A user-friendly GUI for the use of dnamatch2

v2.0.0

Tutorial to get started!

By Øyvind Bleka at Oslo University Hospital

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 (at least 2.0.2)
  - 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

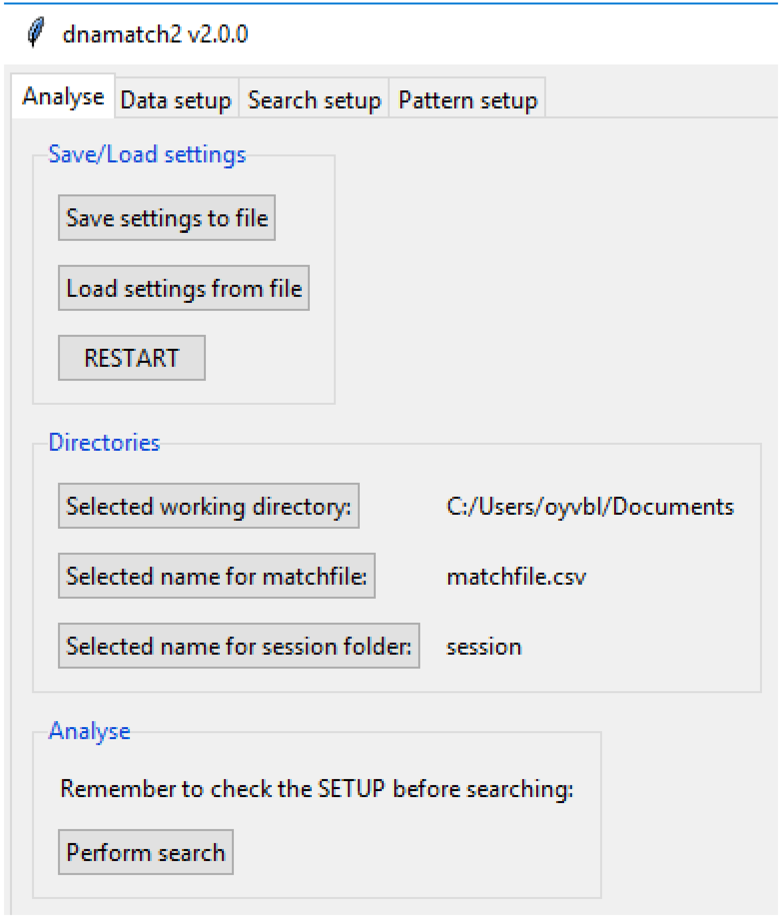
Contains 4 sections:

1) Analyse
 - Executes search

2) Data setup
- Select data in search

3) Search setup
- Settings in search

4) Pattern setup
- Settings for recognizing different ID formats/types (advanced)



Selected working directory: C:/Users/oyvbl/Documents

Selected name for matchfile: matchfile.csv

Selected name for session folder: session

Remember to check the SETUP before searching:

Section 1

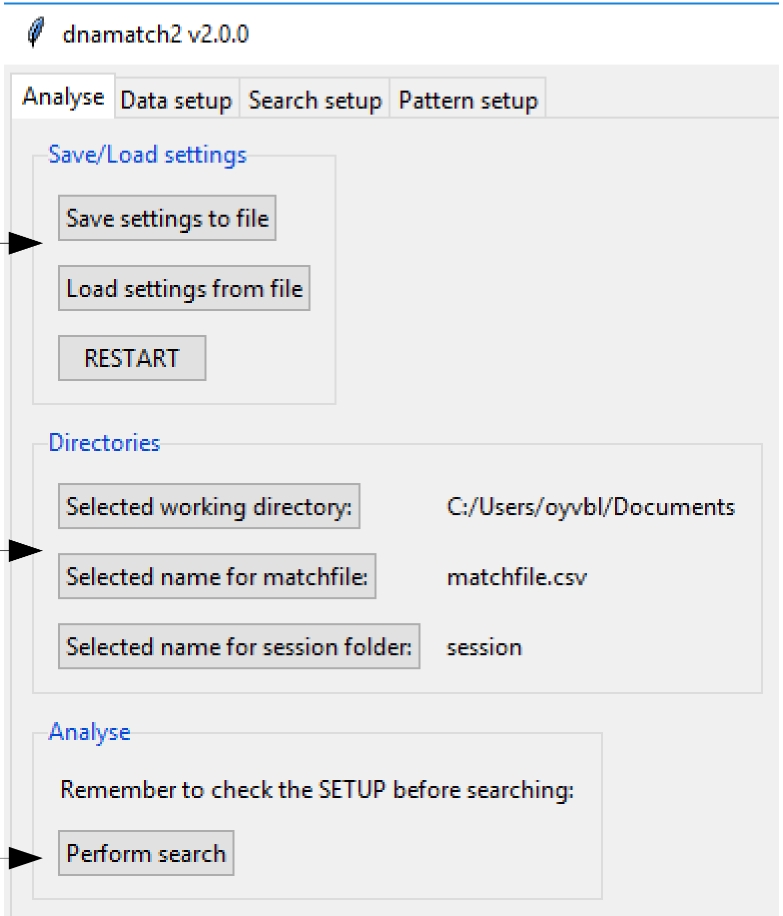
Saving/loading project file

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

Executes the search



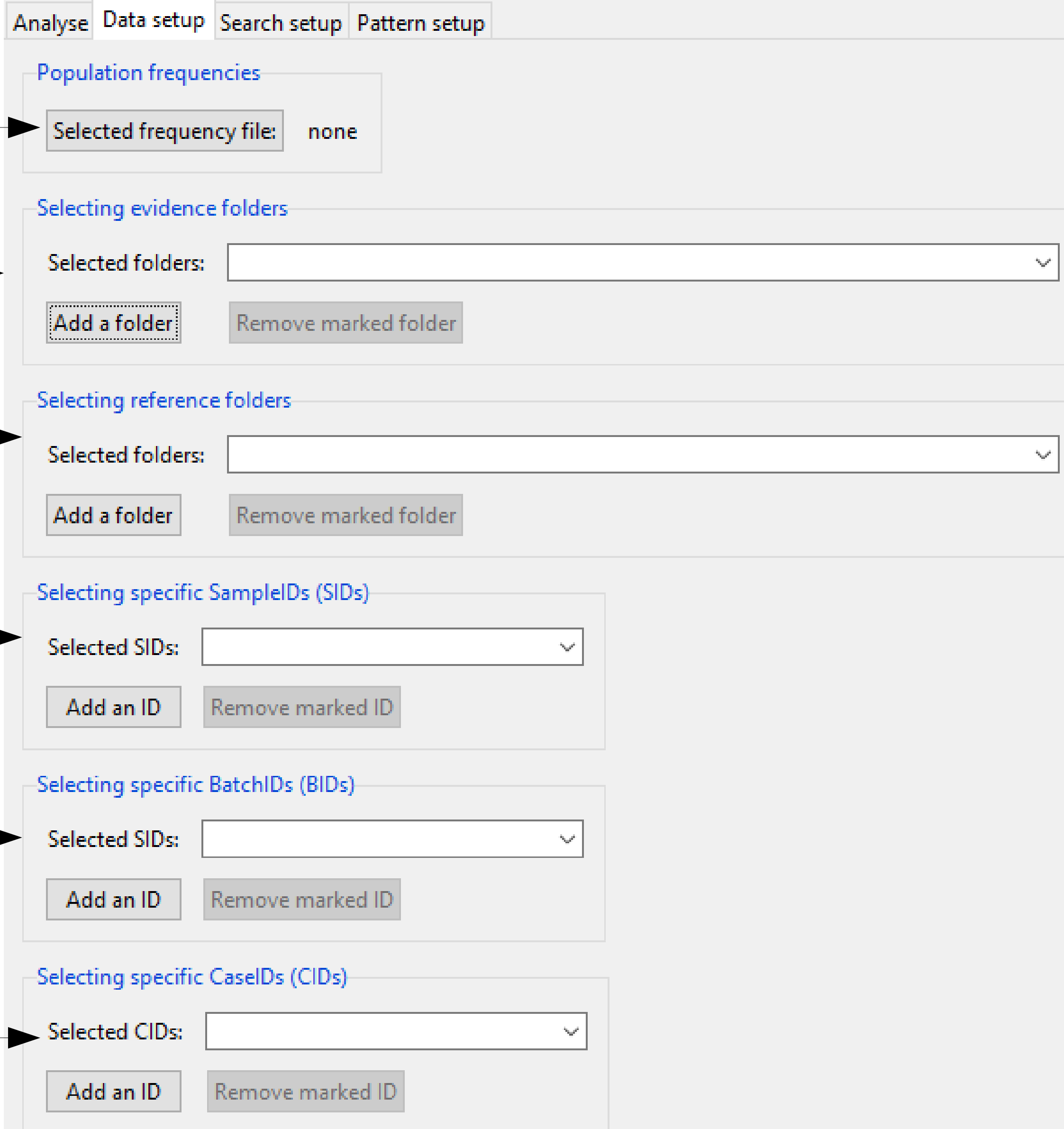
Selected working directory: C:/Users/oyvbl/Documents

