

# EuroForMix

A user-friendly software for evaluating STR/SNP profiles using peak height information

Øyvind Bleka, MSc  
Oyvind.Bleka@fhi.no

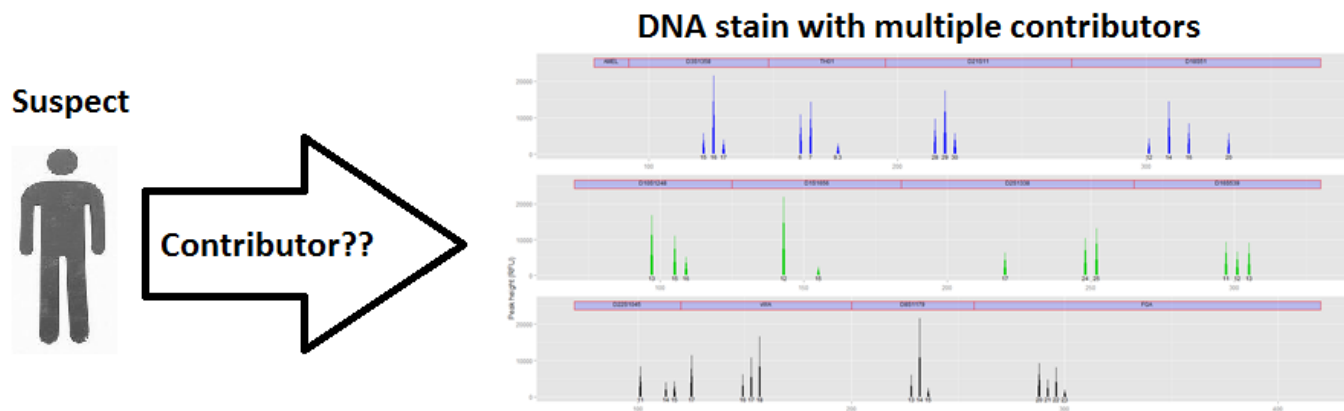


UNIVERSITY  
OF OSLO



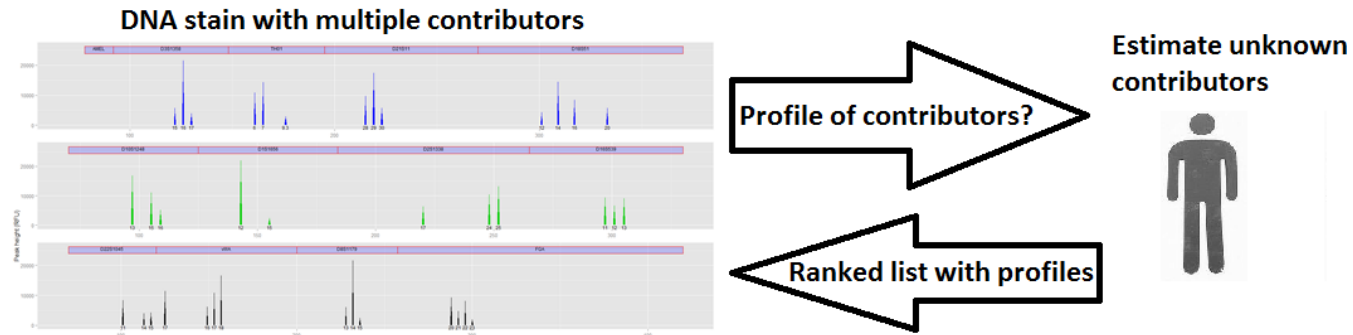
Norwegian Institute of Public Health

- A Graphical User Interface which implements and extends the continuous model from Cowell et.al (2015).
- Parameters for mixture proportion, peak height distribution, stutter proportion and degradation are automatically taken into account.
- No need for calibration, but prior information can be specified.
- Weight-of-evidence (WoE) of an obtained crime sample now uses peak height information!

