

A user-friendly GUI for the use of dnamatch2 v2.2.0

Tutorial to get started!

By Øyvind Bleka at Oslo University Hospital

What's new from v2.0?

- Choose between search strategies:

1. Simple allele comparison
2. + qualitative LR
3. ++ quantitative LR



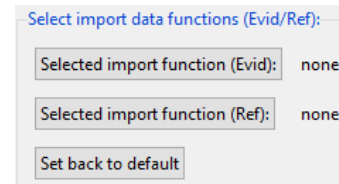
Search strategy: MAC +Qual ++Quan

- Multi-kit compatibility

- Search between samples processed with different kits
 - Both evidence and reference profiles
 - Uses overlapping markers

- Added flexibility

- Customize your own import data function
- Optional to search within sub-folders



Select import data functions (Evid/Ref):

Selected import function (Evid):	none
Selected import function (Ref):	none
<input type="button" value="Set back to default"/>	

The GUI

- Set*/save/load search settings (management).
- Performs a dnamatch2 search based on settings.
- Produces matchfiles from search.

* Settings are remembered after quitting GUI

Installation

Go to <http://euroformix.com/dnamatch2> and follow instructions.

Remember to install the R-packages:

- euroformix (works well with version 3)
- forensim
- gWidgetstcltk
- igraph

How to use

1) Open R where dnamatch2 package was installed.

2) Open GUI in R with:

```
library(dnamatch2);gui()
```

The GUI

Contains 4 sections:

1) Analyse

- Executes search

2) Data setup

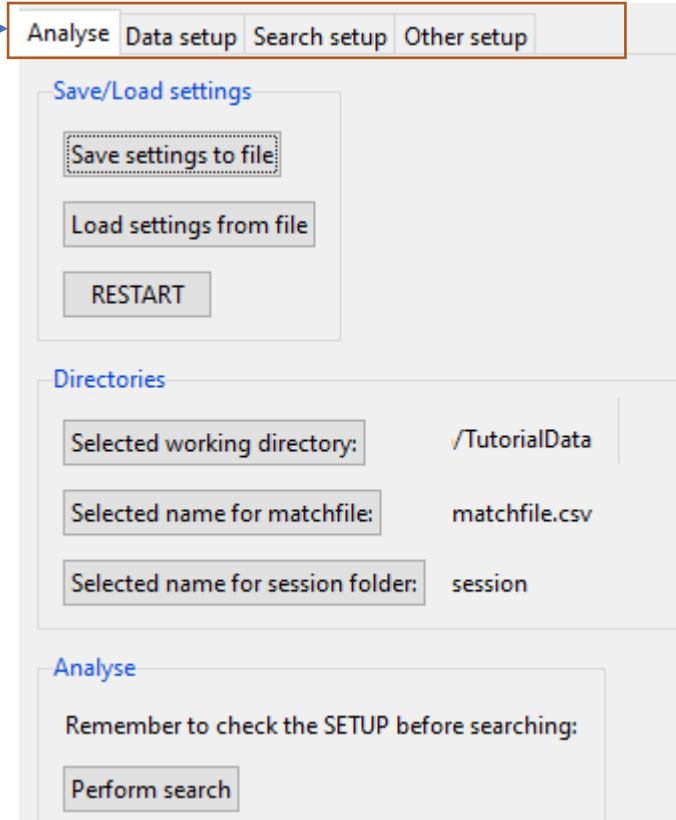
- Select data in search

3) Search setup

- Settings in search

4) Other setup

- Pattern settings for recognizing different ID formats/types (advanced)
- Select your own Import function



Section 1: Analyse

Working directory:

- Search Results will be stored here!

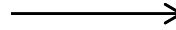
Matchfile:

- Includes full list of match candidates
- The file will be extended with new matches

Session folder:

- Includes detailed search logs

Saving/loading
setting file



Executes the search



Analyse Data setup Search setup Other setup

Save/Load settings

Save settings to file

Load settings from file

RESTART

Directories

Selected working directory: /TutorialData

Selected name for matchfile: matchfile.csv

Selected name for session folder: session

Analyse


Remember to check the SETUP before searching:


Perform search


NB: Remember to save changed settings first in Data setup


Section 2: Data setup

Select population frequency file (mandatory) 