Selected name for matchfile: matchfile.csv

Selected name for session folder: session

Remember to check the SETUP before searching:

Section 2



A population frequency file (mandatory)

One or more folders including files with evidence profiles (mandatory)

Folder(s) including files with reference profiles (optional)

Specific Samples to search (optional)

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

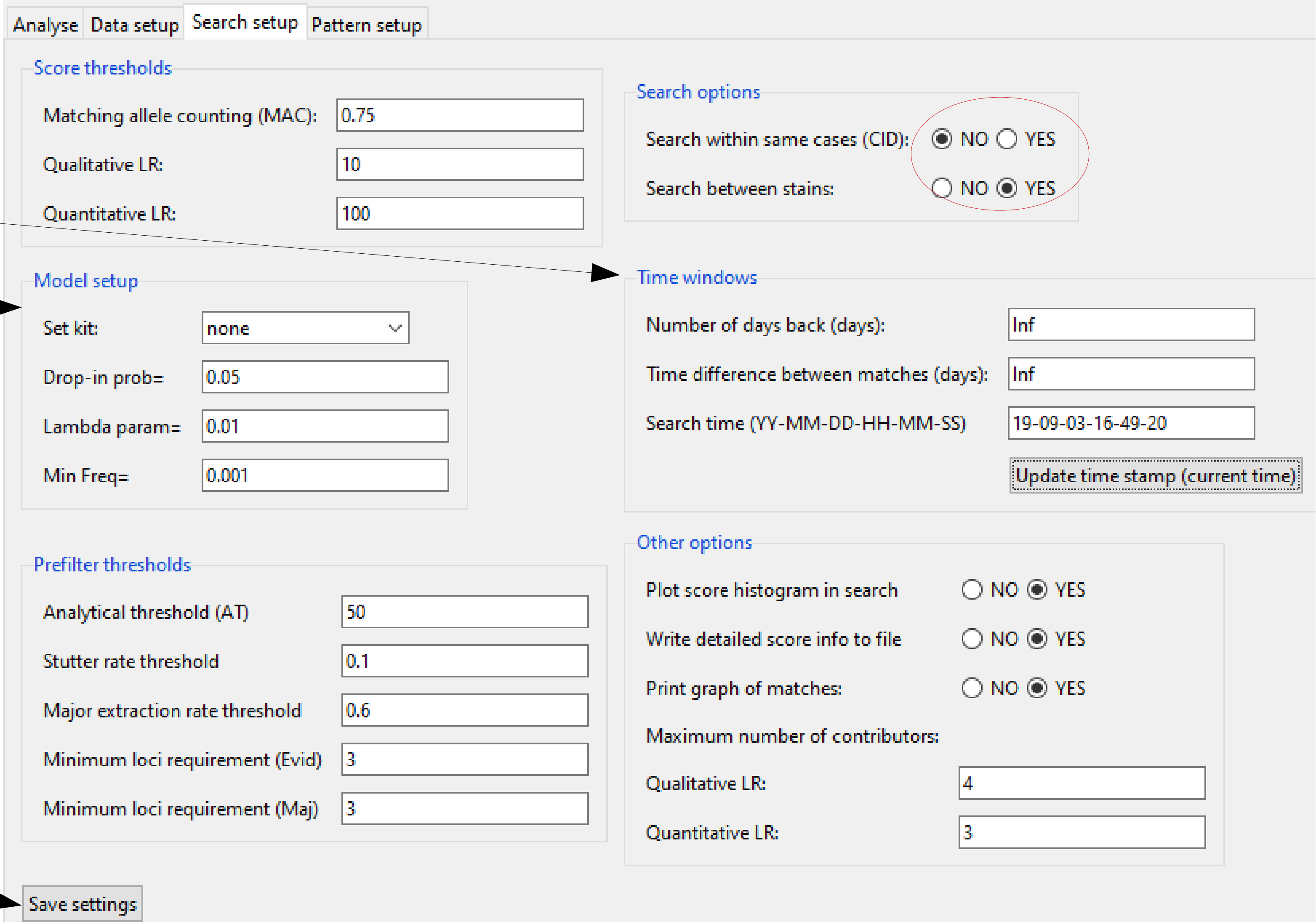
Specific Case(s) to search (optional)

Section 3: Search settings

Possible to change search time (set when GUI opened)

Selecting kit will activate the degradation model

Don’t forget to save the settings!



Score thresholds

Matching allele counting (MAC): 0.75

Qualitative LR: 10

Quantitative LR: 100

Model setup

Set kit: none

Drop-in prob= 0.05

Lambda param= 0.01

Min Freq= 0.001

Prefilter thresholds

Analytical threshold (AT) 50

Stutter rate threshold 0.1

Major extraction rate threshold 0.6

Minimum loci requirement (Evid) 3

Minimum loci requirement (Maj) 3

Search options

Search within same cases (CID): ● NO ○ YES

Search between stains: ○ NO ● YES

Time windows

Number of days back (days): Inf

Time difference between matches (days): Inf

Search time (YY-MM-DD-HH-MM-SS) 19-09-03-16-49-20

Other options

Plot score histogram in search ○ NO ● YES

Write detailed score info to file ○ NO ● YES

Print graph of matches: ○ NO ● YES

Maximum number of contributors:

Qualitative LR: 4

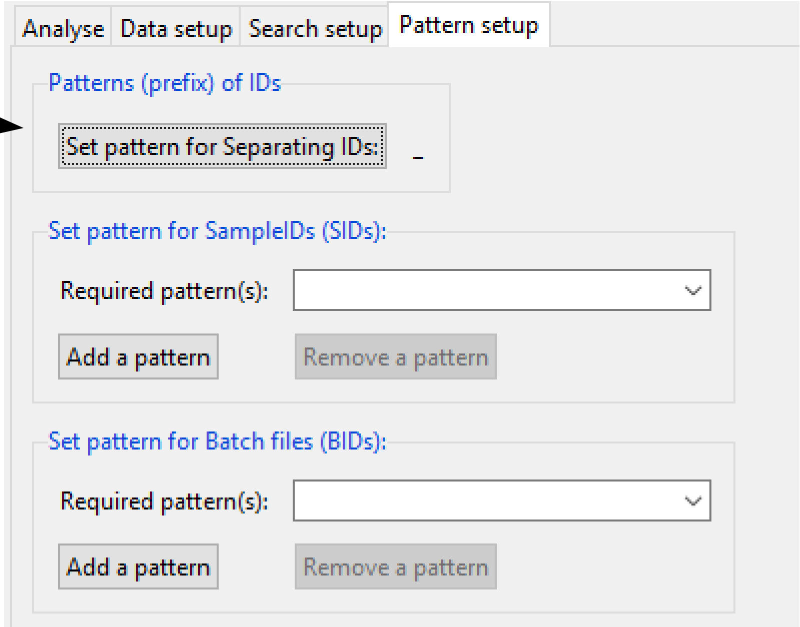
Quantitative LR: 3

Section 4: Pattern setup

Used to recognize (the seperator sign)
SID = SampleID
RID = ReplicateID
CID = CaseID
Format of SampleNames = SID_RID_CID
(mandatory)

