

# 'Internal development' validation of EuroForMix (version 3.4.0)

Numerical tests automatically checked using *testthat* (R-package)  
Manual checks for GUI results

Carried out regarding euroformix R-package release

By Øyvind Bleka (05. February 2022)

# Document updates

- Changed for v3.4.0 from v3.1.0 (05.02.22)
  - Included extended kinship testing
  - All typed individuals are taken into account
  - Changed round-off errors

# Critical to test numerical calculations done by C++

- 4 C++ functions to test

1. likelihood function for 1 replicate, no conditional references or relatedness,

- a. No stutter model `double calcLogLikGammaMarkerUnknown`

- b. Forward/backward models `double calcLogLikGammaMarkerUnknownStutter`

2. Otherwise another likelihood function is run (slower)

```
double calcLogLikGammaMarker
```

3. cumulative likelihood function (all model variants)

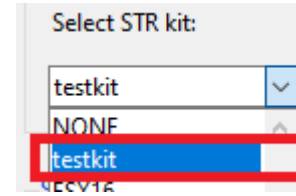
```
void calccumval
```

# “Full model” is tested

- Degradation model
- Backward stutter model
- Forward stutter model
- Dye specific settings for
  - Analytical thresholds (AT)
  - Drop-in probabilities (pC)
  - Drop-in distribution ( $\lambda$ )
  - Theta-correction (Fst)
- Non-mixtures vs mixtures (1 contributor vs more)
- Unknowns vs conditioning
- Unrelated vs kinship for unknown (sibling and child)
- Stutters vs no stutters

# Example data

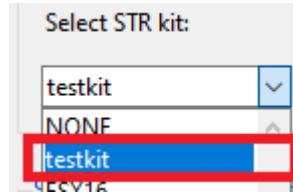
- Constructing the testkit (subset of SGMplus)
  - Two dyes only (blue and yellow)
  - TH01 renamed to 'TH 01' (white space in names)



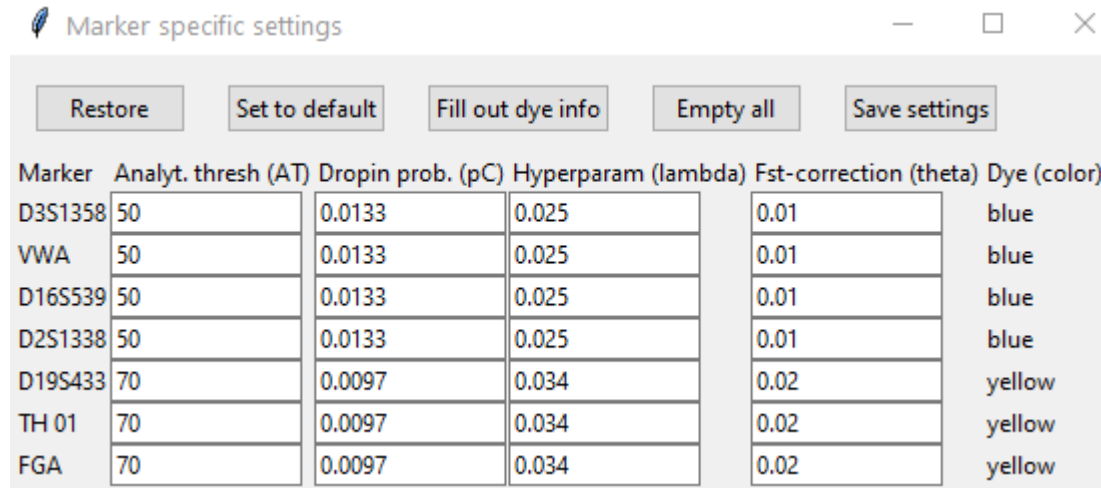
- Frequencies based on 'SGMplus\_Norway'
- 2 example data sets which includes backward and forward stutters
  1. 1 contributor, 1 replicate
  2. 2 contributors, 2 replicates
- 1 example data set including SNPs.
  - 2 contributors, 2 replicates
- Data available in R-package installation folder **euroformix**\*examples*

# Settings

- Selected kit: testkit  
(Not for SNPs)



- Dye specific settings:  
(Not for SNPs)