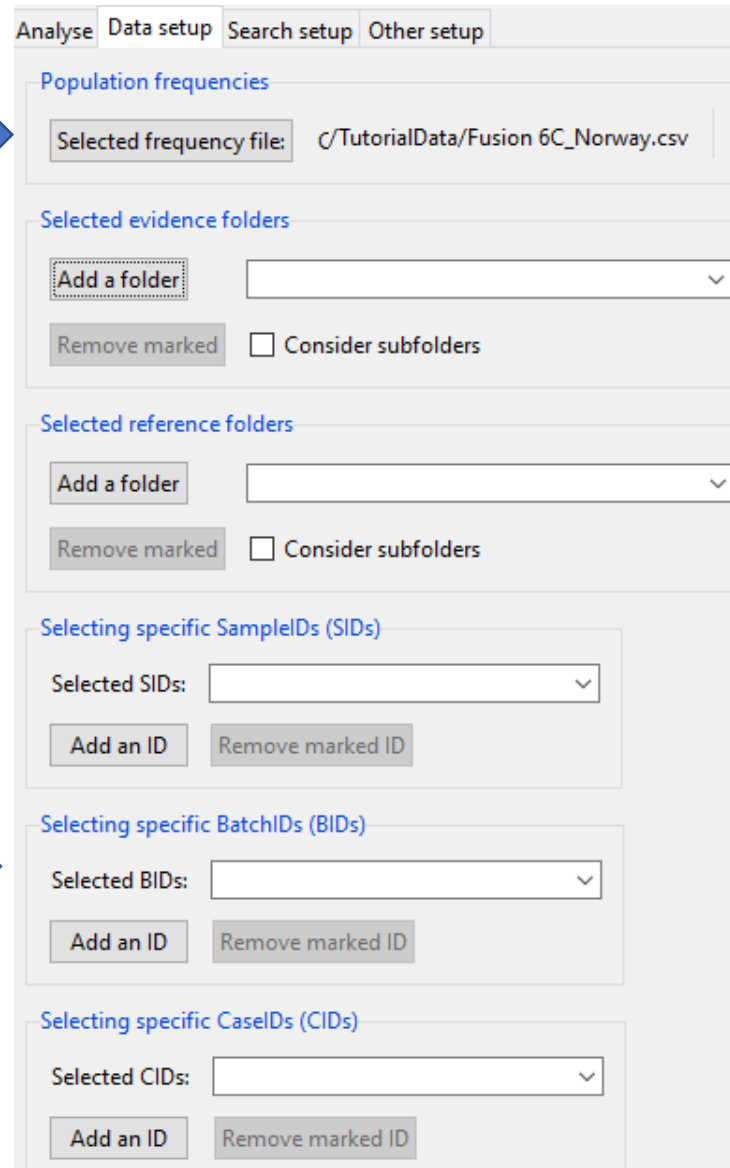
One or more folders including files with evidence profiles (mandatory). Searching subfolders is optional. 

Folder(s) including files with reference profiles (optional). Searching subfolders is optional. 

Specific Samples to search (optional) 

Specific Batch(es) to search (optional) (must be name of evidence profile files) 

Specific Case(s) to search (optional) 



The screenshot shows the 'Data setup' tab of a software interface. It contains several sections for configuring search parameters:

- Population frequencies:** A text input field labeled 'Selected frequency file:' with the value 'C:/TutorialData/Fusion 6C_Norway.csv'.
- Selected evidence folders:** A section with an 'Add a folder' button, a dropdown menu, a 'Remove marked' button, and a checkbox labeled 'Consider subfolders'.
- Selected reference folders:** A section with an 'Add a folder' button, a dropdown menu, a 'Remove marked' button, and a checkbox labeled 'Consider subfolders'.
- Selecting specific SampleIDs (SIDs):** A section with a 'Selected SIDs:' dropdown menu, an 'Add an ID' button, and a 'Remove marked ID' button.
- Selecting specific BatchIDs (BIDs):** A section with a 'Selected BIDs:' dropdown menu, an 'Add an ID' button, and a 'Remove marked ID' button.
- Selecting specific CaseIDs (CIDs):** A section with a 'Selected CIDs:' dropdown menu, an 'Add an ID' button, and a 'Remove marked ID' button.

Section 3: Search setup

Possible to change search time (set when GUI opened)

Selecting kit will activate the degradation model

Dropin-model

Thresholds used to filter data (remove alleles and ignore markers)

Majors are extracted for between stain search

Don't forget to save the settings!

Analyse Data setup Search setup Other setup

Score thresholds

Matching allele counting (MAC):

Qualitative LR:

Quantitative LR:

Search options

Search within same cases (CID): NO YES

Search between stains: NO YES

Search strategy: MAC +Qual ++Quan

Model setup

Set kit:

Drop-in prob=:

Lambda param=:

Min Freq=:

Time windows

Number of days back (days):

Time difference between matches (days):

Search time (YY-MM-DD-HH-MM-SS):

Prefilter thresholds

Analytical threshold (AT):

Stutter rate threshold:

Major extraction rate threshold:

Minimum loci requirement (Evid):

Minimum loci requirement (Maj):

Other options

Plot score histogram in search: NO YES

Write detailed score info to file: NO YES

Print graph of matches: NO YES

Ignore empty markers: NO YES

Max contr.num. (QUAL):

Max contr.num. (QUAN):

Req.optim.num. (QUAN):

EFM option: number of required MLE optimizations

Used when multiple k...

Section 4: Other setup

Symbol used to recognize following (separator sign)

- SID = SampleID
- RID = ReplicateID
- CID = CaseID

Format of SampleNames = SID_RID_CID
(mandatory)

Used to restrict on specific patterns in SID
Useful for filtering specific sample types (optional)

Used to restrict on specific patterns in Batch files (BID)
Useful for filtering specific batch types (optional)

Select a file with a R-function for reading data
which returns a table in EuroForMix format
- Can be separate for Evid and Ref profiles


The screenshot shows the 'Other setup' tab in a software interface. It contains several configuration sections:

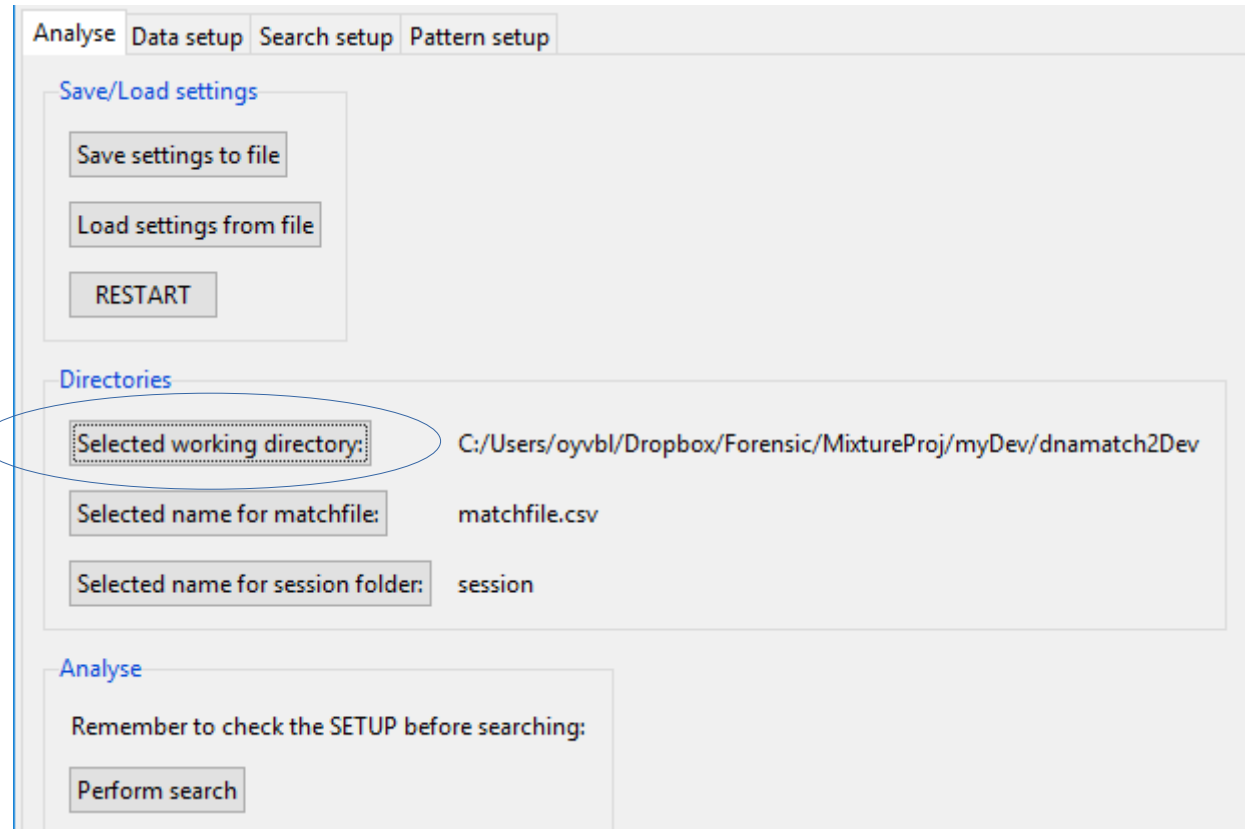
- Patterns (prefix) of IDs:** A section with a label 'Set pattern for Separating IDs:' and a text input field containing a hyphen (-).
- Set pattern for SampleIDs (SIDs):** A section with a label 'Required pattern(s):' and a dropdown menu. Below it are 'Add a pattern' and 'Remove a pattern' buttons.
- Set pattern for Batch files (BIDs):** A section with a label 'Required pattern(s):' and a dropdown menu. Below it are 'Add a pattern' and 'Remove a pattern' buttons.
- Select import data functions (Evid/Ref):** A section with two rows: 'Selected import function (Evid): none' and 'Selected import function (Ref): none'. Below these is a 'Set back to default' button.

Arrows from the text on the left point to the 'Set pattern for Separating IDs' field, the 'Required pattern(s)' dropdown for SIDs, the 'Required pattern(s)' dropdown for BIDs, and the 'Selected import function (Evid)' field.

Tutorial using a Fusion 6C example

Setup (page 1)

Select workdirectory: 
*The match and session-info will
be created under this folder.*



Analyse Data setup Search setup Pattern setup

Save/Load settings

Save settings to file

Load settings from file

RESTART

Directories

Selected working directory: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev