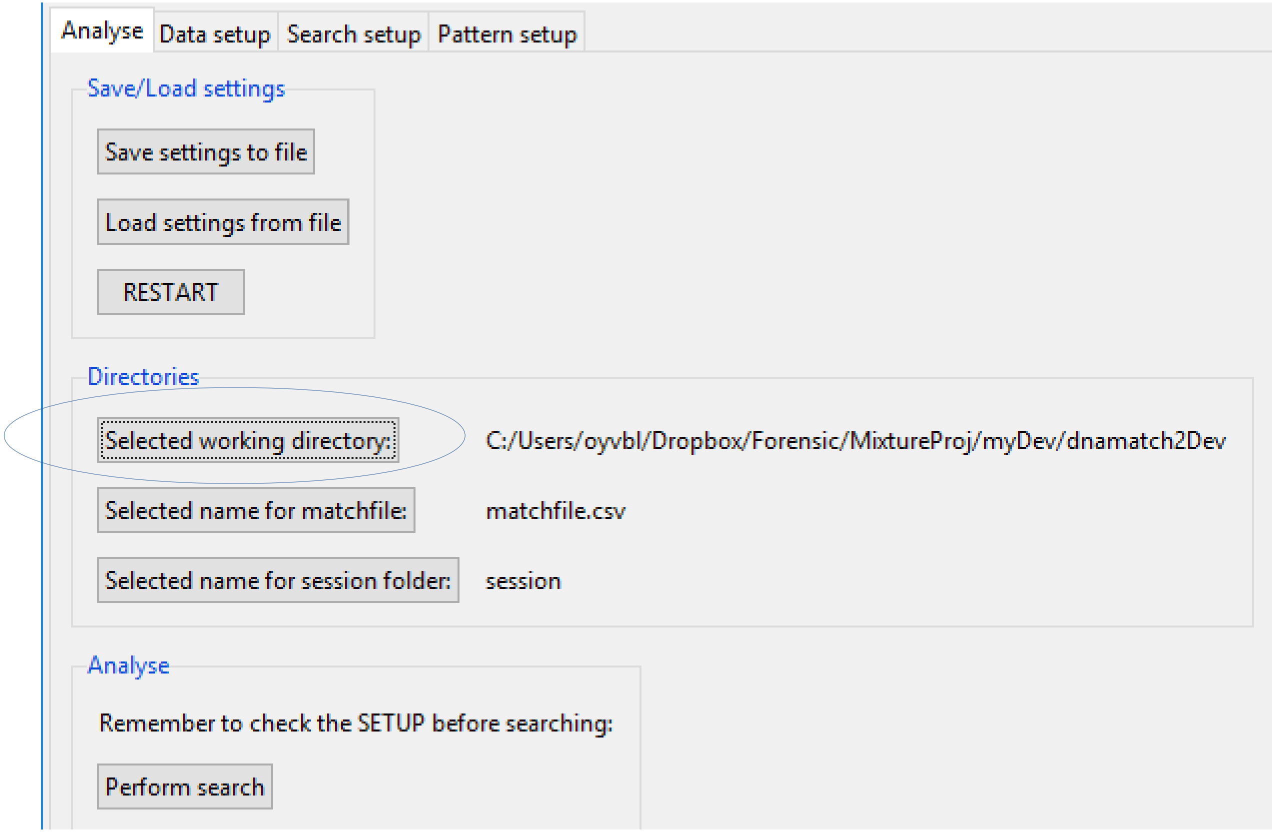
Used to restrict on specific patterns in SID
Useful for filtering specific sample types.
(optional)

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



Tutorial using a Fusion 6C example

Setup (page 1)



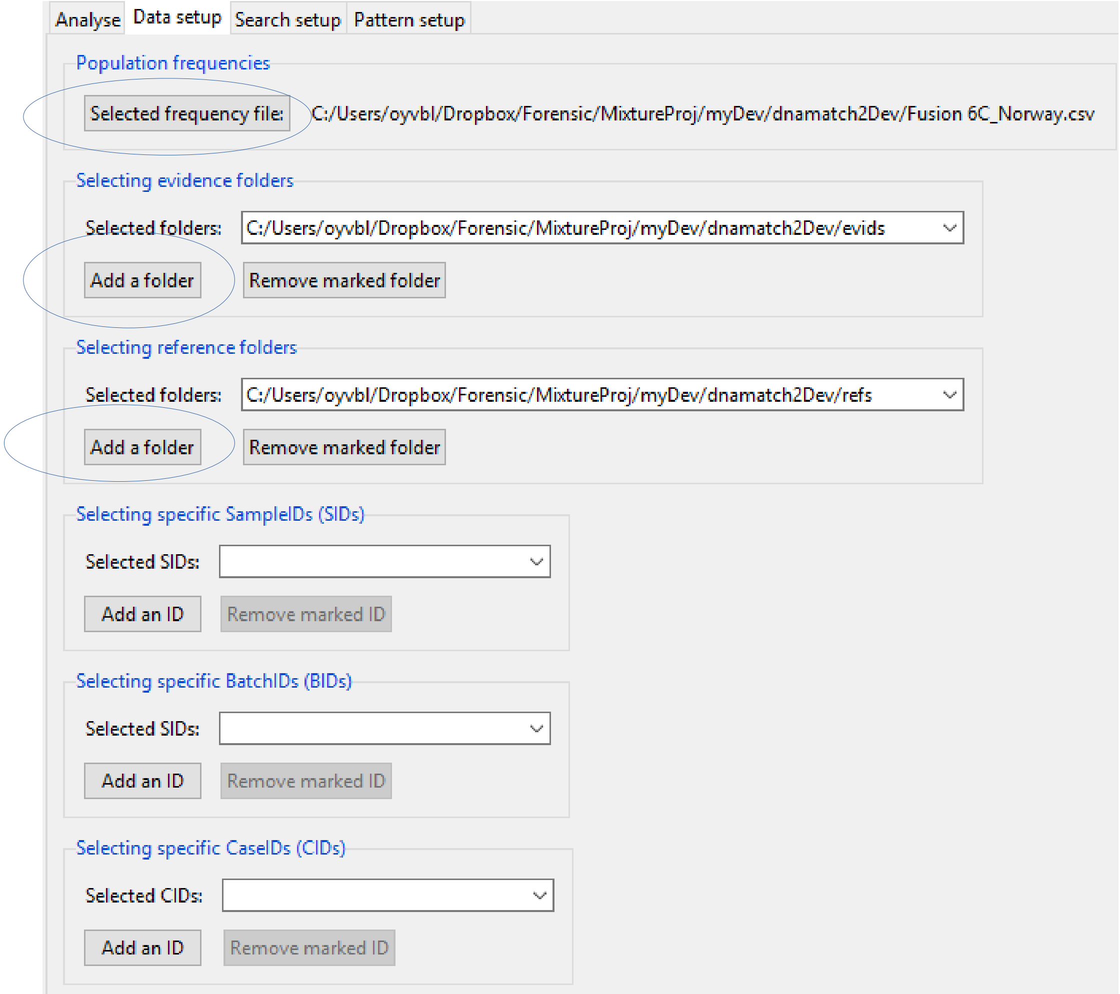
C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev

matchfile.csv

session

Remember to check the SETUP before searching:

Setup (page 2)



Population frequencies

Selected frequency file: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/Fusion 6C_Norway.csv

Selecting evidence folders

Selected folders: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/evids

Selecting reference folders

Selected folders: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/refs

Selecting specific SampleIDs (SIDs)

Selected SIDs:

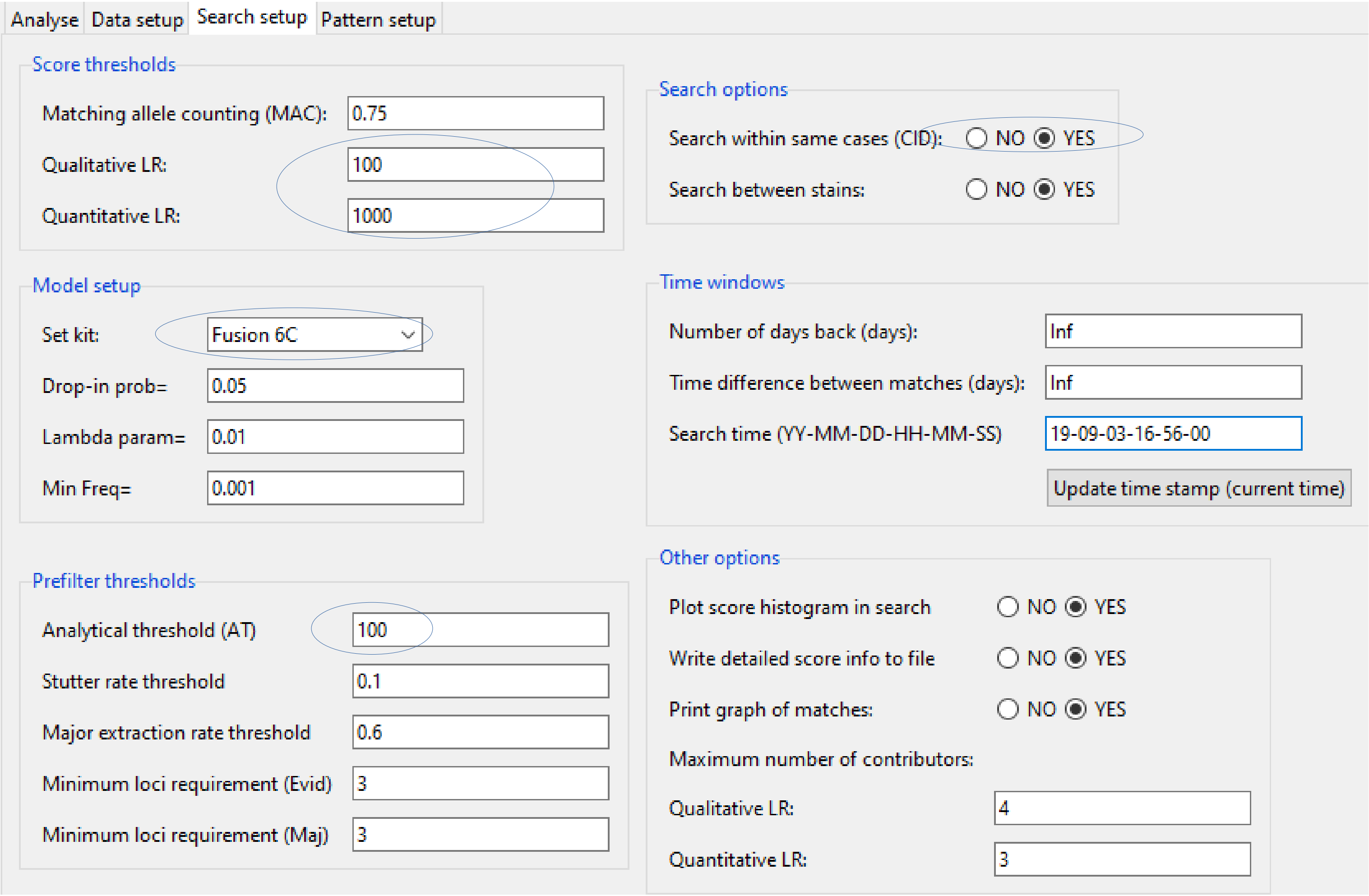
Selecting specific BatchIDs (BIDs)

Selected SIDs:

Selecting specific CaseIDs (CIDs)

Selected CIDs:

Setup (page 3)



Score thresholds

Matching allele counting (MAC): 0.75

Qualitative LR: 100

Quantitative LR: 1000

Model setup

Set kit: Fusion 6C

Drop-in prob= 0.05

Lambda param= 0.01

Min Freq= 0.001

Prefilter thresholds

Analytical threshold (AT) 100

Stutter rate threshold 0.1

Major extraction rate threshold 0.6

Minimum loci requirement (Evid) 3

Minimum loci requirement (Maj) 3

Search options

Search within same cases (CID): ○ NO ● YES

Search between stains: ○ NO ● YES

Time windows

Number of days back (days): Inf

Time difference between matches (days): Inf

Search time (YY-MM-DD-HH-MM-SS) 19-09-03-16-56-00

Other options

Plot score histogram in search ○ NO ● YES

Write detailed score info to file ○ NO ● YES

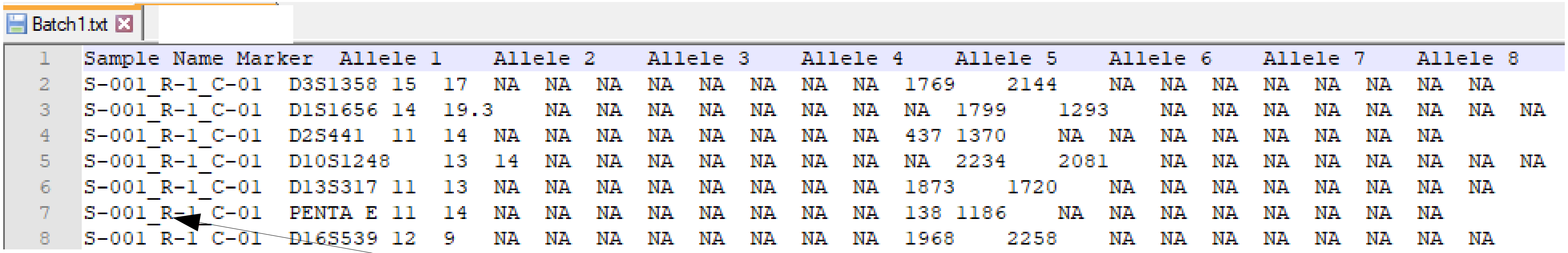
Print graph of matches: ○ NO ● YES

Maximum number of contributors:

Qualitative LR: 4

Quantitative LR: 3

Setup (page 4)



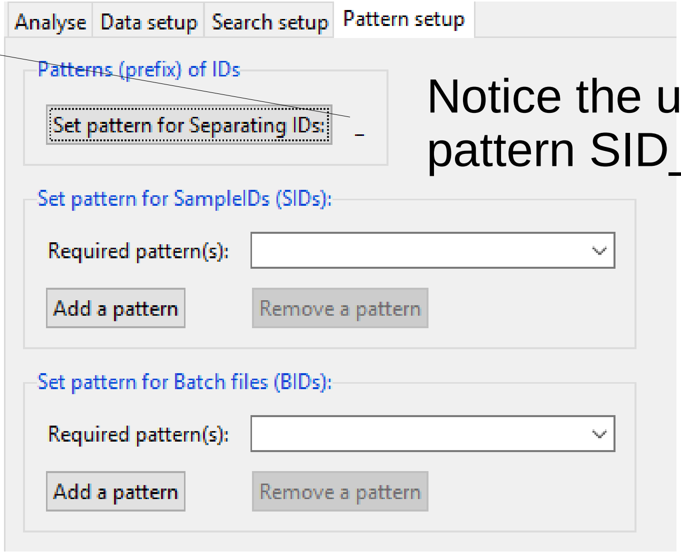
Batch1.txt

```
Sample Name Marker  Allele 1    Allele 2    Allele 3    Allele 4    Allele 5    Allele 6    Allele 7    Allele 8
S-001_R-1_C-01  D3S1358 15  17  NA  NA  NA  NA  NA  NA  NA  NA  1769    2144    NA  NA  NA  NA  NA  NA  NA  NA
S-001_R-1_C-01  D1S1656 14  19.3    NA  NA  NA  NA  NA  NA  NA  NA  1799    1293    NA  NA  NA  NA  NA  NA  NA  NA
S-001_R-1_C-01  D2S441  11  14  NA  NA  NA  NA  NA  NA  NA  NA  437 1370    NA  NA  NA  NA  NA  NA  NA  NA
S-001_R-1_C-01  D10S1248    13  14  NA  NA  NA  NA  NA  NA  NA  NA  2234    2081    NA  NA  NA  NA  NA  NA  NA  NA
S-001_R-1_C-01  D13S317 11  13  NA  NA  NA  NA  NA  NA  NA  NA  1873    1720    NA  NA  NA  NA  NA  NA  NA  NA
S-001_R-1_C-01  PENTA E 11  14  NA  NA  NA  NA  NA  NA  NA  NA  138 1186    NA  NA  NA  NA  NA  NA  NA  NA
S-001_R-1_C-01  D16S539 12  9   NA  NA  NA  NA  NA  NA  NA  NA  1968    2258    NA  NA  NA  NA  NA  NA  NA  NA
```