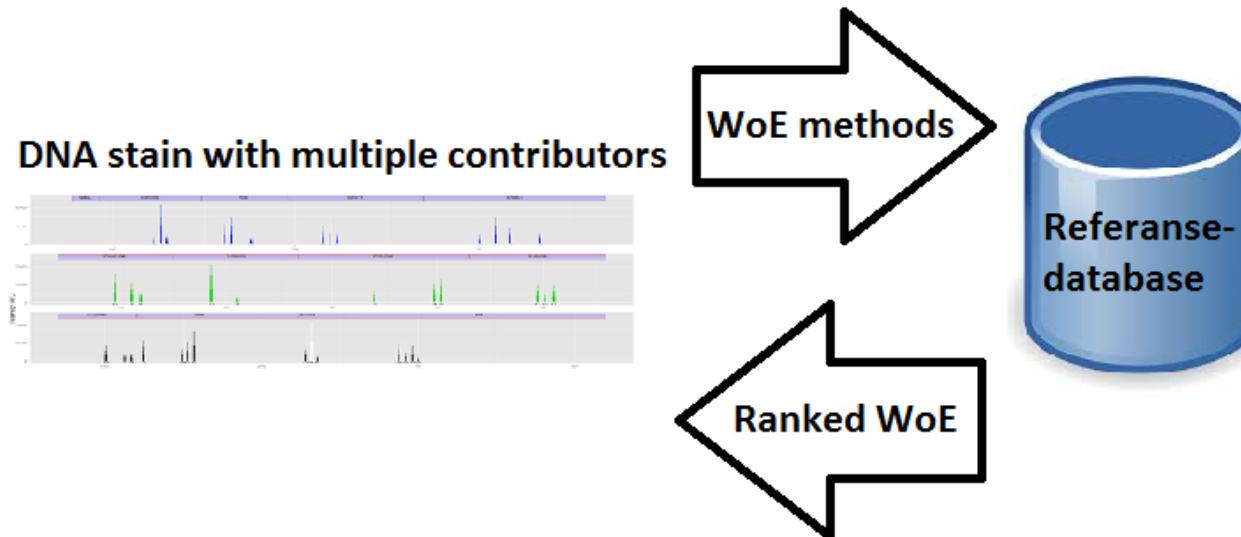


- The continuous model in EuroForMix supports:
  - Multiple contributors in hypothesis
    - Can condition on any number of reference profiles
    - Can specify any number of unknowns (practical limit is 4)
  - Replicated samples
    - No need for making a consensus sample
  - Stutters
  - Allele drop-out
  - Allele drop-in with a peak height model
  - Sub-population effect (coancestry coefficient  $F_{st}$ )
  - Degradation of peak heights over fragment length

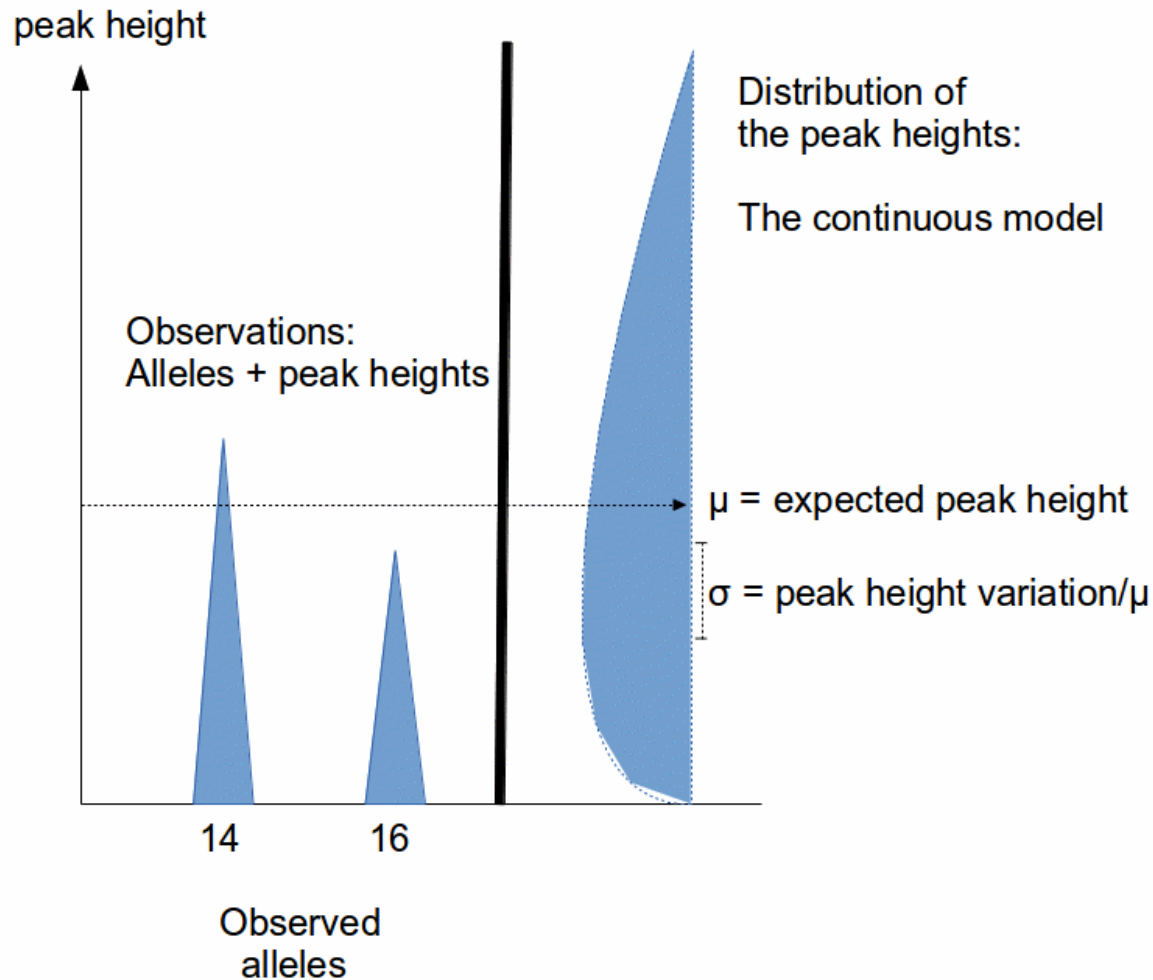
## Deconvolution:



## Database search:



# The continuous model: A gamma distribution



To obtain  $P(E|H)$ , the probability of observed sample  $E$  given hypothesis  $H$ , an inference approach must be applied

- EuroForMix supports two approaches:

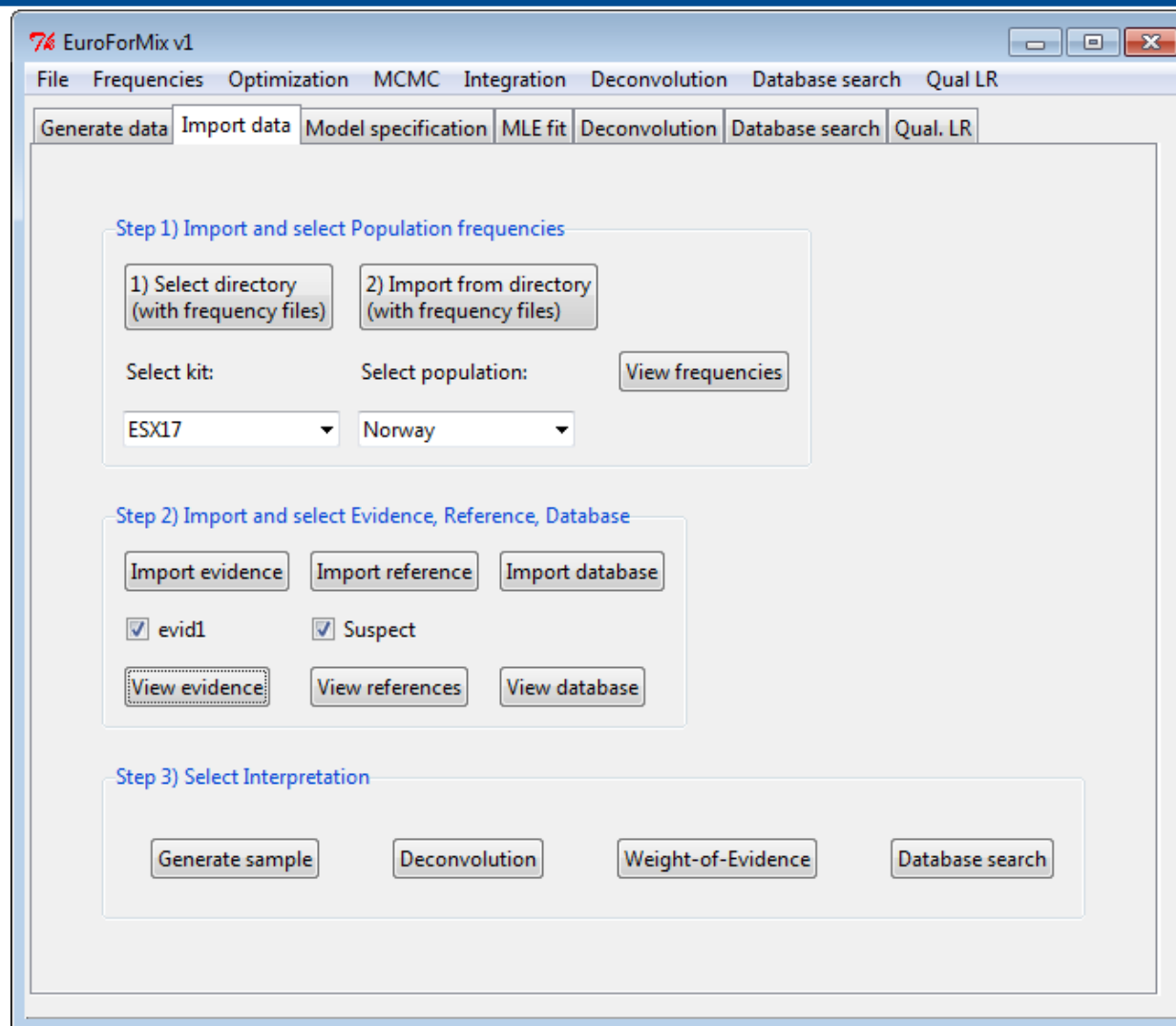
Approach 1) Maximum likelihood estimation

