LMMNORM: A PACKAGE FOR THE NORMALIZATION OF MICROARRAYS USING LINEAR MIXED MODELS

Philippe Haldermans

I-Biostat, Center for Statistics, Hasselt University

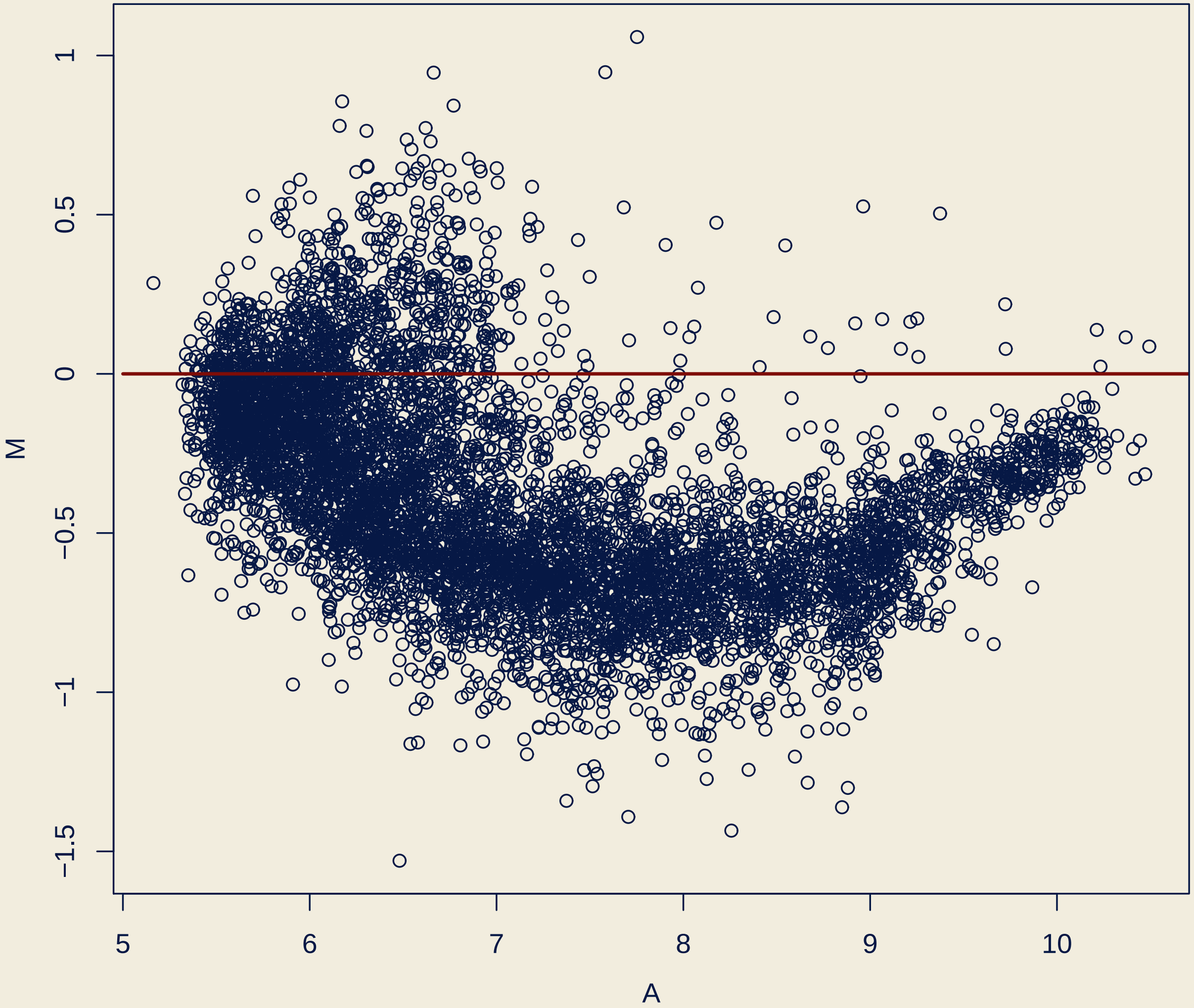
July 8-10, 2009

Normalization: our definition

Normalization: Adjust microarray data for systematic effects which arise from variation in the technology rather than from biological differences between the RNA samples.