Use default values

We could have used
SIDpattern = “S-”

BIDpattern = “Batch”



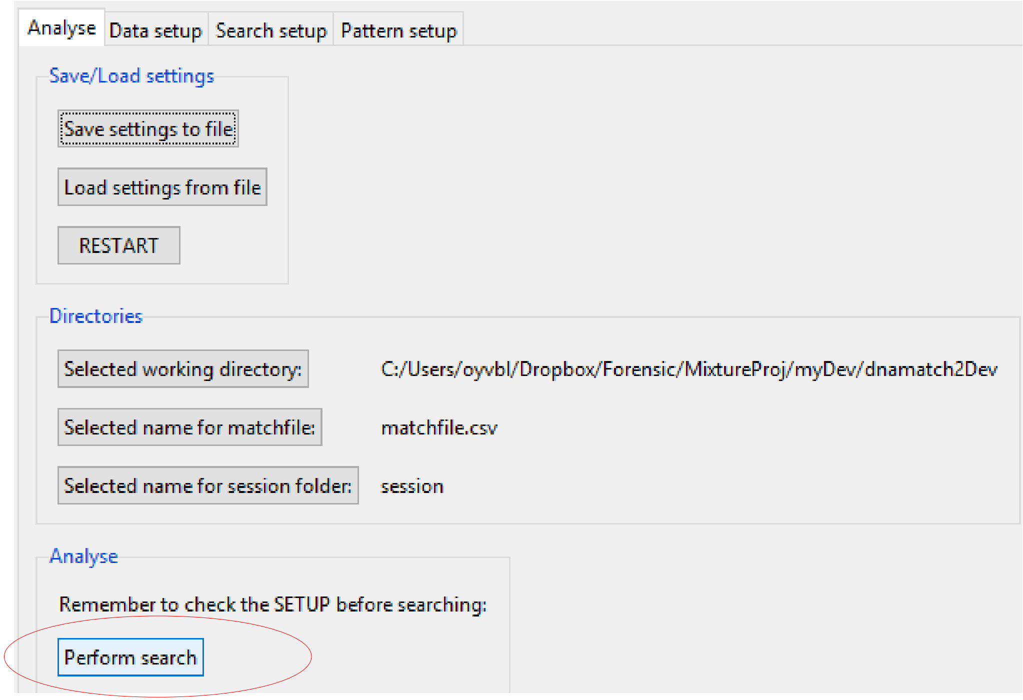
Set pattern for Separating IDs: _

Required pattern(s):

Required pattern(s):

Notice the underscore
pattern SID_RID_CID

Time to perform search!

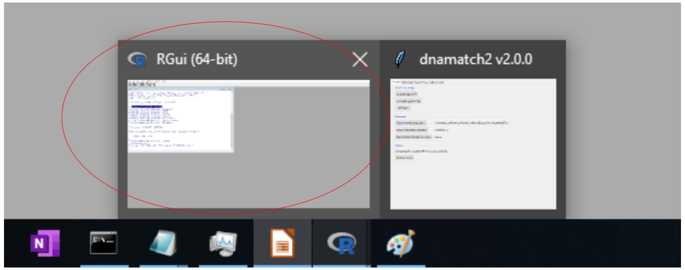


Selected working directory: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev

Selected name for matchfile: matchfile.csv

Selected name for session folder: session

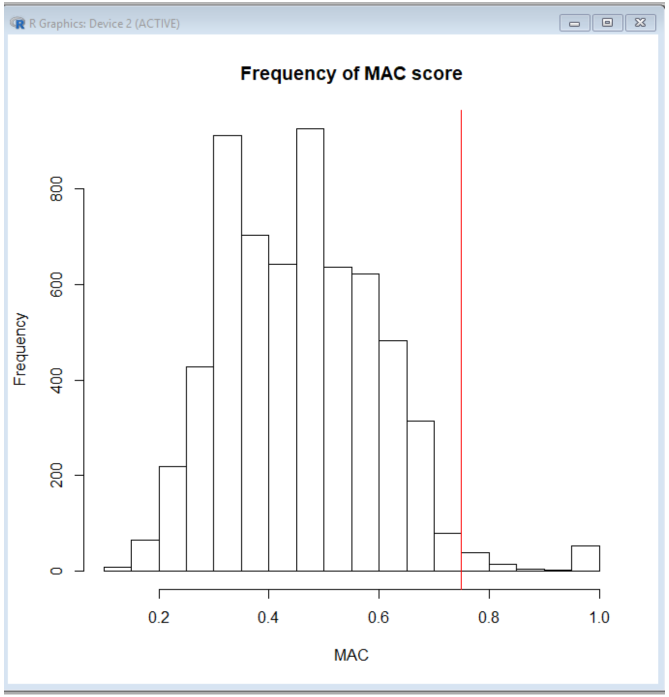
Remember to check the SETUP before searching:



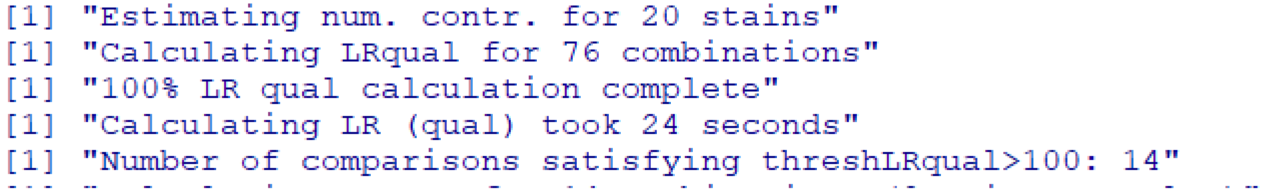
Open Rgui and look on the R-console for progress information!