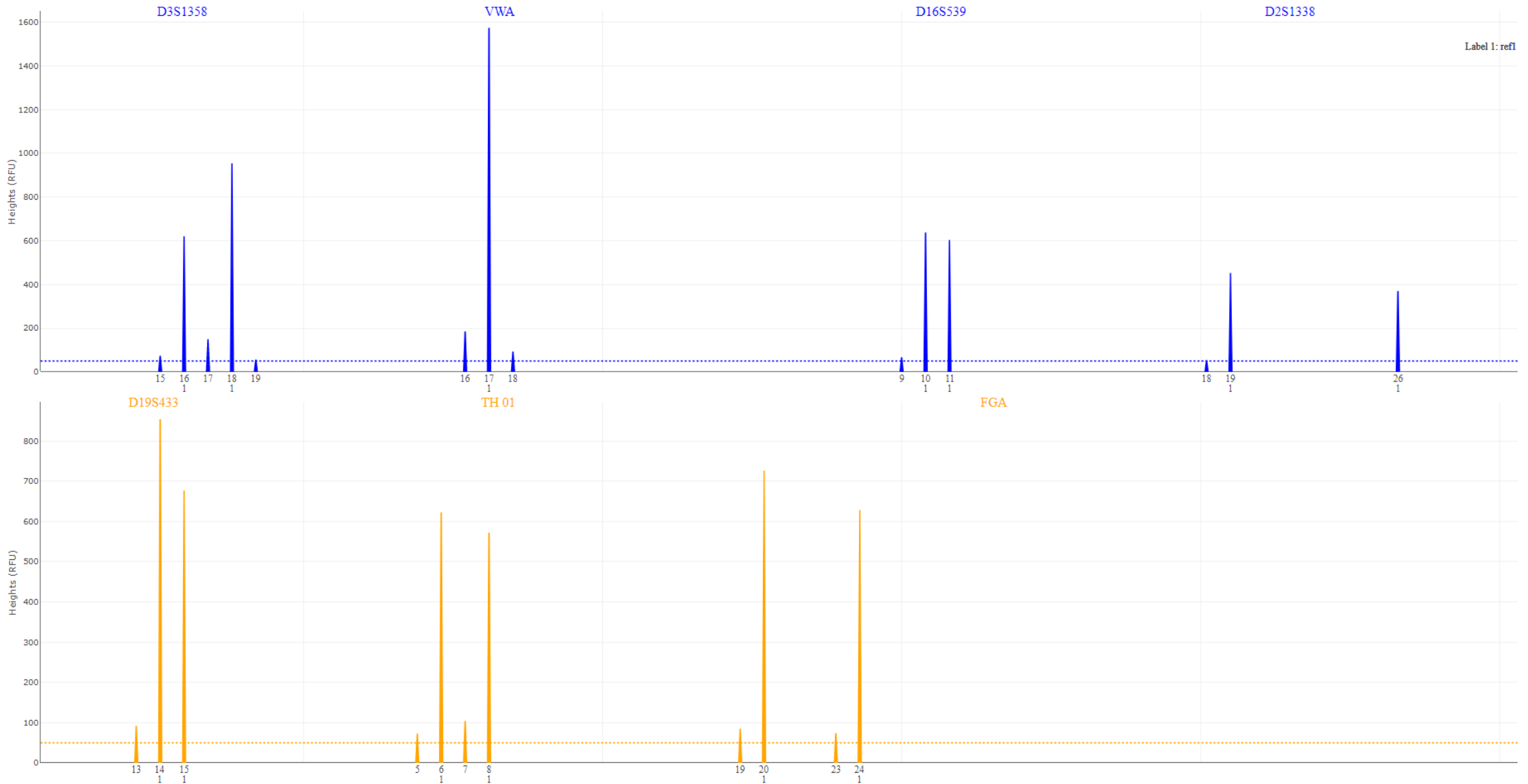
A screenshot of a "Marker specific settings" window. At the top, there are five buttons: "Restore", "Set to default", "Fill out dye info", "Empty all", and "Save settings". Below the buttons is a table with the following columns: "Marker", "Analyt. thresh (AT)", "Dropin prob. (pC)", "Hyperparam (lambda)", "Fst-correction (theta)", and "Dye (color)".

Marker	Analyt. thresh (AT)	Dropin prob. (pC)	Hyperparam (lambda)	Fst-correction (theta)	Dye (color)
D3S1358	50	0.0133	0.025	0.01	blue
VWA	50	0.0133	0.025	0.01	blue
D16S539	50	0.0133	0.025	0.01	blue
D2S1338	50	0.0133	0.025	0.01	blue
D19S433	70	0.0097	0.034	0.02	yellow
TH 01	70	0.0097	0.034	0.02	yellow
FGA	70	0.0097	0.034	0.02	yellow