Selected name for matchfile: matchfile.csv


Selected name for session folder: session

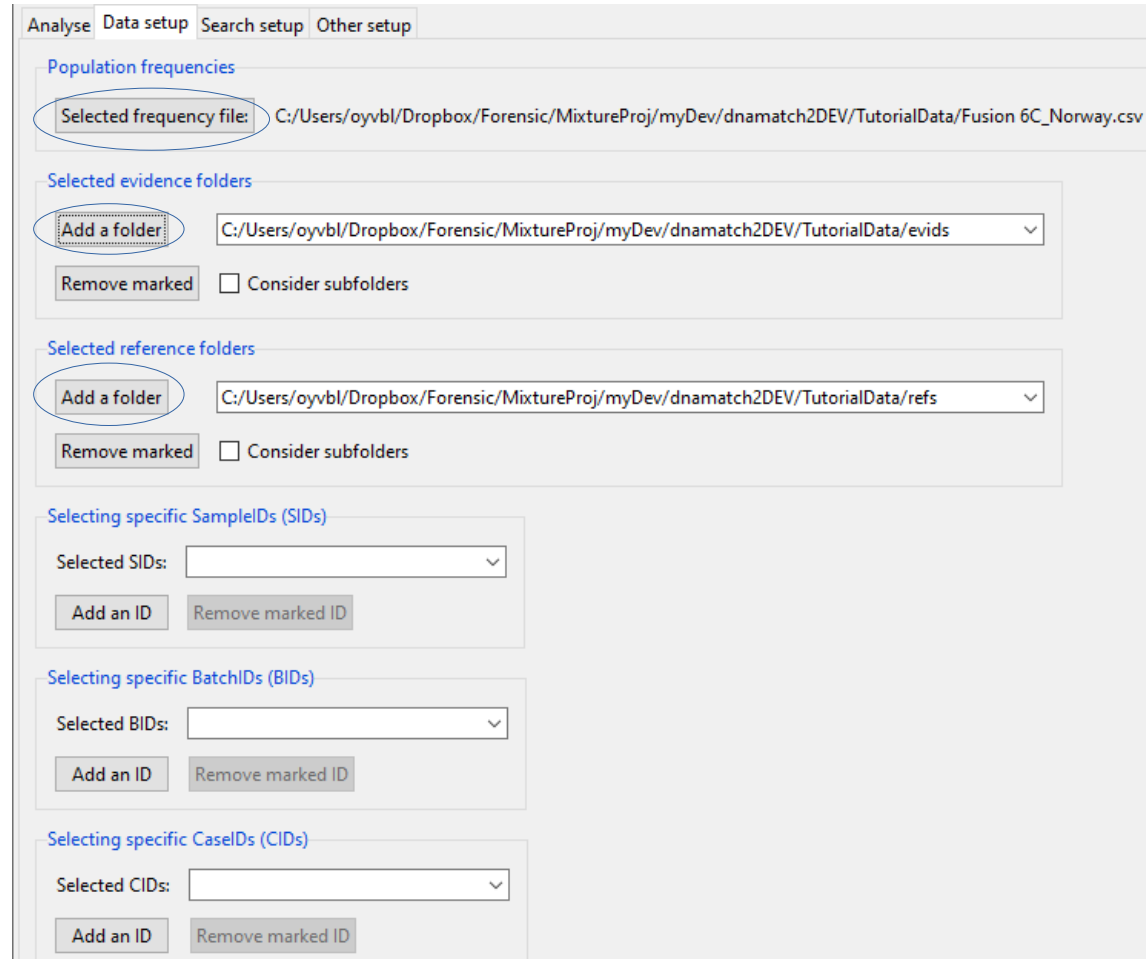
Analyse

Remember to check the SETUP before searching:

Perform search

Setup (page 2)

Select workdirectory:  *The match and session-info will be created under this folder.*



Analysse Data setup Search setup Other setup

Population frequencies

Selected frequency file: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2DEV/TutorialData/Fusion 6C_Norway.csv

Selected evidence folders

Add a folder C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2DEV/TutorialData/evids

Remove marked Consider subfolders

Selected reference folders

Add a folder C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2DEV/TutorialData/refs

Remove marked Consider subfolders

Selecting specific SampleIDs (SIDs)

Selected SIDs:

Add an ID Remove marked ID

Selecting specific BatchIDs (BIDs)

Selected BIDs:

Add an ID Remove marked ID

Selecting specific CaselDs (CIDs)

Selected CIDs:

Add an ID Remove marked ID

Setup (page 3)

Analyse Data setup Search setup Other setup

Score thresholds

Matching allele counting (MAC):

Qualitative LR:

Quantitative LR:

Search options

Search within same cases (CID): NO YES

Search between stains: NO YES

Search strategy: MAC +Qual ++Quan

Model setup

Set kit:

Drop-in prob=

Lambda param=

Min Freq=

Time windows

Number of days back (days):

Time difference between matches (days):

Search time (YY-MM-DD-HH-MM-SS):

Other options

Plot score histogram in search NO YES

Write detailed score info to file NO YES

Print graph of matches: NO YES

Ignore empty markers: NO YES

Max contr.num. (QUAL):

Max contr.num. (QUAN):

Req.optim.num. (QUAN):

Prefilter thresholds

Analytical threshold (AT)

Stutter rate threshold

Major extraction rate threshold

Minimum loci requirement (Evid)

Minimum loci requirement (Maj)

Setup (page 4)

	Sample Name	Marker	Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6	Allele 7	Allele 8
1	S-001_R-1_C-01	D3S1358	15	17	NA	NA	NA	NA	NA	NA
2	S-001_R-1_C-01	D1S1656	14	19.3	NA	NA	NA	NA	NA	NA
3	S-001_R-1_C-01	D2S441	11	14	NA	NA	NA	NA	NA	NA
4	S-001_R-1_C-01	D10S1248	13	14	NA	NA	NA	NA	NA	NA
5	S-001_R-1_C-01	D13S317	11	13	NA	NA	NA	NA	NA	NA
6	S-001_R-1_C-01	PENTA E	11	14	NA	NA	NA	NA	NA	NA
7	S-001_R-1_C-01	D16S539	12	9	NA	NA	NA	NA	NA	NA
8	S-001_R-1_C-01	D16S539	12	9	NA	NA	NA	NA	NA	NA

Use default values

We could have used SIDpattern = "S-"

BIDpattern = "Batch"

Analyse Data setup Search setup **Pattern setup**

Patterns (prefix) of IDs

Set pattern for Separating IDs:

Set pattern for SampleIDs (SIDs):

Required pattern(s):