Temporary results: MAC

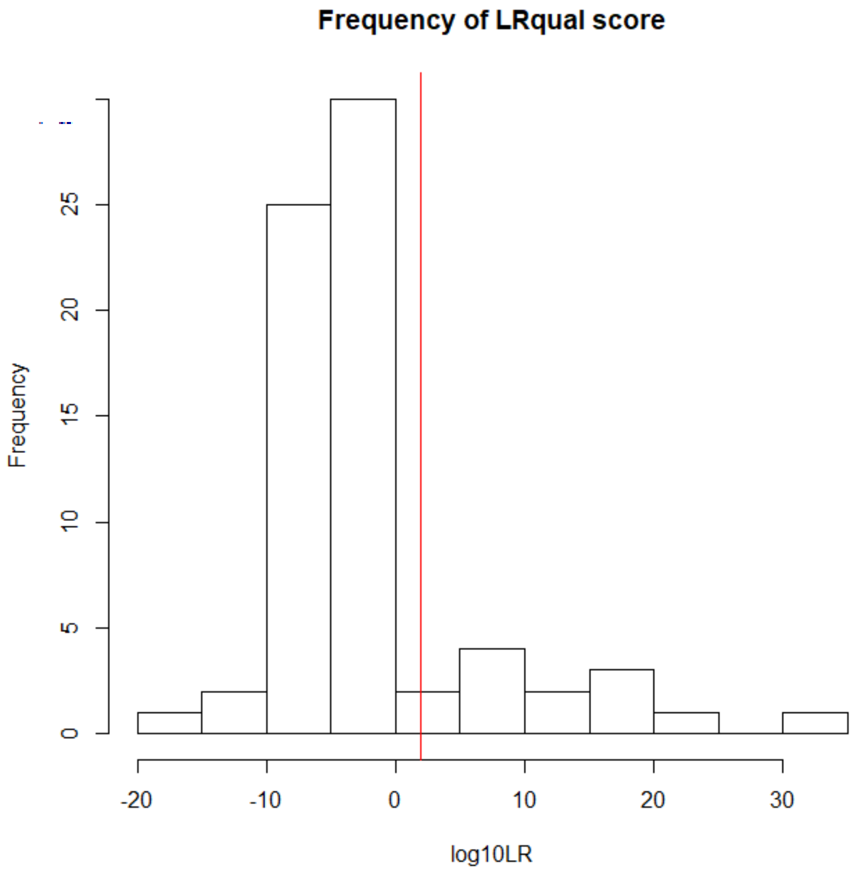
```
[1] "Number of loci to use: 23"
[1] "Number of imported samples: 43"
[1] "Number of imported references: 143"
[1] "Calculating MAC for all 6149 comparisons: All refs against all stains"
[1] "Calculating MAC scores took 1 seconds"
[1] "Number of comparisons satisfying (after filters) threshMAC=0.75: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



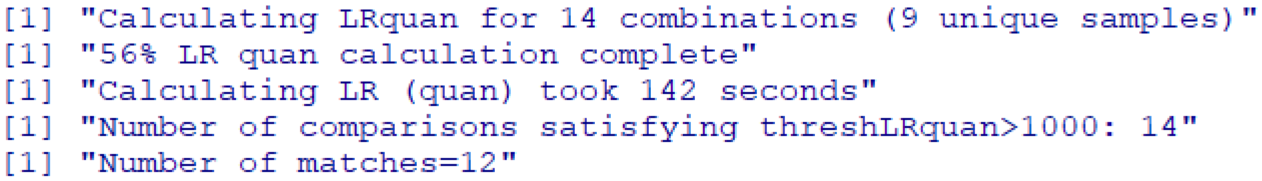
Temporary results: LRqual



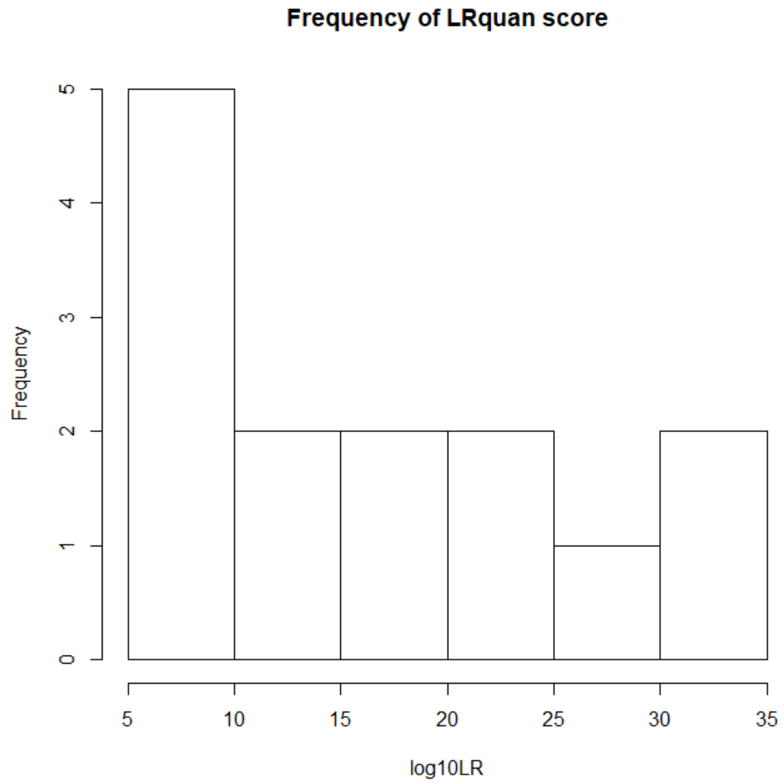
```
[1] "Estimating num. contr. for 20 stains"
[1] "Calculating LRqual for 76 combinations"
[1] "100% LR qual calculation complete"
[1] "Calculating LR (qual) took 24 seconds"
[1] "Number of comparisons satisfying threshLRqual>100: 14"
```



Final results: LRquan

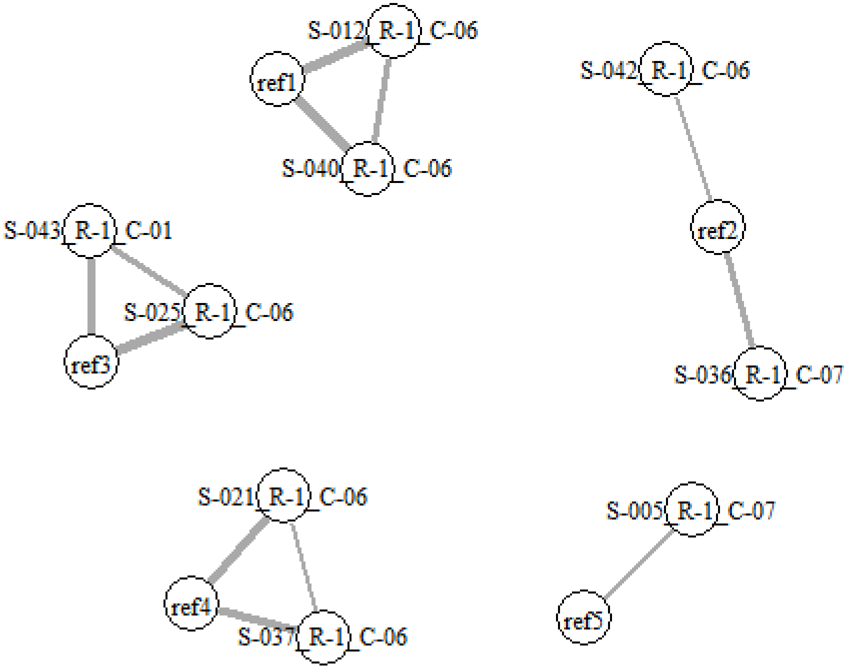