Normalization: our definition

- MA-plot



Different models

- Global normalization
  - $M = \beta_0 + \varepsilon$
- Linear normalization
  - $M = \beta_0 + \beta_1 A + \varepsilon$
- Nonlinear normalization
  - $M(A_i) = \beta_0 + \beta_1 A_i + \sum_{k=1}^{K} b_k (A_i - t_k)_+ + \varepsilon_i$

Different models

- Nonlinear normalization with replicated genes
  - $M = X\beta + Z_\ell b_\ell + u_i + \varepsilon$
- Nonlinear normalization with different smoothers per print-tip
  - $M = X\beta + Z^* b^* + \varepsilon$

$$\boldsymbol{Z}^* = \begin{bmatrix} \boldsymbol{Z}_1 & 0 & 0 & \dots & 0 \\ 0 & \boldsymbol{Z}_2 & 0 & \dots & 0 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & \boldsymbol{Z}_L \end{bmatrix}. \tag{1}$$

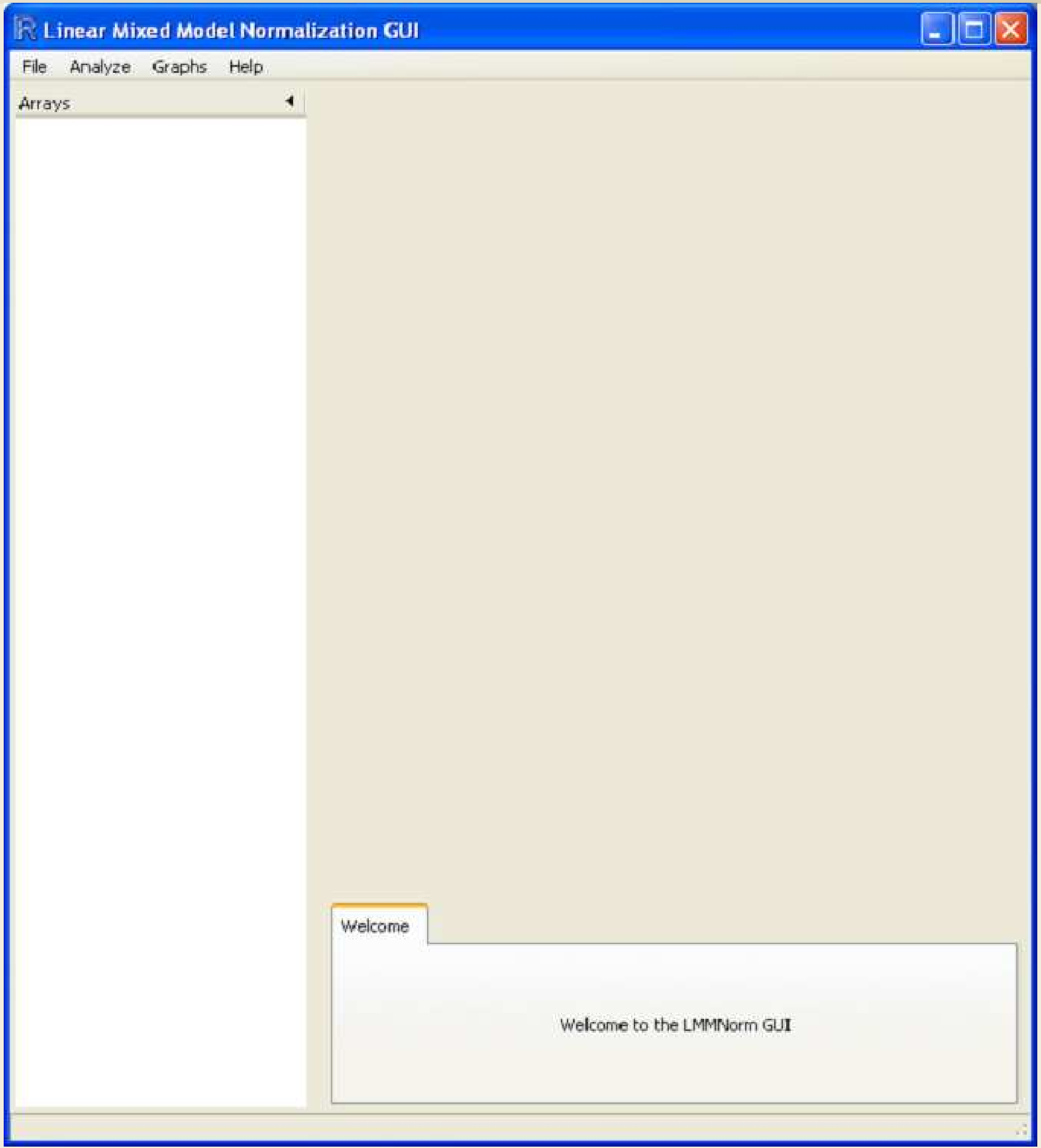
Different models

- Compare the different models and choose the best one based on:
  - AIC
  - BIC
  - LRT
  - ...

LMMNorm package

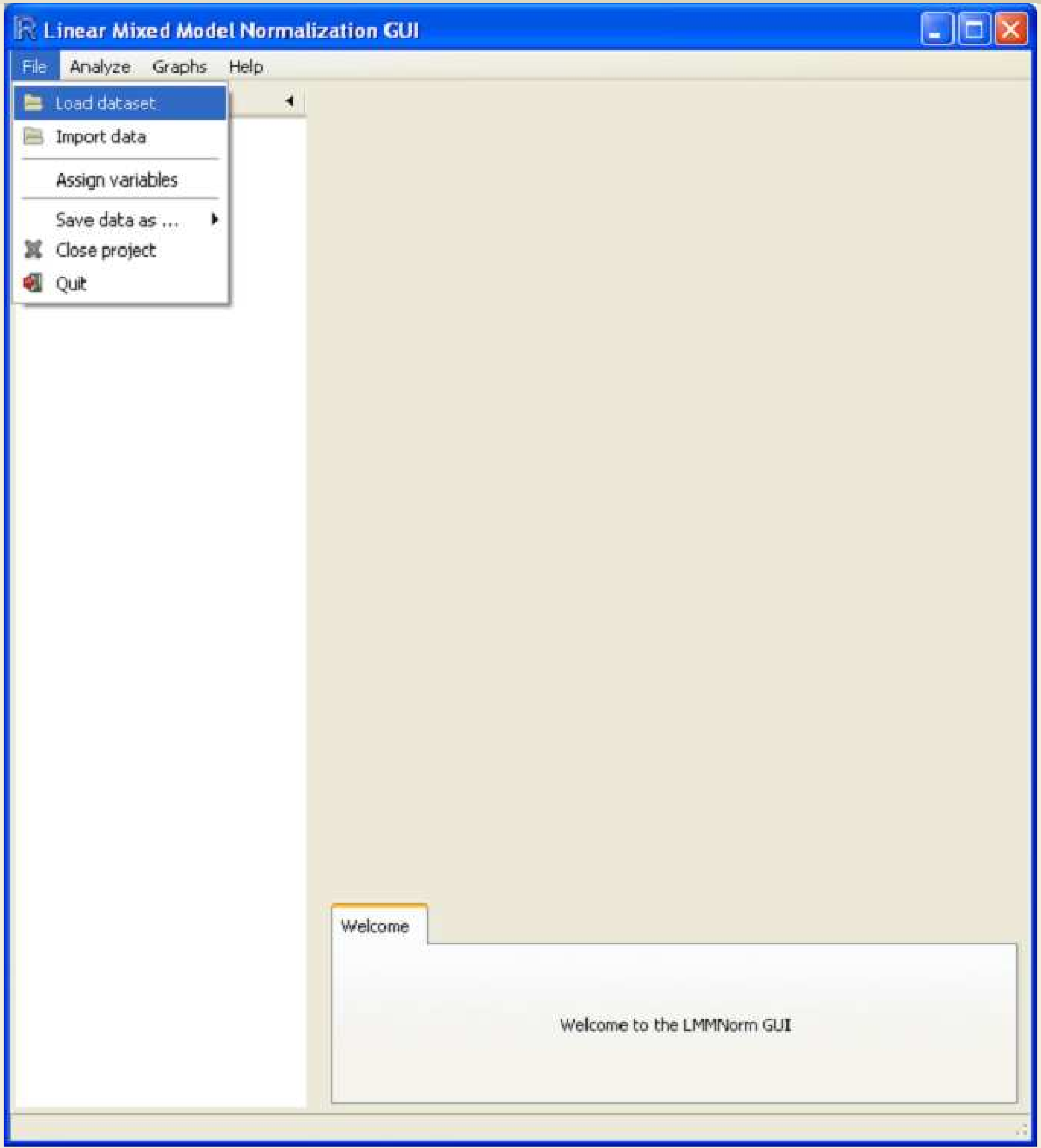
- Command-line functions
  - mixednorm(log2R, log2G, geneID, printTip, row, column, proc = c("global", "linear", "nonlinear", "repl", "fixed.pin", "diff.pin", "diffpin.rep"))
- Graphical User Interface
  - LMMNormGUI()