$P(E|H)$  estimated with  $\max_{\theta} p(E|H, \theta)$

Approach 2) Bayesian (integrates out model parameters)

$P(E|H)$  estimated with  $\int_{\theta} p(E|H, \theta)p(\theta)d\theta$

# The GUI: Import

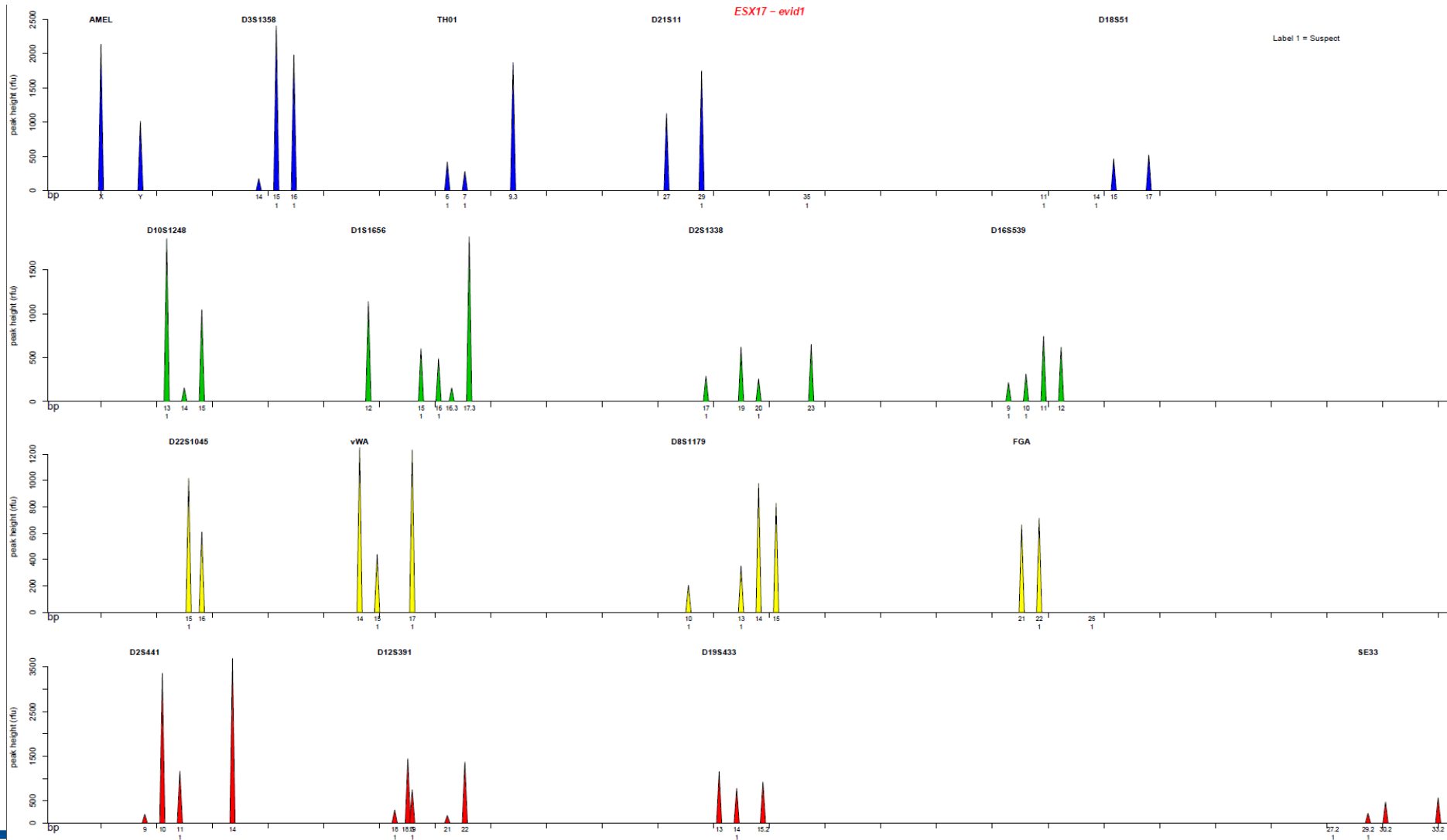


The screenshot shows the EuroForMix v1 software interface. The window title is "76 EuroForMix v1". The menu bar includes "File", "Frequencies", "Optimization", "MCMC", "Integration", "Deconvolution", "Database search", and "Qual LR". The main menu has tabs for "Generate data", "Import data", "Model specification", "MLE fit", "Deconvolution", "Database search", and "Qual LR".