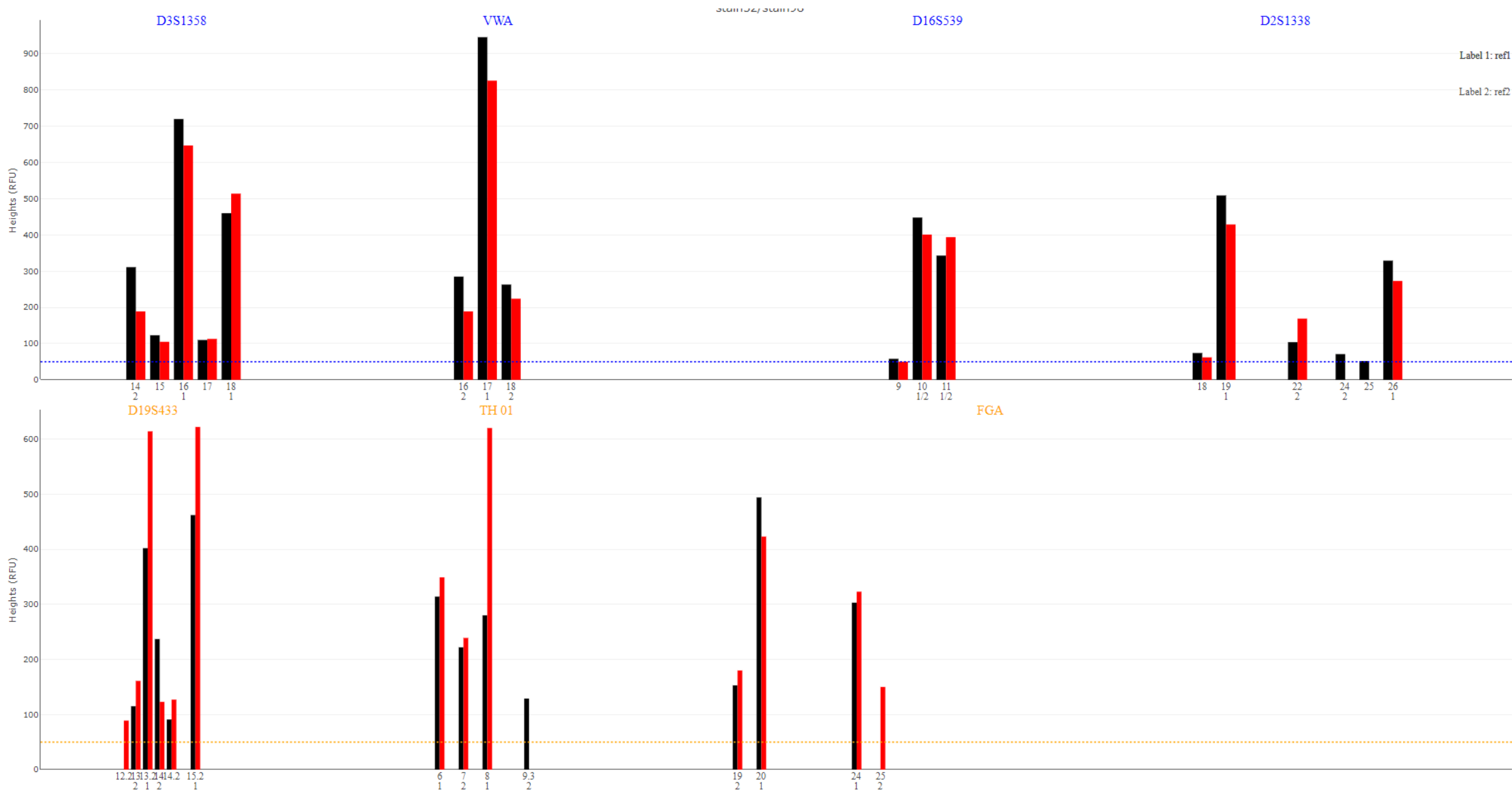
- Optimization settings:

```
-----Optimalisation setting-----  
Required number of (identical) optimizations: 2  
Accuracy of optimisations (steptol): 0.001  
Seed for optimisations: 1
```