LMMNormGUI: a demo



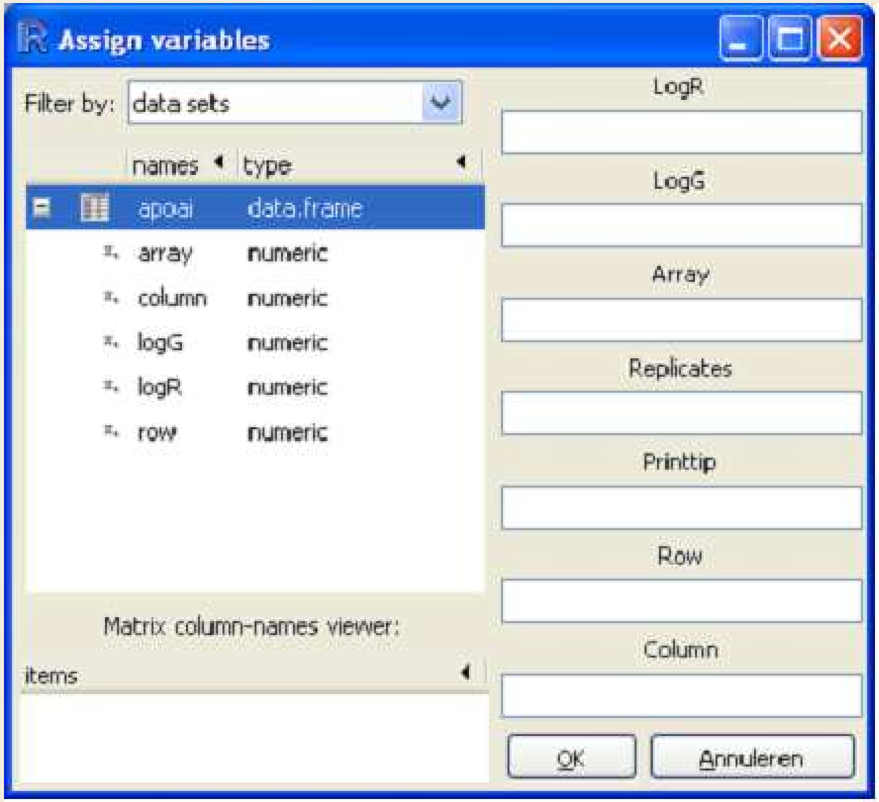
Welcome to the LMMNorm GUI

LMMNormGUI: a demo



Welcome to the LMMNorm GUI

LMMNormGUI: a demo



Filter by: data sets

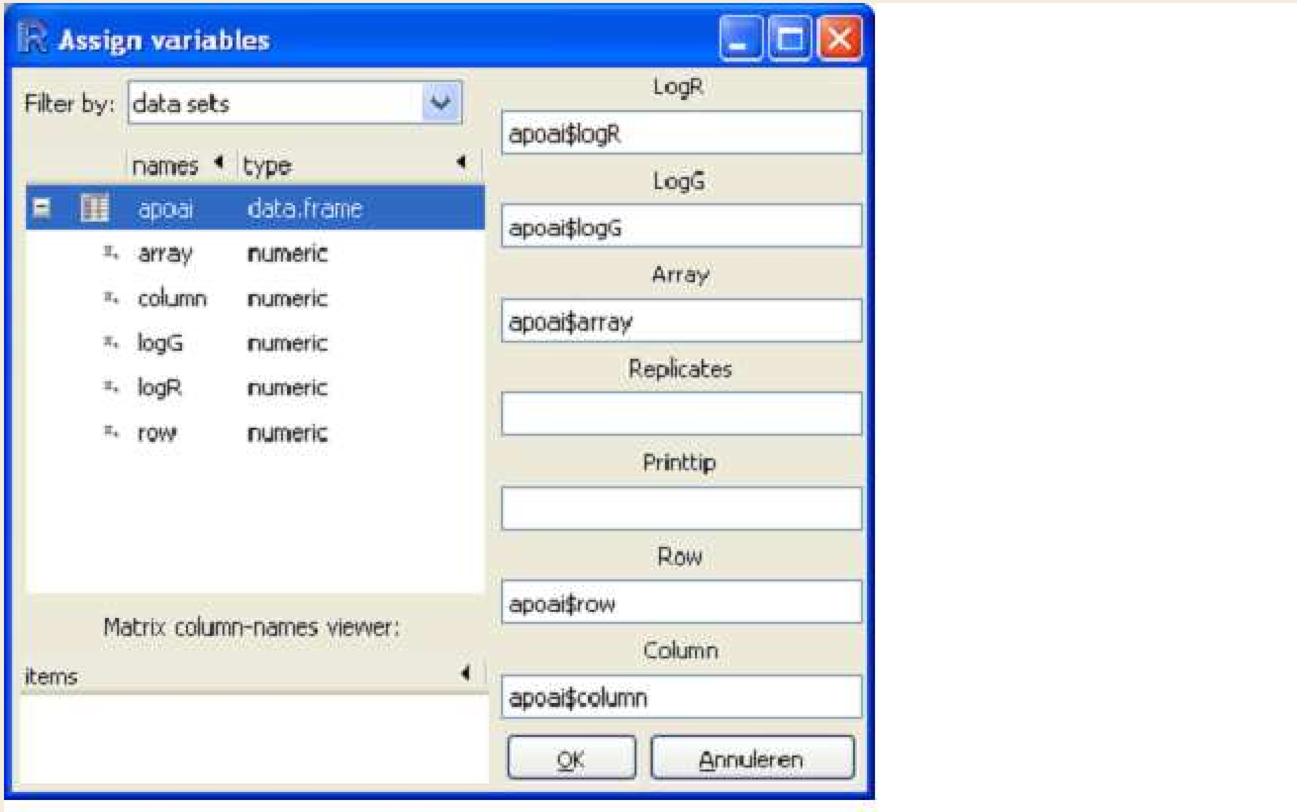
names	type
apoai	data.frame
array	numeric
column	numeric
logG	numeric
logR	numeric
row	numeric

Matrix column-names viewer:

items

LogR
LogG
Array
Replicates
Printtip
Row
Column

LMMNormGUI: a demo



Filter by: data sets

names	type
apoai	data.frame
array	numeric
column	numeric
logG	numeric
logR	numeric
row	numeric