The "Import data" tab is active, showing three steps:

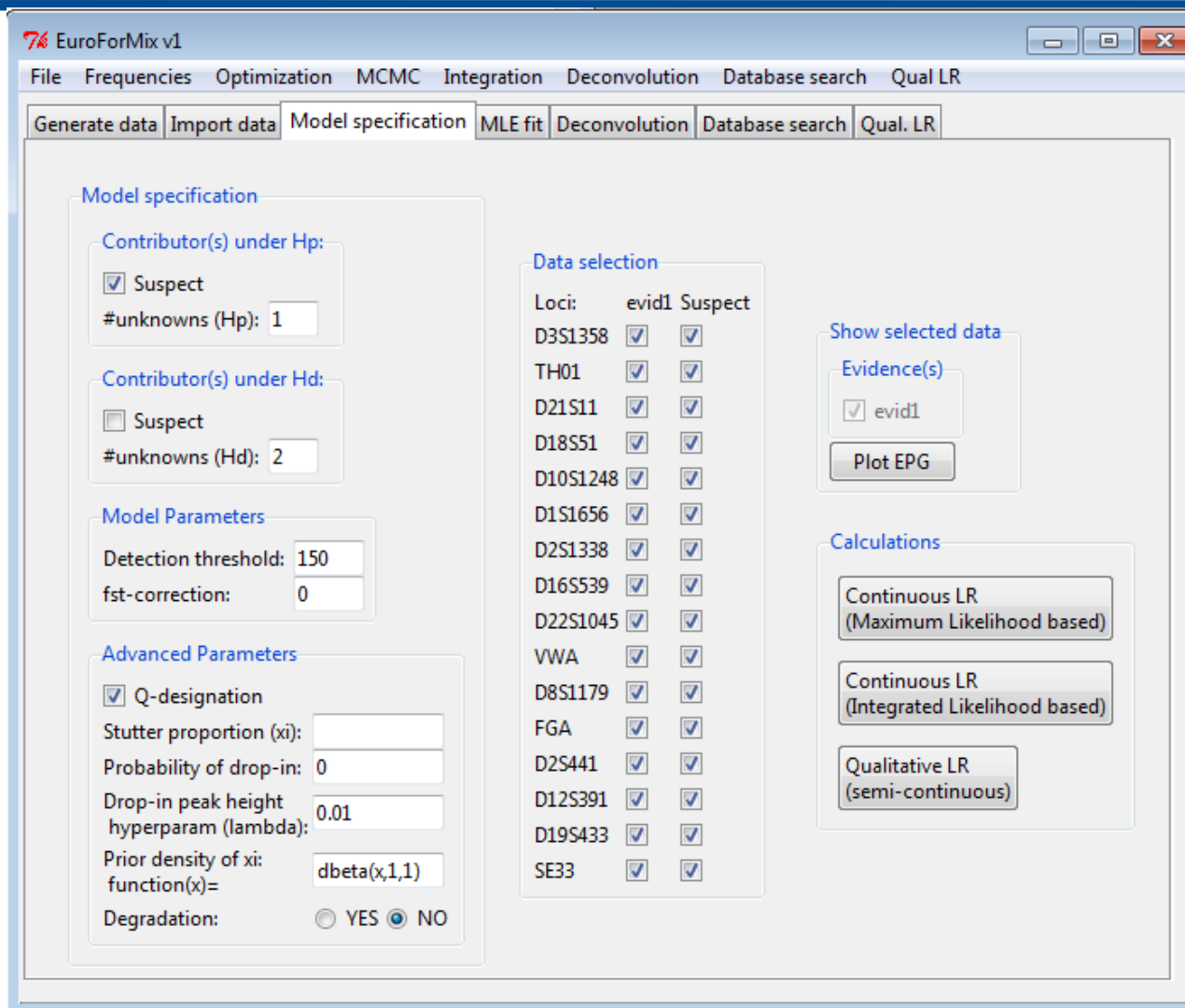
- Step 1) Import and select Population frequencies**
  - Buttons: "1) Select directory (with frequency files)" and "2) Import from directory (with frequency files)".
  - Fields: "Select kit:" with a dropdown menu showing "ESX17", and "Select population:" with a dropdown menu showing "Norway".
  - Button: "View frequencies".
- Step 2) Import and select Evidence, Reference, Database**
  - Buttons: "Import evidence", "Import reference", and "Import database".
  - Checkboxes:  "evid1" and  "Suspect".
  - Buttons: "View evidence", "View references", and "View database".
- Step 3) Select Interpretation**
  - Buttons: "Generate sample", "Deconvolution", "Weight-of-Evidence", and "Database search".