Add a pattern Remove a pattern

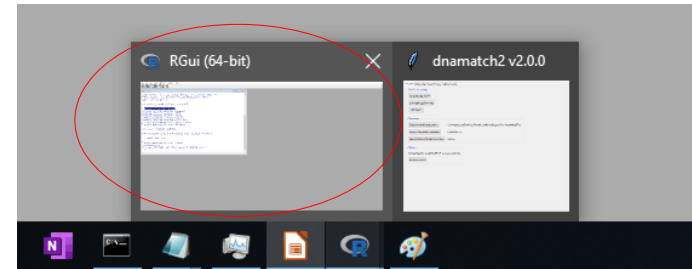
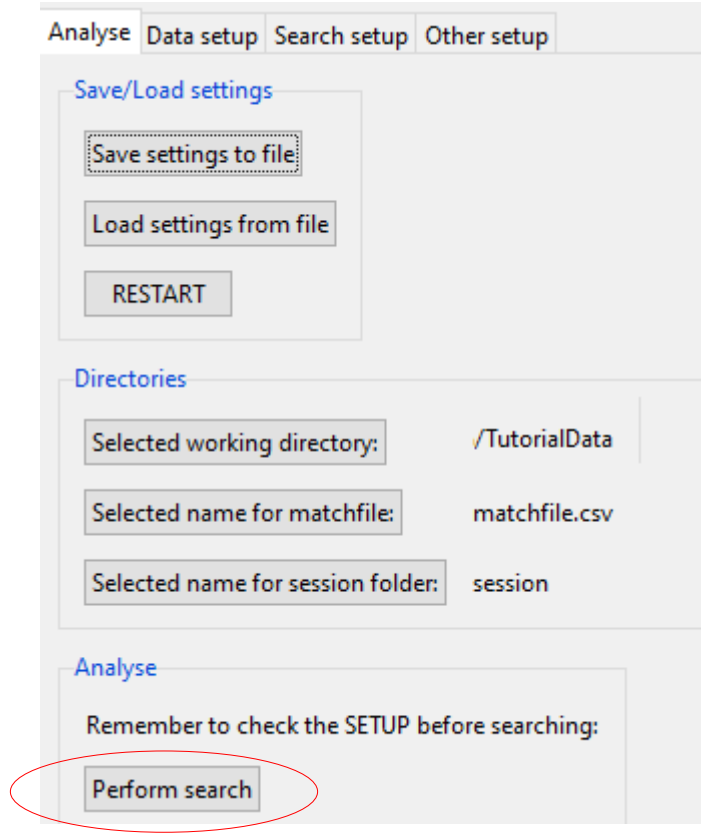
Set pattern for Batch files (BIDs):

Required pattern(s):

Add a pattern Remove a pattern

Notice the underscore pattern

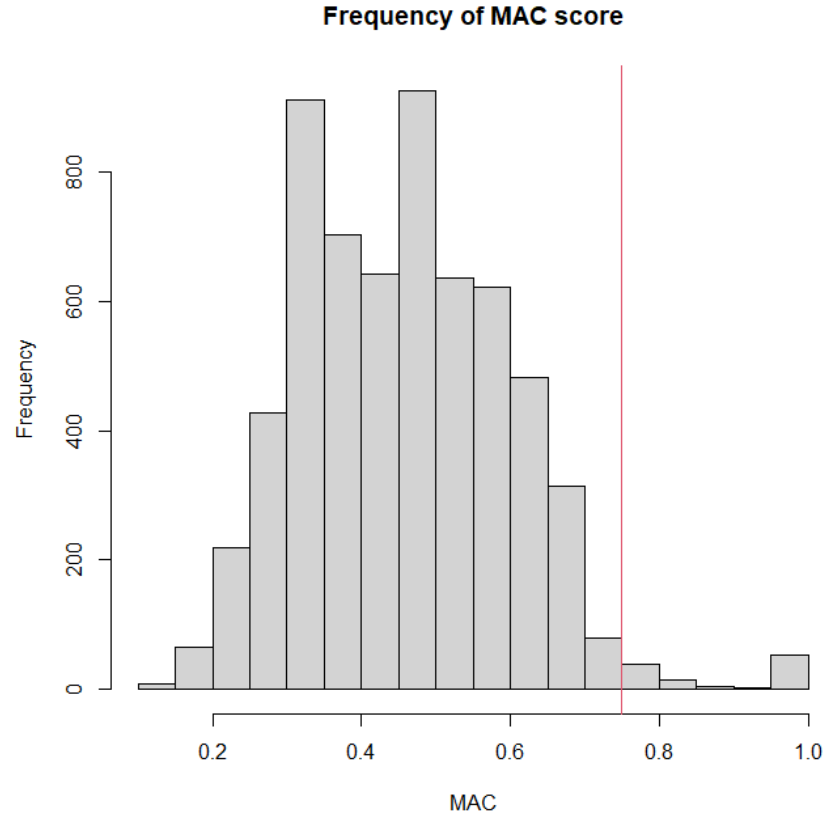
Time to perform search!



Open Rgui and look on the R-console for progress information!

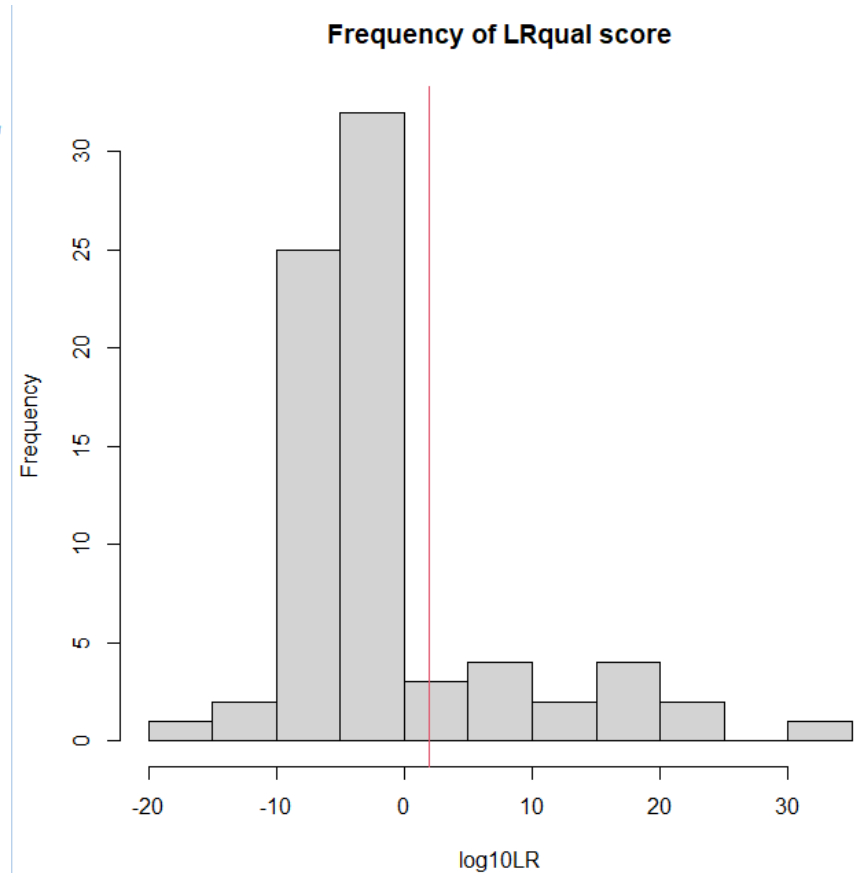
Temporary results: MAC

```
[1] "Number of imported references: 100"  
[1] "Number of loci to use: 23"  
[1] "Number of samples to search: 43"  
[1] "Number of reference profiles to search: 143"  
[1] "Calculating MAC for all 6149 comparisons: All refs against all stains"  
[1] "Calculating MAC scores took 1 seconds"  
[1] "Number of comparisons satisfying (after filters) threshMAC=0.75: 76"
```