# Example data 1: stain34

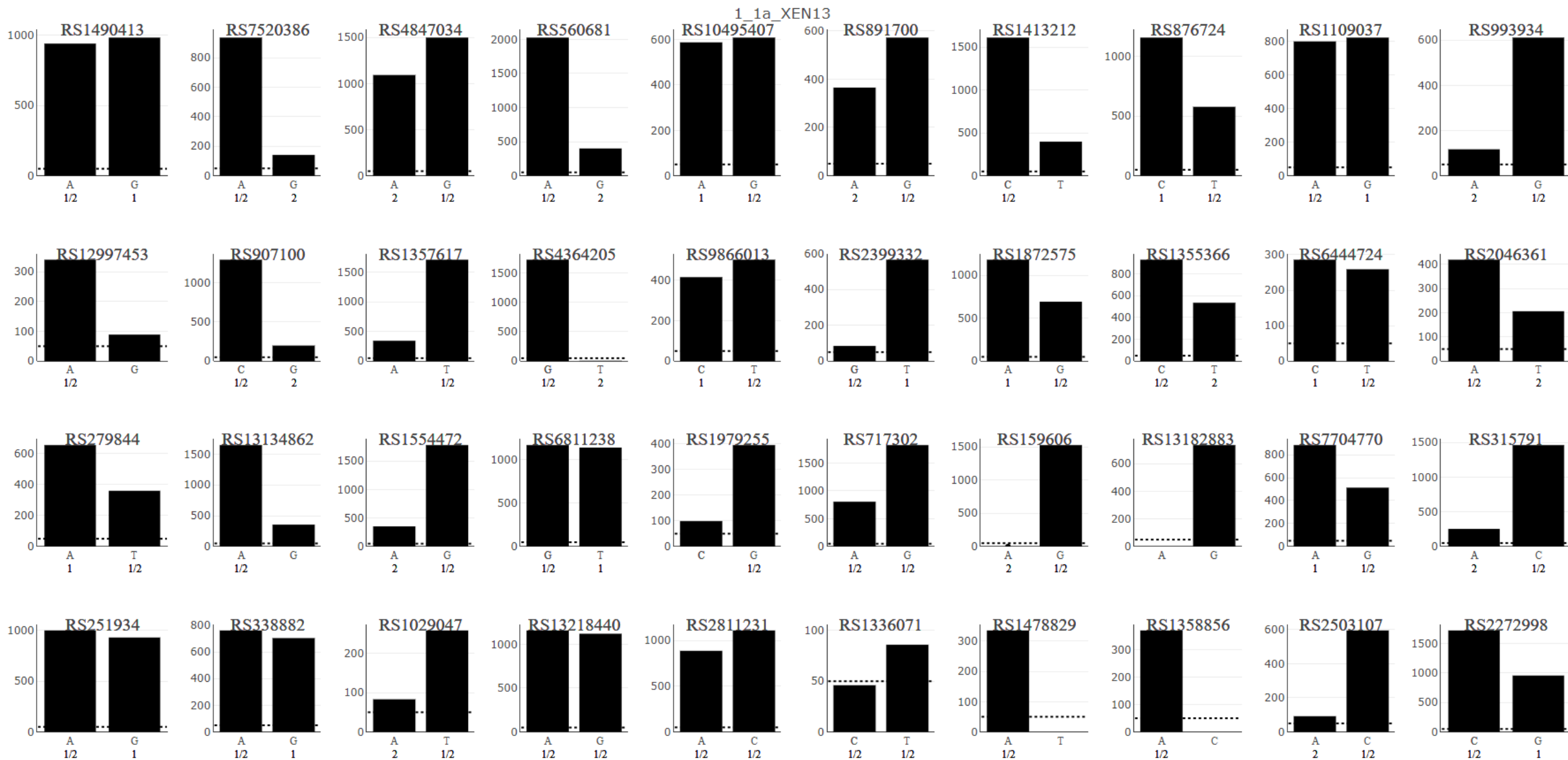


# Example data 2: stain52/stain98





# Example data 3: 1\_1a\_XEN13/ 1\_1b\_XEN13 (SNPs)



# Checking R-functions depending on C++

- contLikMLE: Obtain maximum likelihood value/parameter estimates
- logLiki: Calculate per-marker likelihood
- validMLEmodel: Calculates cumulative probabilities (of PHs)
- deconvolve: Calculates posterior probabilities of unknown genotypes
- contLikINT: Bayesian approach (integration)
- contLikMCMC: MCMC simulation
- Functions which are indirectly checked:
  - tableReader: Imports tables from files
  - freqImport: Imports allele frequencies
  - sample\_tableToList: Converts table data to list format
  - prepareData: Restructures the data prepared as input
  - calcRMPfst: Used to obtain “Upper boundary” of LR
  - Qassignate: Performs allele frequency imputation and Q-designation
  - prepareC: Used to prepare data for running C++ code.

# Propositions tested for dataset 1: *test\_logLik1contr*

- Hyp 1: ref1 is contributor
- Hyp 2: 1 (unrelated) unknown is contributor, ref1 is known non-contributor
- Hyp 3: 1 unknown sibling of ref1 is contributor, ref1 is known non-contributor

# GUI check Hp: Hyp 1 vs Hd: Hyp 2

**Model specification**

Contributor(s) under Hp:  
 ref1  
#unknowns (Hp): 0

Contributor(s) under Hd:  
 ref1  
#unknowns (Hd): 1

Last unknown is  
Unrelated

to

**Model options**

Degradation:  YES  NO  
BW Stutter:  YES  NO  
FW Stutter:  YES  NO

**Data selection**

Loci:	stain34	ref1
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH 01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

**Estimates under Hd**

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	1.000	0.000
P.H.expectation	921.931	39.723
P.H.variability	0.099	0.014
Degrad. slope	0.771	0.036
BWstutt-prop.	0.104	0.010
FWstutt-prop.	0.051	0.010

Maximum Likelihood value

logLik= -155.34  
adj.loglik= -160.34  
Lik= 3.43e-68

**Estimates under Hp**

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	1.000	0.000
P.H.expectation	921.881	39.718
P.H.variability	0.099	0.014
Degrad. slope	0.771	0.036
BWstutt-prop.	0.104	0.010
FWstutt-prop.	0.051	0.010

Maximum Likelihood value

logLik= -134.1  
adj.loglik= -139.1  
Lik= 5.79e-59

**Joint LR**

LR= 1.69e+09  
log10LR= 9.228  
Upper boundary= 9.228

**Non-contributor analysis**

Select reference to replace with non-contributor:  
ref1

Sample maximum based  
Sample integrated based

**LR for each locus**

VWA	9.46
D16S539	26.05
D2S1338	141.1
D19S433	7.263
TH 01	22.85
FGA	20.5
D3S1358	14.29

# GUI check Hp: Hyp 1 vs Hd: Hyp 3

### Model specification

Contributor(s) under Hp:

ref1

#unknowns (Hp): 0

Contributor(s) under Hd:

ref1

#unknowns (Hd): 1

Last unknown is

Sibling

to

ref1

### Model options

Degradation:  YES  NO

BW Stutter:  YES  NO

FW Stutter:  YES  NO

### Data selection

Loci:	stain34	ref1
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH 01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