# The GUI: View data





# The GUI: Specification of the model



The screenshot displays the EuroForMix v1 software interface. The main window is titled "EuroForMix v1" and features a menu bar with options: File, Frequencies, Optimization, MCMC, Integration, Deconvolution, Database search, and Qual LR. Below the menu bar is a toolbar with buttons for "Generate data", "Import data", "Model specification", "MLE fit", "Deconvolution", "Database search", and "Qual. LR".

The "Model specification" panel is active and contains several sections:

- Contributor(s) under Hp:** Includes a checked checkbox for "Suspect" and a text input field for "#unknowns (Hp):" with the value "1".
- Contributor(s) under Hd:** Includes an unchecked checkbox for "Suspect" and a text input field for "#unknowns (Hd):" with the value "2".
- Model Parameters:** Includes a text input field for "Detection threshold:" with the value "150" and another for "fst-correction:" with the value "0".
- Advanced Parameters:** Includes a checked checkbox for "Q-designation", a text input field for "Stutter proportion (xi):", a text input field for "Probability of drop-in:" with the value "0", a text input field for "Drop-in peak height hyperparam (lambda):" with the value "0,01", a text input field for "Prior density of xi: function(x)=" with the value "dbeta(x,1,1)", and a radio button selection for "Degradation:" with "NO" selected.

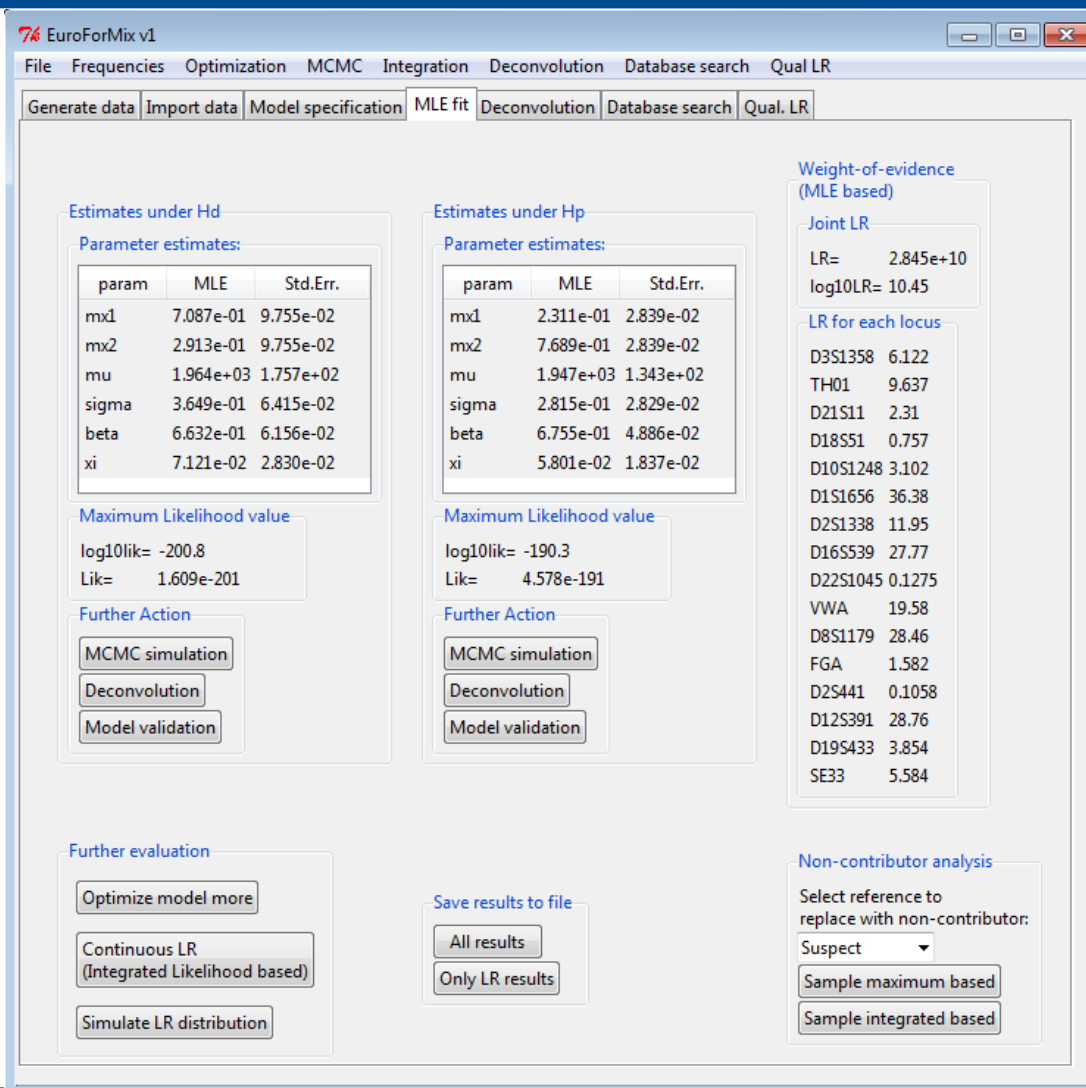
The "Data selection" panel shows a table of loci and their selection status for "evid1" and "Suspect":