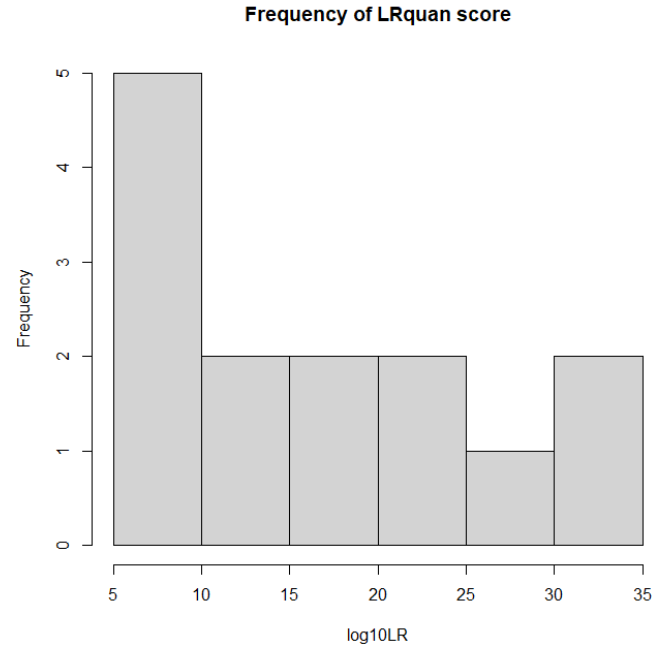
Temporary results: LRqual

```
"Estimating num. contr. for 20 stains"  
"Calculating LRqual for 76 combinations"  
"50% LR qual calculation complete"  
"100% LR qual calculation complete"  
"Calculating LR (qual) took 25 seconds"  
"Number of comparisons satisfying threshLRqual>100: 14"
```



Final results: LRquan

```
"Calculating LRquan for 14 combinations (9 unique samples)"  
"56% LR quan calculation complete"  
"Calculating LR (quan) took 52 seconds"  
"Number of comparisons satisfying threshLRquan>1000: 14"  
"Search completed! Storing results..."  
"Number of matches=12"
```



Match graph

True matchlist:

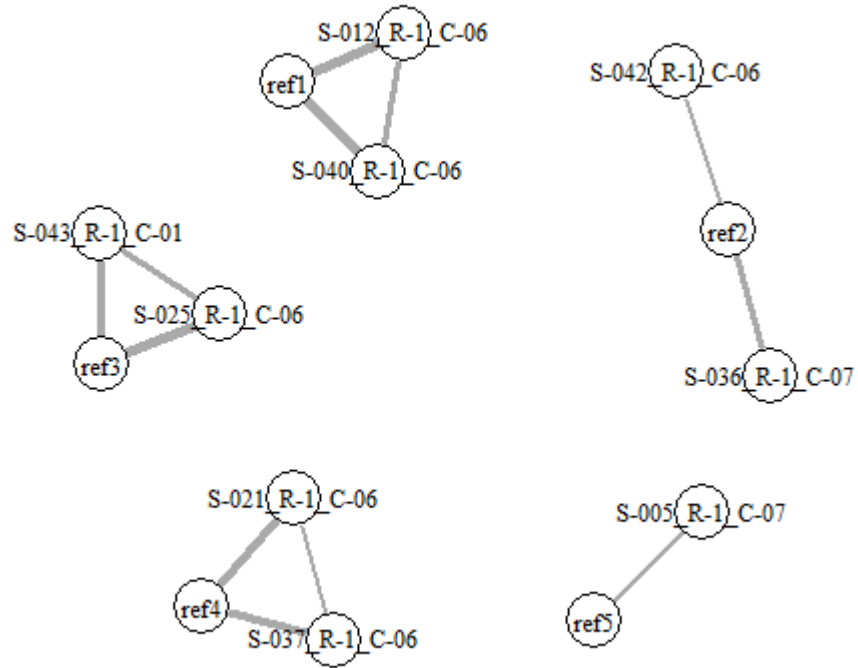
Ref1: S40,S12

Ref2: S42,S36

Ref3: S43,S25,S3 (missing)