```
[1] "Calculating LRquan for 14 combinations (9 unique samples)"
[1] "56% LR quan calculation complete"
[1] "Calculating LR (quan) took 142 seconds"
[1] "Number of comparisons satisfying threshLRquan>1000: 14"
[1] "Number of matches=12"
```



Frequency of LRquan score

Match graph



True matchlist:

Ref1: S40,S12

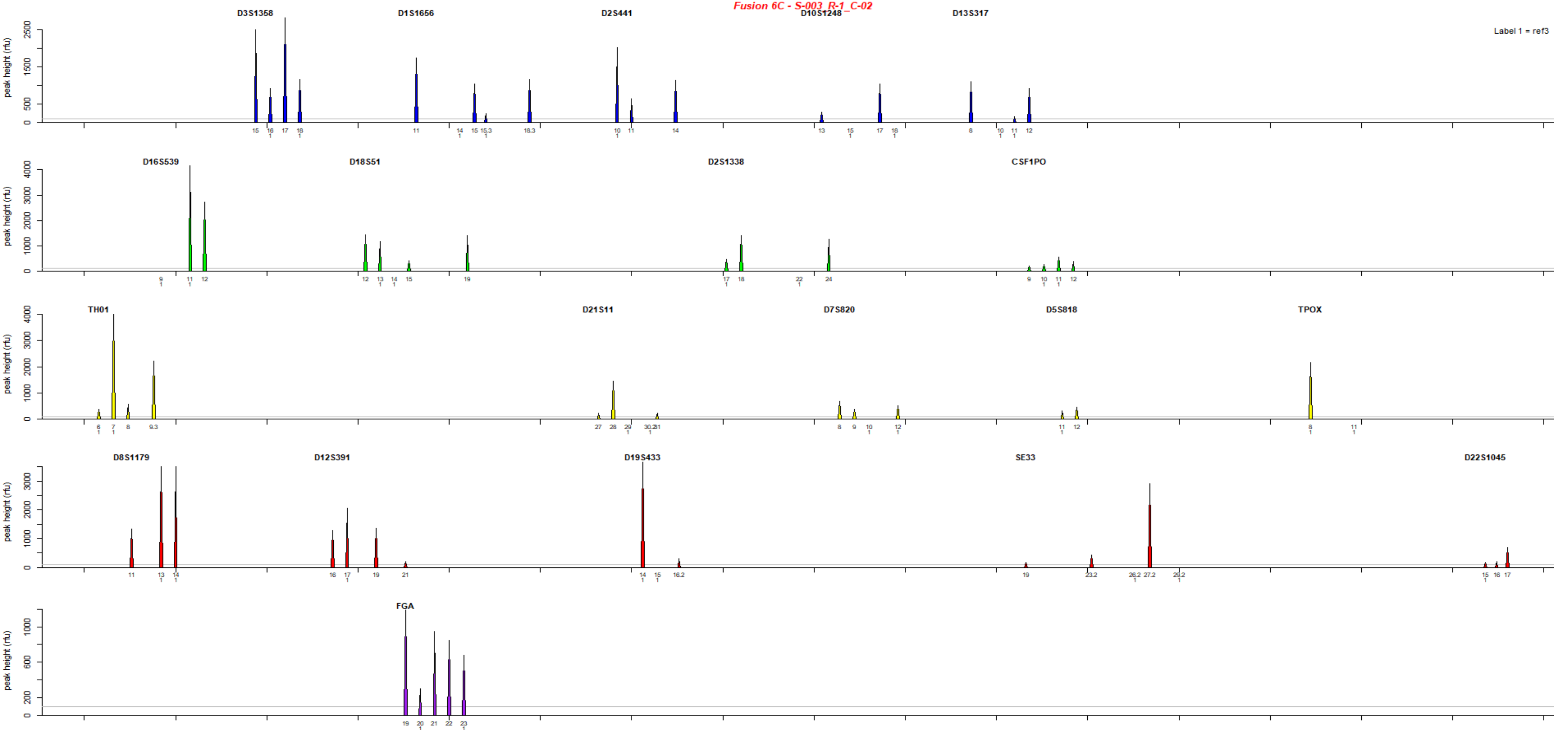
Ref2: S42,S36

Ref3: S43,S25,S3 (missing)

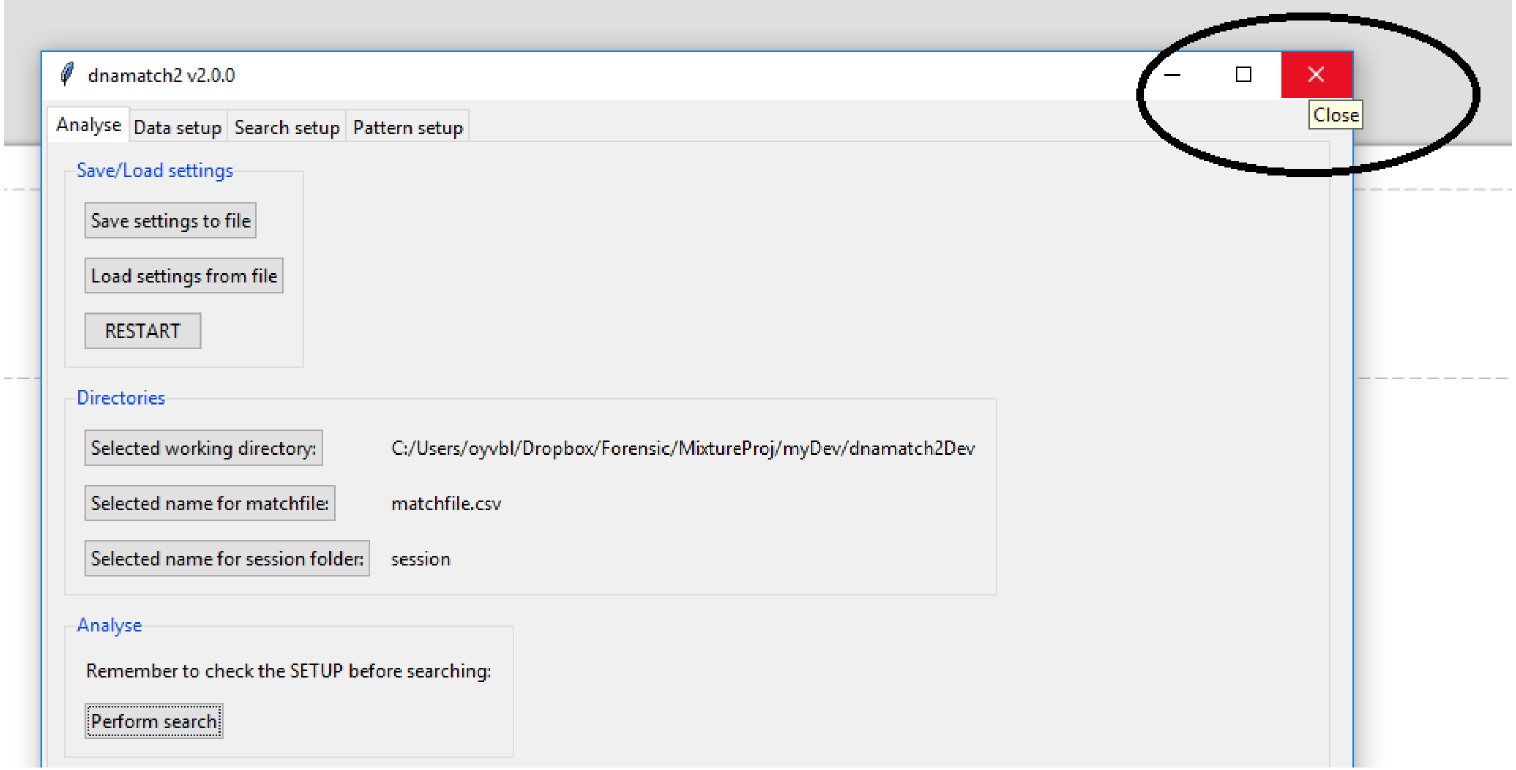
Ref4: S37,S21

Ref5: S5

S3 (false negative)



Close GUI after use



C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev

matchfile.csv

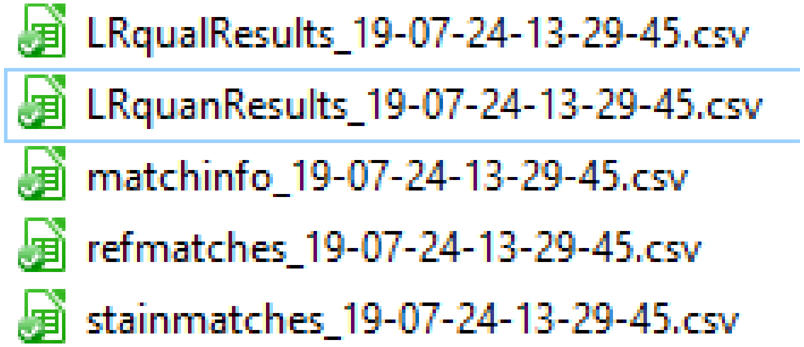
session

Remember to check the SETUP before searching:

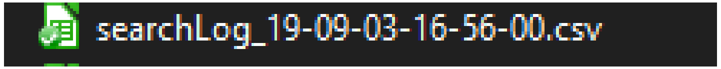
Result and Log files

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

In session folder: Log files given by the time stamp

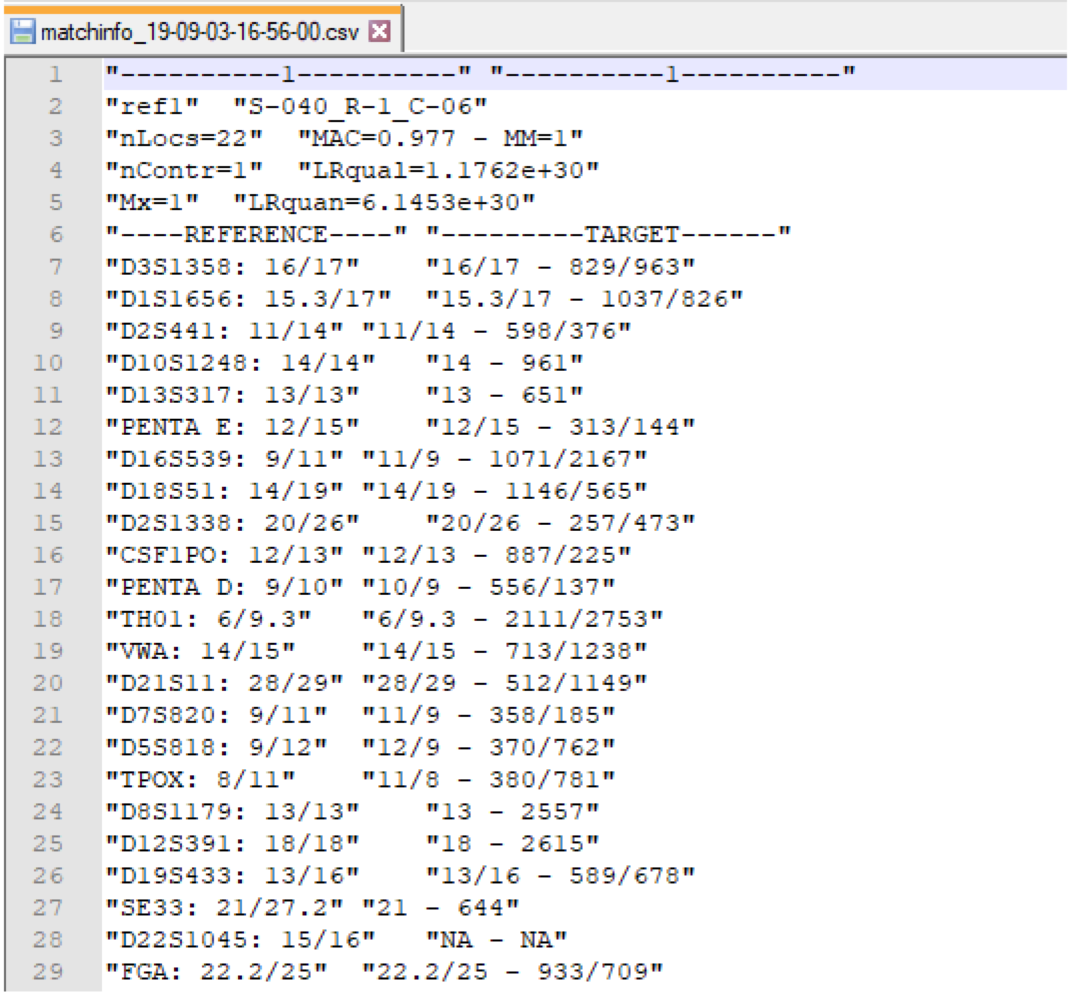


LRqualResults_19-07-24-13-29-45.csv	Overview of all LRs in QUAN comparison
LRquanResults_19-07-24-13-29-45.csv	Overview of all LRs in QUAN comparison
matchinfo_19-07-24-13-29-45.csv	Profiles details for matches