Loci:	evid1	Suspect
D3S1358	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TH01	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D21S11	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D18S51	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D10S1248	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D1S1656	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S1338	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D16S539	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D22S1045	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
VWA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D8S1179	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FGA	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D2S441	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D12S391	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D19S433	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
SE33	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Additional panels include:

- Show selected data:** A section titled "Evidence(s)" with a checked checkbox for "evid1" and a "Plot EPG" button.
- Calculations:** Three buttons for "Continuous LR (Maximum Likelihood based)", "Continuous LR (Integrated Likelihood based)", and "Qualitative LR (semi-continuous)".

# The GUI: Maximum Likelihood estimates



**Estimates under Hd**

Parameter estimates:

param	MLE	Std.Err.
mx1	7.087e-01	9.755e-02
mx2	2.913e-01	9.755e-02
mu	1.964e+03	1.757e+02
sigma	3.649e-01	6.415e-02
beta	6.632e-01	6.156e-02
xi	7.121e-02	2.830e-02

Maximum Likelihood value  
log10lik= -200.8  
Lik= 1.609e-201

Further Action

MCMC simulation  
Deconvolution  
Model validation

Further evaluation

Optimize model more  
Continuous LR (Integrated Likelihood based)  
Simulate LR distribution

**Estimates under Hp**

Parameter estimates:

param	MLE	Std.Err.
mx1	2.311e-01	2.839e-02
mx2	7.689e-01	2.839e-02
mu	1.947e+03	1.343e+02
sigma	2.815e-01	2.829e-02
beta	6.755e-01	4.886e-02
xi	5.801e-02	1.837e-02

Maximum Likelihood value  
log10lik= -190.3  
Lik= 4.578e-191

Further Action

MCMC simulation  
Deconvolution  
Model validation

Save results to file

All results  
Only LR results

**Weight-of-evidence (MLE based)**

Joint LR  
LR= 2.845e+10  
log10LR= 10.45

LR for each locus

D3S1358	6.122
TH01	9.637
D21S11	2.31
D18S51	0.757
D10S1248	3.102
D1S1656	36.38
D2S1338	11.95
D16S539	27.77
D22S1045	0.1275
VWA	19.58
D8S1179	28.46
FGA	1.582
D2S441	0.1058
D12S391	28.76
D19S433	3.854
SE33	5.584

Non-contributor analysis

Select reference to replace with non-contributor:  
Suspect  
Sample maximum based  
Sample integrated based

Requiring 3 optimizations:  
Calculates MLE ~12 seconds

Can do further analysis from here:

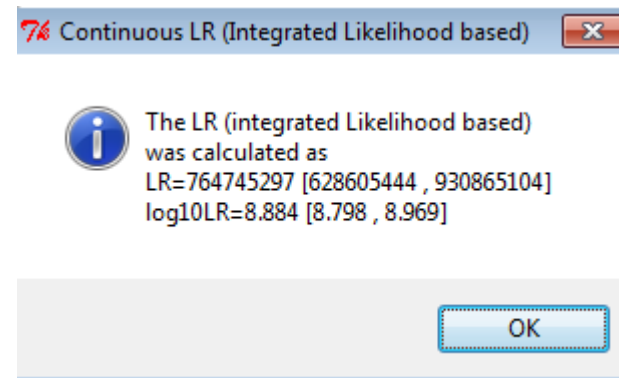
- MCMC simulation of the posterior distribution of model parameters.
- Model validation
- Deconvolution