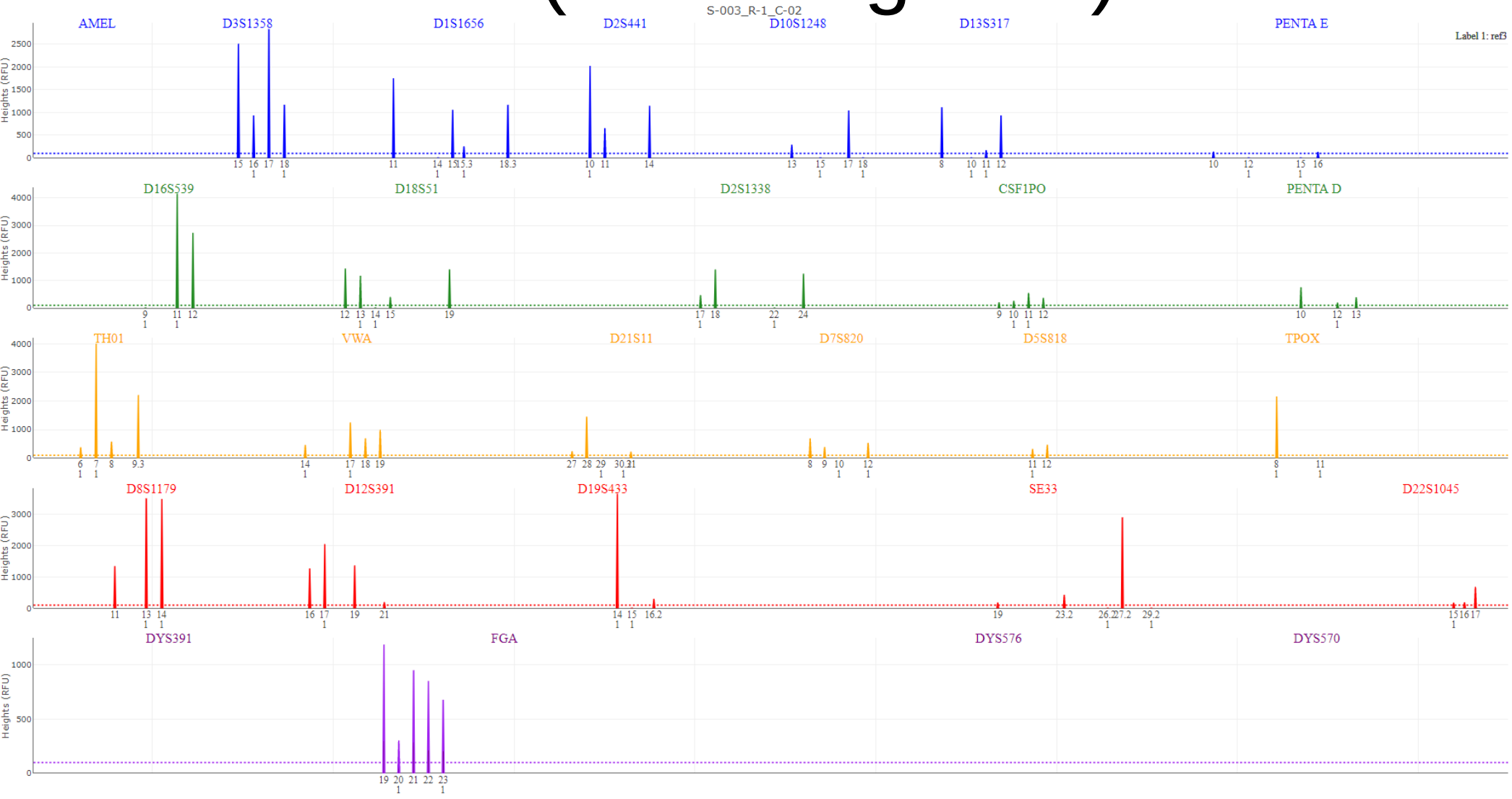
Ref4: S37,S21

Ref5: S5

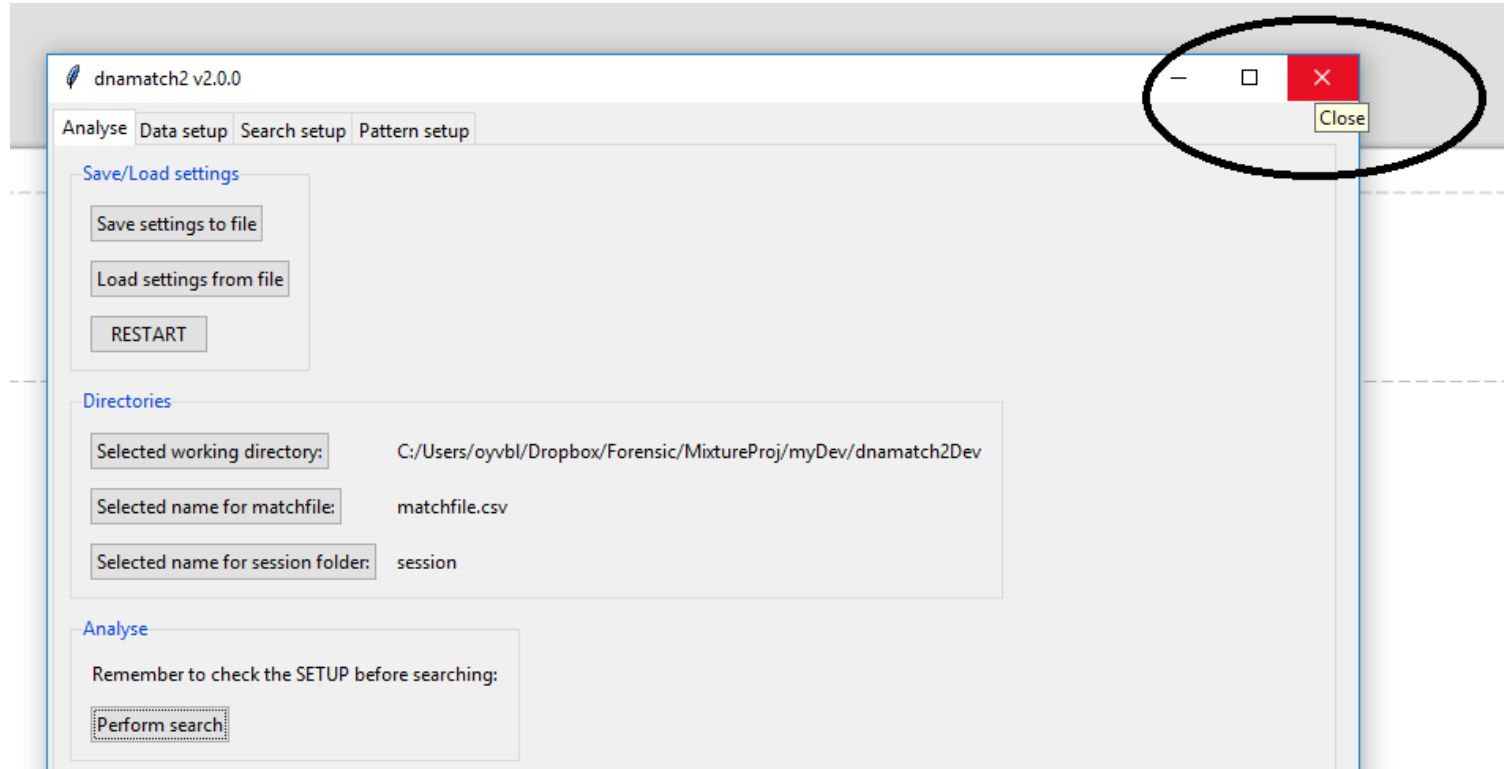


S3 (false negative)

16 dropouts!









Close GUI after use



Result and Log files

In work directory: “matchfile.csv“
(includes details about match candidates)

In session folder: Log files given by the time stamp

 LRqualResults_19-07-24-13-29-45.csv	Overview of all LRs in QUAN comparison
 LRquanResults_19-07-24-13-29-45.csv	Overview of all LRs in QUAN comparison
 matchinfo_19-07-24-13-29-45.csv	Profiles details for matches
 refmatches_19-07-24-13-29-45.csv	List of ref-evid matches
 stainmatches_19-07-24-13-29-45.csv	List of evid-evid matches
 searchLog_19-09-03-16-56-00.csv	Log with argument given to search

Matchinfo

Assumed number
of contributors
used for LR

Mixture proportion of POI

```
matchinfo_19-09-03-16-56-00.csv
1  "-----1-----" "-----1-----"
2  "ref1" "S-040_R-1_C-06"
3  "nLocs=22" "MAC=0.977 - MM=1"
4  "nContr=1" "LRqual=1.1762e+30"
5  "Mx=1" "LRquan=6.1453e+30"
6  "----REFERENCE----" "-----TARGET-----"
7  "D3S1358: 16/17" "16/17 - 829/963"
8  "D1S1656: 15.3/17" "15.3/17 - 1037/826"
9  "D2S441: 11/14" "11/14 - 598/376"
10 "D10S1248: 14/14" "14 - 961"
11 "D13S317: 13/13" "13 - 651"
12 "PENTA E: 12/15" "12/15 - 313/144"
13 "D16S539: 9/11" "11/9 - 1071/2167"
14 "D18S51: 14/19" "14/19 - 1146/565"
15 "D2S1338: 20/26" "20/26 - 257/473"
16 "CSF1PO: 12/13" "12/13 - 887/225"