### Estimates under Hd

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	1.000	0.000
P.H.expectation	921.931	39.723
P.H.variability	0.099	0.014
Degrad. slope	0.771	0.036
BWstutt-prop.	0.104	0.010
FWstutt-prop.	0.051	0.010

### Maximum Likelihood value

logLik= -155.34

adj.loglik= -160.34

Lik= 3.43e-68

### Estimates under Hp

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	1.000	0.000
P.H.expectation	921.881	39.718
P.H.variability	0.099	0.014
Degrad. slope	0.771	0.036
BWstutt-prop.	0.104	0.010
FWstutt-prop.	0.051	0.010

### Maximum Likelihood value

logLik= -134.1

adj.loglik= -139.1

Lik= 5.79e-59

### Joint LR

LR= 1.69e+09

log10LR= 9.228

Upper boundary= 9.228

### Non-contributor analysis

Select reference to replace with non-contributor:

ref1

Sample maximum based

Sample integrated based

### LR for each locus

VWA	9.46
D16S539	26.05
D2S1338	141.1
D19S433	7.263
TH 01	22.85
FGA	20.5
D3S1358	14.29

# Propositions tested for dataset 2: *test\_logLik2contr2Rep*

- Hyp 1: ref1 + ref2 are contributors
- Hyp2: ref 1 + 1 (unrelated) unknown are contributors, ref2 is known non-contributor
- Hyp3: ref 1 + 1 related (sibling) unknown of ref2 are contributors, ref2 is known non-contributor

# GUI check Hp: Hyp 1 vs Hd: Hyp 2

**Model specification**

Contributor(s) under Hp:

ref1  
 ref2  
 #unknowns (Hp): 0

Contributor(s) under Hd:

ref1  
 ref2  
 #unknowns (Hd): 1

Last unknown is  
 Unrelated

to

**Model options**

Degradation:  YES  NO  
 BW Stutter:  YES  NO  
 FW Stutter:  YES  NO

**Data selection**

Loci:	stain52	stain98	ref1	ref2
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH 01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Estimates under Hd			Estimates under Hp			Joint LR		LR for each locus
Parameter estimates:			Parameter estimates:					
Param.	MLE	Std.Err.	Param.	MLE	Std.Err.	LR=		
Mix-prop. C1	0.764	0.022	Mix-prop. C1	0.796	0.016	1.99e+08		VWA 7.085
Mix-prop. C2	0.236	0.022	Mix-prop. C2	0.204	0.016	log10LR= 8.299		D16S539 15.17
P.H.expectation	796.601	38.269	P.H.expectation	794.885	38.658	Upper boundary= 9.668		D2S1338 164.4
P.H.variability	0.148	0.017	P.H.variability	0.151	0.014			D19S433 7.156
Degrad. slope	0.740	0.040	Degrad. slope	0.747	0.040			TH 01 6.239
BWstutt-prop.	0.119	0.018	BWstutt-prop.	0.119	0.015			FGA 64.68
FWstutt-prop.	0.050	0.020	FWstutt-prop.	0.055	0.017			D3S1358 3.897
Maximum Likelihood value			Maximum Likelihood value			Non-contributor analysis		
logLik=	-349.57		logLik=	-330.46		Select reference to replace with non-contributor:		
adj.loglik=	-355.57		adj.loglik=	-336.46		ref2		
Lik=	1.52e-152		Lik=	3.03e-144		Sample maximum based		
						Sample integrated based		

