+S6+S7	0.986479	0.977966	0.993867	0.999559	0.006088	0.000431	0.014519	0.985481
POI+S3+S6	0.876004	0.8705	0.995986	0.988088	0.003531	0.010494	0.119735	0.880265
POI+S3+S7	0.870292	0.861979	0.995085	0.988205	0.004298	0.010288	0.126571	0.873429
POI+S6+S7	0.902089	0.897052	0.982149	0.995326	0.016396	0.004212	0.090125	0.909875
POI+S3	H1 and H2 are equivalent							
POI+S6	0.752624	0.722805	0.987113	0.981186	0.009826	0.01386	0.250443	0.749557
POI+S7	0.745145	0.716247	0.985693	0.982539	0.010815	0.012728	0.257532	0.742468

Showing 1 to 7 of 7 entries

9) LR calculation for single case

Again, take the case in 5) as an example. Click *Kinship* → *Real cases* → *Browse* (Input genotype data) and upload following text file (hb.txt). Note that pedigrees in this page are updated automatically. Press “calculate LRs” and then press “LRs” to display the result, i.e., LRs at each locus, combined LR (CLR), a verbal equivalent, and posterior probability. Press “pairwise LRs” to display LRs for all pairs of references.

hb.txt - 记事本

sample	S3	S6	S7	POI
D2S441	11,12	11,14	11,12	11,11
TPOX	8,8	8,8	8,8	8,8
D22S1045	17,17	15,17	15,17	15,17
D7S820	12,12	12,12	12,12	12,12
D1S1656	13,17	13,15	15,17	13,15
PentaE	13,19	12,13	13,16	12,13
D10S1248	15,15	15,16	15,16	13,15
D8S1179	13,15	11,15	13,13	11,13
D5S818	11,13	11,13	11,11	11,13
D19S433	14,16.2	14,16.2	13,14	14,14
D16S539	9,11	9,11	9,9	9,9
CSF1PO	13,13	10,13	10,13	9,13
PentaD	10,13	9,13	12,13	9,10
D3S1358	16,16	15,16	16,17	15,16
vWA	18,19	17,18	16,18	16,18
D2S1338	21,23	19,21	19,21	19,21
D18S51	15,19	13,15	14,15	14,19
D6S1043	11,18	13,18	11,11	11,18
D13S317	8,11	11,13	8,10	10,11

EasyKin

Home

Kinship

Settings

Progress

Real cases

Manual

Contact

Input genotype data

Browse...

h1.txt

Upload complete

Frequency of novel alleles

0.01

Prior probability

0.5

old labels

POI

new labels

S3

Calculate LRs

Clear labels

Assign labels

updated automatically

Genotype

LRs

Pairwise LRs

LRs for all pairs of references

genotypes and LRs

Search:

	S3	S6	S7	POI	LR
D2S441	11,12	11,14	11,12	11,11	1.034214
TPOX	8,8	8,8	8,8	8,8	1.561304
D22S1045	17,17	15,17	15,17	15,17	1.264008
D7S820	12,12	12,12	12,12	12,12	3.123945
D1S1656	13,17	13,15	15,17	13,15	1.51452
PentaE	13,19	12,13	13,16	12,13	2.084603
D10S1248	15,15	15,16	15,16	13,15	0.455806
D8S1179	13,15	11,15	13,13	11,13	3.23023

Showing 1 to 8 of 15 entries

Support

H1 true

Weight of evidence

extremely strong

Combined likelihood ratio (CLR)

33811.576472416

log10(CLR)

4.52906542031649

Posterior probability

0.99997042520552

18