17 "PENTA D: 9/10" "10/9 - 556/137"
18 "TH01: 6/9.3" "6/9.3 - 2111/2753"
19 "VWA: 14/15" "14/15 - 713/1238"
20 "D21S11: 28/29" "28/29 - 512/1149"
21 "D7S820: 9/11" "11/9 - 358/185"
22 "D5S818: 9/12" "12/9 - 370/762"
23 "TPOX: 8/11" "11/8 - 380/781"
24 "D8S1179: 13/13" "13 - 2557"
25 "D12S391: 18/18" "18 - 2615"
26 "D19S433: 13/16" "13/16 - 589/678"
27 "SE33: 21/27.2" "21 - 644"
28 "D22S1045: 15/16" "NA - NA"
29 "FGA: 22.2/25" "22.2/25 - 933/709"
```

Number of allele
mismatches

Peak heights


```
1 THIS IS A LOG FOR A dnamatch2 RUN
2 dnamatch2 version: 2.0.0
3 Other packages: (euroformix_2.2.1, forensim_4.3)
4 R-version used: R version 3.5.1 (2018-07-02)
5 User: oyvbl
6 Created: 2019-09-03 16:56:25
7
8 -----FUNCTION CALL-----
9 evidfold: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/evids
10 freqfile: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/Fusion 6C_Norway.csv
11 reffold: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/refs
12 sameCID: TRUE
13 betweensamples: TRUE
14 Thist: Inf
15 threshMAC: 0.75
16 threshLR: 100/1000
17 threshHeight: 100
18 threshStutt: 0.1
19 threshMaj: 0.6
20 minLocStain: 3
21 minLocMaj: 3
22 pC: 0.05
23 lambda: 0.01
24 kit: Fusion 6C
25 minFreq: 0.001
26 searchtime: 2019-09-03 16:56:00
27 SIDvec:
28 BIDvec:
29 CIDvec:
30 timediff: Inf
31 IDsep: _
32 BIDptrn:
33 SIDptrn:
34 printHistPlots: TRUE
35 writeScores: TRUE
36 maxK: 4/3
37 matchfile: matchfile.csv
38 sessionfold: session
```