# GUI check Hp: Hyp 1 vs Hd: Hyp 3

**Model specification**

**Contributor(s) under Hp:**

ref1  
 ref2  
 #unknowns (Hp): 0

**Contributor(s) under Hd:**

ref1  
 ref2  
 #unknowns (Hd): 1

Last unknown is  
 Sibling  
 to  
 ref2

**Model options**

Degradation:  YES  NO  
 BW Stutter:  YES  NO  
 FW Stutter:  YES  NO

**Data selection**

Loci:	stain52	stain98	ref1	ref2
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH 01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Estimates under Hd			Estimates under Hp			Joint LR		LR for each locus	
Param.	MLE	Std.Err.	Param.	MLE	Std.Err.	LR=			
Mix-prop. C1	0.782	0.033	Mix-prop. C1	0.796	0.016	832.8		VWA	2.296
Mix-prop. C2	0.218	0.033	Mix-prop. C2	0.204	0.016	log10LR= 2.921		D16S539	2.525
P.H.expectation	794.664	38.924	P.H.expectation	794.885	38.658	Upper boundary= 9.668		D2S1338	3.806
P.H.variability	0.152	0.015	P.H.variability	0.151	0.014			D19S433	2.216
Degrad. slope	0.746	0.041	Degrad. slope	0.747	0.040			TH 01	1.97
BWstutt-prop.	0.119	0.016	BWstutt-prop.	0.119	0.015			FGA	3.405
FWstutt-prop.	0.051	0.019	FWstutt-prop.	0.055	0.017			D3S1358	2.54
<b>Maximum Likelihood value</b>			<b>Maximum Likelihood value</b>			<b>Non-contributor analysis</b>			
logLik=	-337.19		logLik=	-330.46		Select reference to replace with non-contributor:			
adj.loglik=	-343.19		adj.loglik=	-336.46		ref2			
Lik=	3.64e-147		Lik=	3.03e-144		<input type="button" value="Sample maximum based"/>			
						<input type="button" value="Sample integrated based"/>			



# Propositions tested for dataset 2' *test\_logLik2contr1RepNoStutter*

- Considering 1 replicate (stain52 only)
- Stutter models are turned off, degrad still turned on.
- Hyp 1: ref1 and ref2 are contributor
- Hyp 2: 2 (unrelated) unknowns are contributors, ref1 and ref2 are known non-contributors

# GUI check Hd: Hyp 1 versus Hd: Hyp 2

**Model specification**

Contributor(s) under Hp:

ref1  
 ref2  
#unknowns (Hp): 0

Contributor(s) under Hd:

ref1  
 ref2  
#unknowns (Hd): 2

Last unknown is  
Unrelated

to

**Model options**

Degradation:  YES  NO  
BW Stutter:  YES  NO  
FW Stutter:  YES  NO

**Data selection**

Loci:	stain52	ref1	ref2
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH 01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

**Estimates under Hd**

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	0.719	0.035
Mix-prop. C2	0.281	0.035
P.H.expectation	699.019	71.264
P.H.variability	0.231	0.041
Degrad. slope	0.771	0.084

Maximum Likelihood value

logLik= -234.1  
adj.loglik= -238.1  
Lik= 2.16e-102

**Estimates under Hp**

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	0.724	0.028
Mix-prop. C2	0.276	0.028
P.H.expectation	688.458	59.936
P.H.variability	0.203	0.030
Degrad. slope	0.774	0.070

Maximum Likelihood value

logLik= -203.63  
adj.loglik= -207.63  
Lik= 3.65e-89

**Joint LR**

LR= 1.695e+13  
log10LR= 13.23  
Upper boundary= NA

**Non-contributor analysis**

Select reference to replace with non-contributor:  
ref1

Sample maximum based  
Sample integrated based

**LR for each locus**

VWA	58.92
D16S539	2.049
D2S1338	6649
D19S433	1817
TH 01	150.3
FGA	97.67
D3S1358	0.7917

# Kinship propositions tested for dataset 2'

## *test\_logLik2contr1RepNoStutter*

- Use Deconvolution module instead of Weight-of-Evidence
  - Considering 1 replicate (stain52 only)
  - Stutter models are turned off, degrad still turned on.
  - ref1 and ref2 are typed individuals
- 
- Hyp 1: 2 unknowns (unrelated)
  - Hyp 2: 1<sup>st</sup> unknown is unrelated + 2<sup>nd</sup> unknown is sibling of ref2
  - Hyp 3: 1<sup>st</sup> unknown is unrelated + 2<sup>nd</sup> unknown is child of ref2

# GUI check: Hyp 1

**Model specification**

**Contributor(s) under Hd:**

ref1  
 ref2

#unknowns (Hd):

Last unknown is

to

**Model options**

Degradation:  YES  NO  
 BW Stutter:  YES  NO  
 FW Stutter:  YES  NO

**Data selection**

Loci:	stain52	ref1	ref2
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH 01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

**Estimates under Hd**

**Parameter estimates:**

Param.	MLE	Std.Err.
Mix-prop. C1	0.719	0.035
Mix-prop. C2	0.281	0.035
P.H.expectation	699.019	71.264
P.H.variability	0.231	0.041
Degrad. slope	0.771	0.084

**Maximum Likelihood value**

logLik= -234.1  
 adj.loglik= -238.1  
 Lik= 2.16e-102

Locus	TopGenoty <sub>1</sub>	probal	ratioToNext	TopGenotyp	probability_C2	ratioToNextGenotype_C2
VWA	17/17	0.6267	2.804	16/18	0.6269	2.915
D16S539	10/11	0.9966	326.9	9/11	0.5201	1.493
D2S1338	19/26	0.9966	419.5	22/24	0.3149	1.08
D19S433	13.2/15.2	0.9824	72.09	13/14	0.752	3.354
TH 01	6/8	0.7842	5.67	7/9.3	0.7833	5.672
FGA	20/24	0.9687	58.35	19/20	0.5257	1.551
D3S1358	16/18	0.9682	31.58	14/15	0.5703	1.494