# The GUI: Integrated Likelihood (Bayesian)

**Using relative error 0.1:**

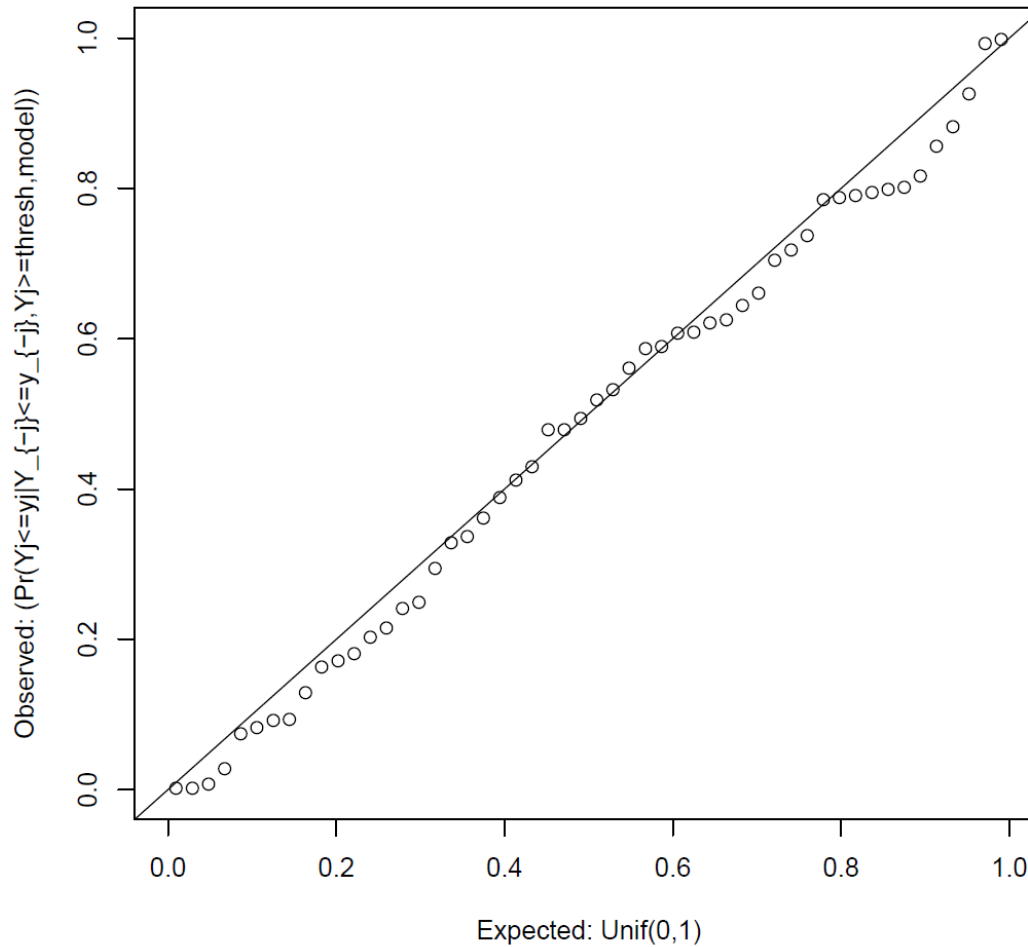
Uses 2:27min receiving

$\log_{10}LR = 8.884 [8.798, 8.969]$



# Plot from GUI: Model validation

PP-plot between fitted model and theoretical model



# The GUI: Deconvolution

EuroForMix v1

File Frequencies Optimization MCMC Integration Deconvolution Database search Qual LR

Generate data Import data Model specification MLE fit Deconvolution Database search Qual LR

rank	D3S135	TH01_g	D21S11	D18S51	D10S1248	D1S1656	D2S1338	D16S539	D22S1045	VWA_g2	D8S1179	FGA_g2	D2S441	D12S391	D19S433	SE33_g2	posterior
1	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.972
2	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/99	10/14	18.3/22	13/15.2	30.2/33.2	0.0114
3	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/21	10/14	18.3/22	13/15.2	30.2/33.2	0.00697
4	15/16	6/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.00153
5	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	22/22	10/14	18.3/22	13/15.2	30.2/33.2	0.00142
6	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	22/99	10/14	18.3/22	13/15.2	30.2/33.2	0.0012
7	15/15	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.000764
8	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/22	10/14	18.3/22	14/15.2	30.2/33.2	0.000719
9	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	12/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.00067
10	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	16/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.000452
11	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/24	10/14	18.3/22	13/15.2	30.2/33.2	0.00037
12	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	20/23	11/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.000328
13	15/16	9.3/9.3	27/29	16/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.000225
14	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/14	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.000194
15	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	20/21	10/14	18.3/22	13/15.2	30.2/33.2	0.00019
16	15/16	7/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.000147
17	15/16	9.3/9.3	27/29	15/17	15/15	12/17.3	19/23	11/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.000136
18	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	10/12	15/16	14/17	14/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	0.000102
19	15/16	9.3/9.3	27/29	15/17	13/15	12/17.3	19/23	11/12	15/16	14/17	15/15	21/22	10/14	18.3/22	13/15.2	30.2/33.2	9.03e-05

- EuroForMix is open-source and freely available through the R-package **euroformix** which is downloadable from R-forge.
- Homepage:
  - [www.euroformix.com](http://www.euroformix.com)
  - Here is installation instruction, tutorial and manual.