refmatches_19-07-24-13-29-45.csv	List of ref-evid matches
stainmatches_19-07-24-13-29-45.csv	List of evid-evid matches

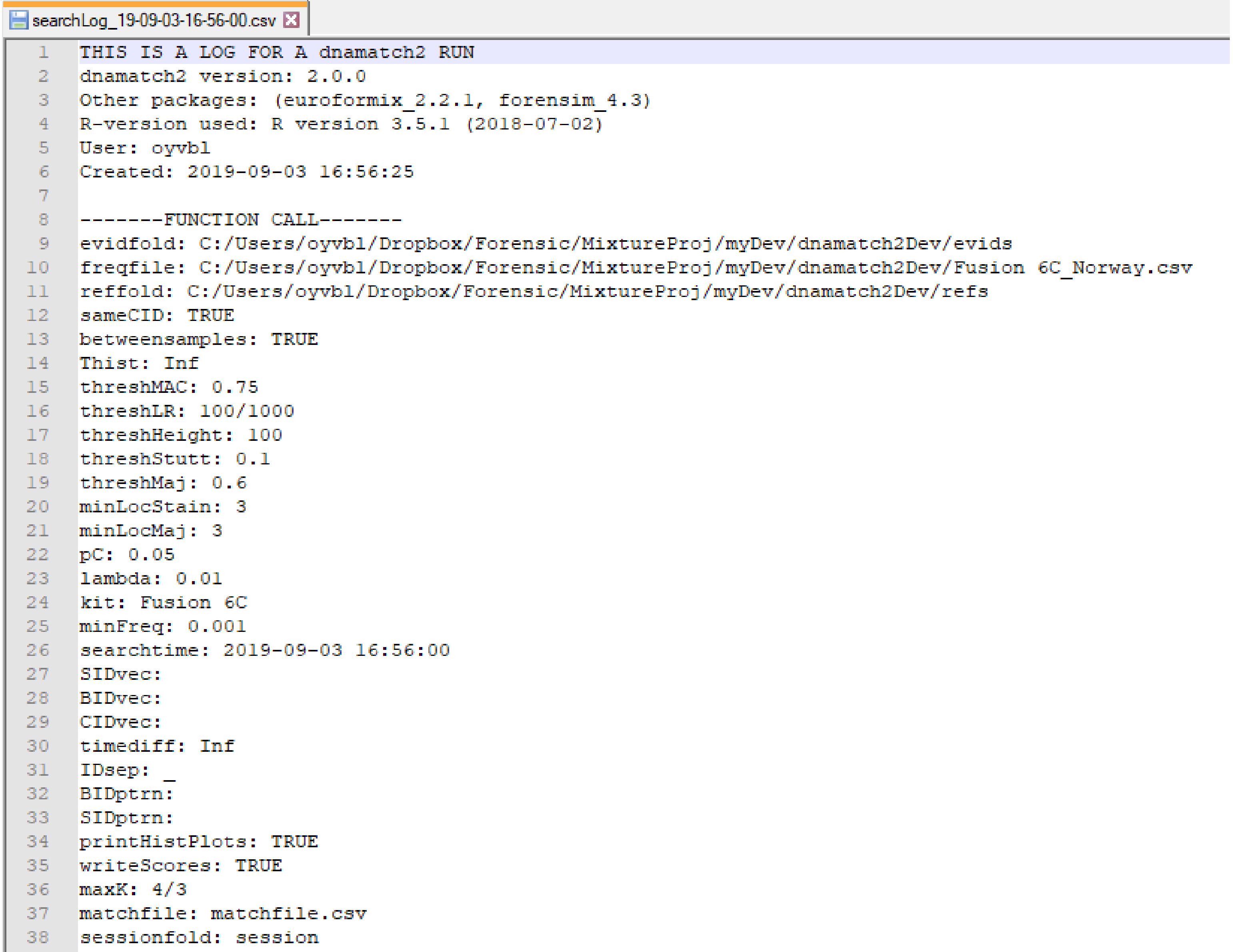


searchLog_19-09-03-16-56-00.csv — Log with argument given to search

Matchinfo



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



```
THIS IS A LOG FOR A dnamatch2 RUN
dnamatch2 version: 2.0.0
Other packages: (euroformix_2.2.1, forensim_4.3)
R-version used: R version 3.5.1 (2018-07-02)
User: oyvbl
Created: 2019-09-03 16:56:25

-------FUNCTION CALL-------
evidfold: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/evids
freqfile: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/Fusion 6C_Norway.csv
reffold: C:/Users/oyvbl/Dropbox/Forensic/MixtureProj/myDev/dnamatch2Dev/refs
sameCID: TRUE
betweensamples: TRUE
Thist: Inf
threshMAC: 0.75
threshLR: 100/1000
threshHeight: 100
threshStutt: 0.1
threshMaj: 0.6
minLocStain: 3
minLocMaj: 3
pC: 0.05
lambda: 0.01
kit: Fusion 6C
minFreq: 0.001
searchtime: 2019-09-03 16:56:00
SIDvec:
BIDvec:
CIDvec:
timediff: Inf
IDsep: _
BIDptrn:
SIDptrn:
printHistPlots: TRUE
writeScores: TRUE
maxK: 4/3
matchfile: matchfile.csv
sessionfold: session
```