# GUI check Hyp 2

**Model specification**

**Contributor(s) under Hd:**

ref1  
 ref2

#unknowns (Hd): 2

Last unknown is  
 Sibling

to  
 ref2

**Model options**

Degradation:  YES  NO  
 BW Stutter:  YES  NO  
 FW Stutter:  YES  NO

**Data selection**

Loci:	stain52	ref1	ref2
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH 01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

**Estimates under Hd**

**Parameter estimates:**

Param.	MLE	Std.Err.
Mix-prop. C1	0.720	0.030
Mix-prop. C2	0.280	0.030
P.H.expectation	701.804	65.423
P.H.variability	0.213	0.033
Degrad. slope	0.771	0.076

**Maximum Likelihood value**

logLik= -226.54  
 adj.loglik= -230.54  
 Lik= 4.12e-99

Locus	TopGenot	probal	ratioTot	TopGei	probal	ratio1
VWA	17/17	0.9287	21.43	16/18	0.9288	22.37
D16S539	10/11	0.9987	883.1	9/10	0.648	2.203
D2S1338	19/26	0.9999	22620	22/24	0.8636	11.05
D19S433	13.2/15.2	0.9978	566.6	13/14	0.94	16.77
TH 01	6/8	0.9649	40.8	7/9.3	0.9646	40.81
FGA	20/24	0.9947	232.7	19/99	0.5576	1.486
D3S1358	16/18	0.9979	496.9	14/15	0.5943	1.54

# GUI check Hyp 3

**Model specification**

Contributor(s) under Hd:

ref1  
 ref2

#unknowns (Hd): 2

Last unknown is  
 Child

to  
 ref2

**Model options**

Degradation:  YES  NO  
 BW Stutter:  YES  NO  
 FW Stutter:  YES  NO

**Data selection**

Loci:	stain52	ref1	ref2
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH 01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

**Estimates under Hd**

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	0.718	0.031
Mix-prop. C2	0.282	0.031
P.H.expectation	698.023	66.171
P.H.variability	0.216	0.034
Degrad. slope	0.773	0.079

Maximum Likelihood value

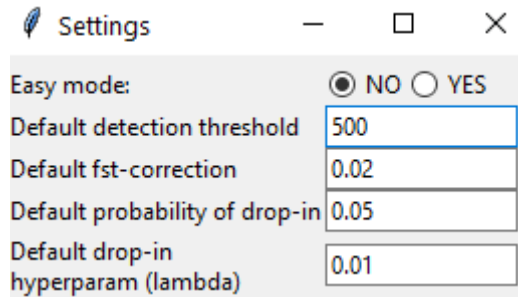
logLik= -227.44  
 adj.loglik= -231.44  
 Lik= 1.67e-99

Locus	TopGenoty <sub>1</sub>	probability_C1	ratioToNext	TopGenoty	probability	ratio1
VWA	17/17	0.8248	7.837	16/18	0.8246	8.205
D16S53!	10/11	0.9986	879.5	9/10	0.7391	3.022
D2S133!	19/26	0.9996	6376	22/24	0.4623	1.469
D19S43:	13.2/15.2	0.9952	255.1	13/14	0.892	8.972
TH 01	6/8	0.9302	20.05	7/9.3	0.9294	20.04
FGA	20/24	0.9925	138.4	19/20	0.5495	1.573
D3S135!	16/18	1	33320	14/15	0.5997	1.536

# Propositions tested for dataset 3 (SNPs):

## *test\_logLik2contr1RepSNP*

Global settings for all markers:



Easy mode:	<input checked="" type="radio"/> NO <input type="radio"/> YES
Default detection threshold	500
Default fst-correction	0.02
Default probability of drop-in	0.05
Default drop-in hyperparam (lambda)	0.01

Note: Profile P1 has missing markers  
RS13182883 and RS576261  
(substituting with additional unknown)

Only 1<sup>st</sup> replicate

- Hyp 1: P1 + 1U are contributors
- Hyp2: 1U + 2U (unrelated) unknowns are contributors, P1 is known non-contributor

# GUI check Hp: Hyp 1 vs Hd: Hyp 2

**Model specification**

Contributor(s) under Hp:  
 P1  
#unknowns (Hp): 1

Contributor(s) under Hd:  
 P1  
#unknowns (Hd): 2

Last unknown is  
Unrelated  
to

**Model options**

Degradation:  YES  NO  
BW Stutter:  YES  NO  
FW Stutter:  YES  NO

Show selected data  
Evidence(s)  
 1\_1a\_XEN13  
Show

Data selection

**Estimates under Hd**

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	0.500	0.093
Mix-prop. C2	0.500	0.093
P.H.expectation	743.964	30.688
P.H.variability	0.669	0.039

Maximum Likelihood value

logLik= -1470.4  
adj.loglik= -1473.4  
Lik= 2.6e-639

Further Action

MCMC simulation  
Deconvolution  
Model validation  
Model fitted P.H.

**Estimates under Hp**

Parameter estimates:

Param.	MLE	Std.Err.
Mix-prop. C1	0.801	0.043
Mix-prop. C2	0.199	0.043
P.H.expectation	730.588	28.490
P.H.variability	0.631	0.034

Maximum Likelihood value

logLik= -1409.45  
adj.loglik= -1412.45  
Lik= 7.64e-613

Further Action

MCMC simulation  
Deconvolution  
Model validation  
Model fitted P.H.

**Joint LR**

LR= 2.935e+26  
log10LR= 26.47  
Upper boundary= 55.35

**Non-contributor analysis**

Select reference to replace with non-contributor:  
P1

Sample maximum based  
Sample integrated based

**Further**

LR sensitivity  
Create report

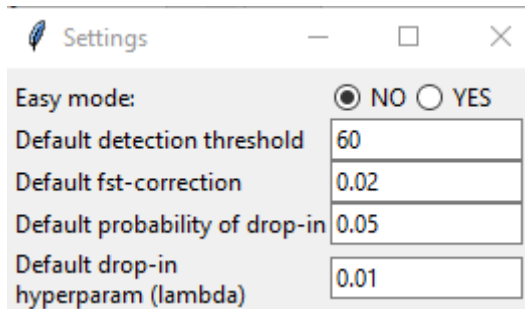


# Propositions tested for dataset 3 (SNPs):

## *test\_logLik2contr2RepSNP*

Note: Profiles P1+P3 are missing markers RS13182883 and RS576261 (substituting with additional unknown)

Global settings for all markers:



Easy mode:	<input checked="" type="radio"/> NO <input type="radio"/> YES
Default detection threshold	60
Default fst-correction	0.02
Default probability of drop-in	0.05
Default drop-in hyperparam (lambda)	0.01

Using **both** replicates

- Hyp 1: P1 + P3 are contributors
- Hyp2: P1 + 1 (unrelated) unknown are contributors, P3 is known non-contributor

# GUI check      Hp: Hyp 1 vs Hd: Hyp 2

**Model specification**

**Contributor(s) under Hp:**

P1  
 P3  
#unknowns (Hp): 0

**Contributor(s) under Hd:**

P1  
 P3  
#unknowns (Hd): 1

Last unknown is  
Unrelated  
to

**Model options**

Degradation:  YES  NO  
BW Stutter:  YES  NO  
FW Stutter:  YES  NO

**Show selected data**

**Evidence(s)**

1\_1a\_XEN13  
 1\_1b\_XEN13

Show

Data selection

Estimates under Hd			Estimates under Hp		
Parameter estimates:			Parameter estimates:		
Param.	MLE	Std.Err.	Param.	MLE	Std.Err.
Mix-prop. C1	0.666	0.025	Mix-prop. C1	0.784	0.017
Mix-prop. C2	0.334	0.025	Mix-prop. C2	0.216	0.017
P.H.expectation	815.870	17.955	P.H.expectation	797.860	20.221
P.H.variability	0.513	0.016	P.H.variability	0.591	0.018
Maximum Likelihood value			Maximum Likelihood value		
logLik=	-3807.18		logLik=	-3980.05	
adj.loglik=	-3810.18		adj.loglik=	-3983.05	
Lik=	3.66e-1654		Lik=	3.05e-1729	

# Additional: Testing the **qualitative** model (tested for dataset 2, using both replicates)

Different drop-out probability models are tested:

- Hyp 1: Ref1 + Ref2, common drop-out parameter
- Hyp 2: Ref1 + Ref2, Ref1 drop-out parameter set to 0.
- Hyp 3: Ref1 + Ref2, 2 drop-out parameters

Ref2 known non-contr for hyps Hyp '4-6':

- Hyp 4: Ref1 + 1U, common drop-out parameter.
- Hyp 5: Ref1 + 1U, Ref1 drop-out parameter set to 0.
- Hyp 6: Ref1 + 1U, 2 drop-out parameters

# GUI check

Only common drop-out parameter implemented in GUI

Model specification

Contributor(s) under Hp:

ref1  
 ref2

#unknowns (Hp): 0

Contributor(s) under Hd:

ref1  
 ref2

#unknowns (Hd): 1

1st unknown is  
Unrelated

Confirm

Results of MLE based LR:  
LR=2.62e-08  
log10LR=-7.582  
Estimated pD under Hp=0.07301  
Estimated pD under Hd=0.0381

Do you want to export the results?

Yes No

Hyp 1 (Hp) : R1+R2 (common drop-out parameter):  
logLik = -95.1168543762792  
d = 0.0730126, 0.0730126

Hyp 4 (Hd): R1+1U (common drop-out parameter):  
logLik = -77.6593191143117  
d = 0.03809141, 0.03809141

log10LR = -7.581711

# Additional validation of EFM GUI

- The GUI is manually inspected for concise results whenever new versions are released – based on following data:
  - The output from the manual (tutorial data)
    - Also covers the Bayesian approach
  - Workshop material:
    - ENFSI exercises (ENFSI1 and ENFSI2)
    - EFMexercise26