LogR: apoai$logR

LogG: apoai$logG

Array: apoai$array

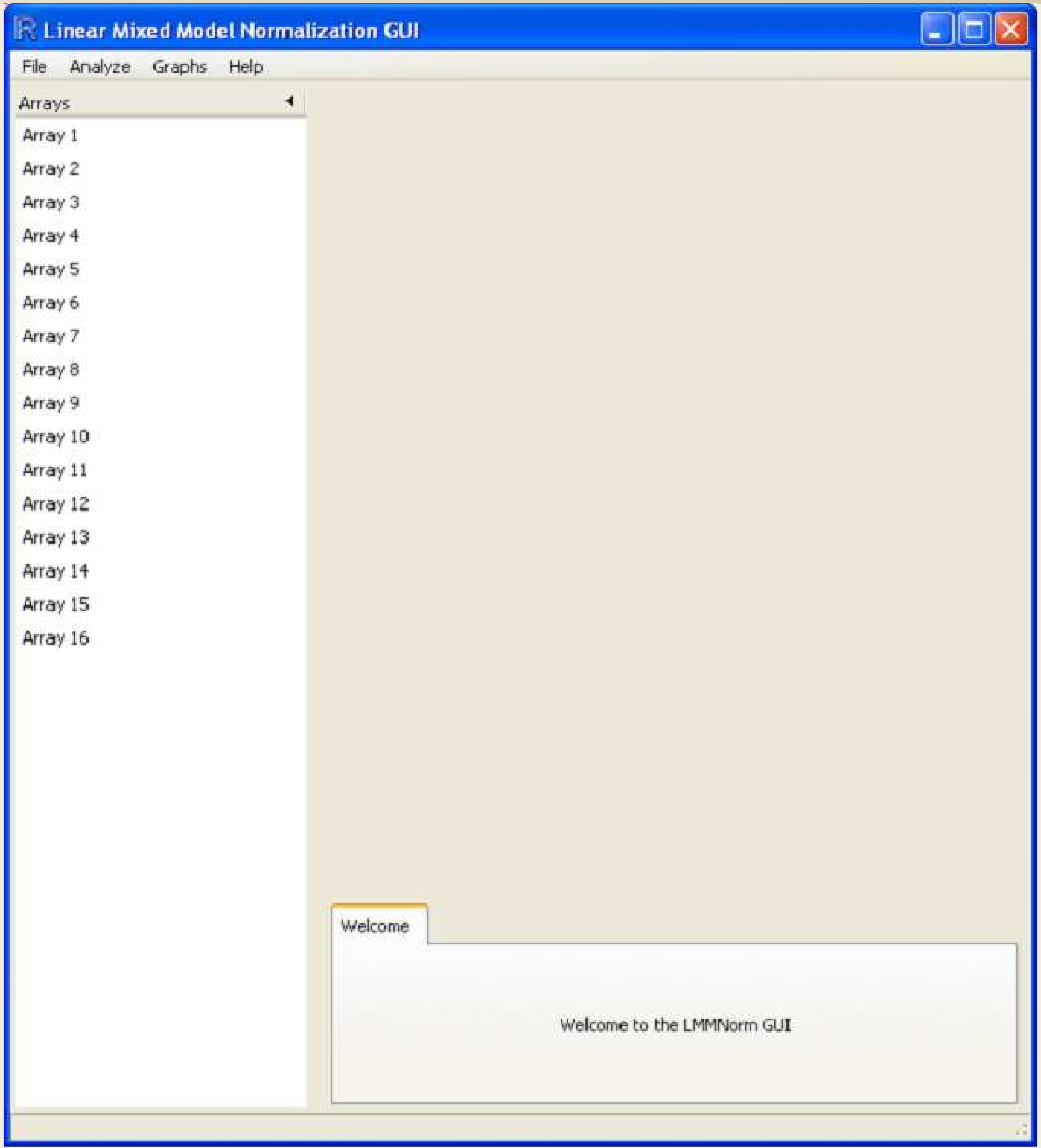
Replicates:

Printtip:

Row: apoai$row

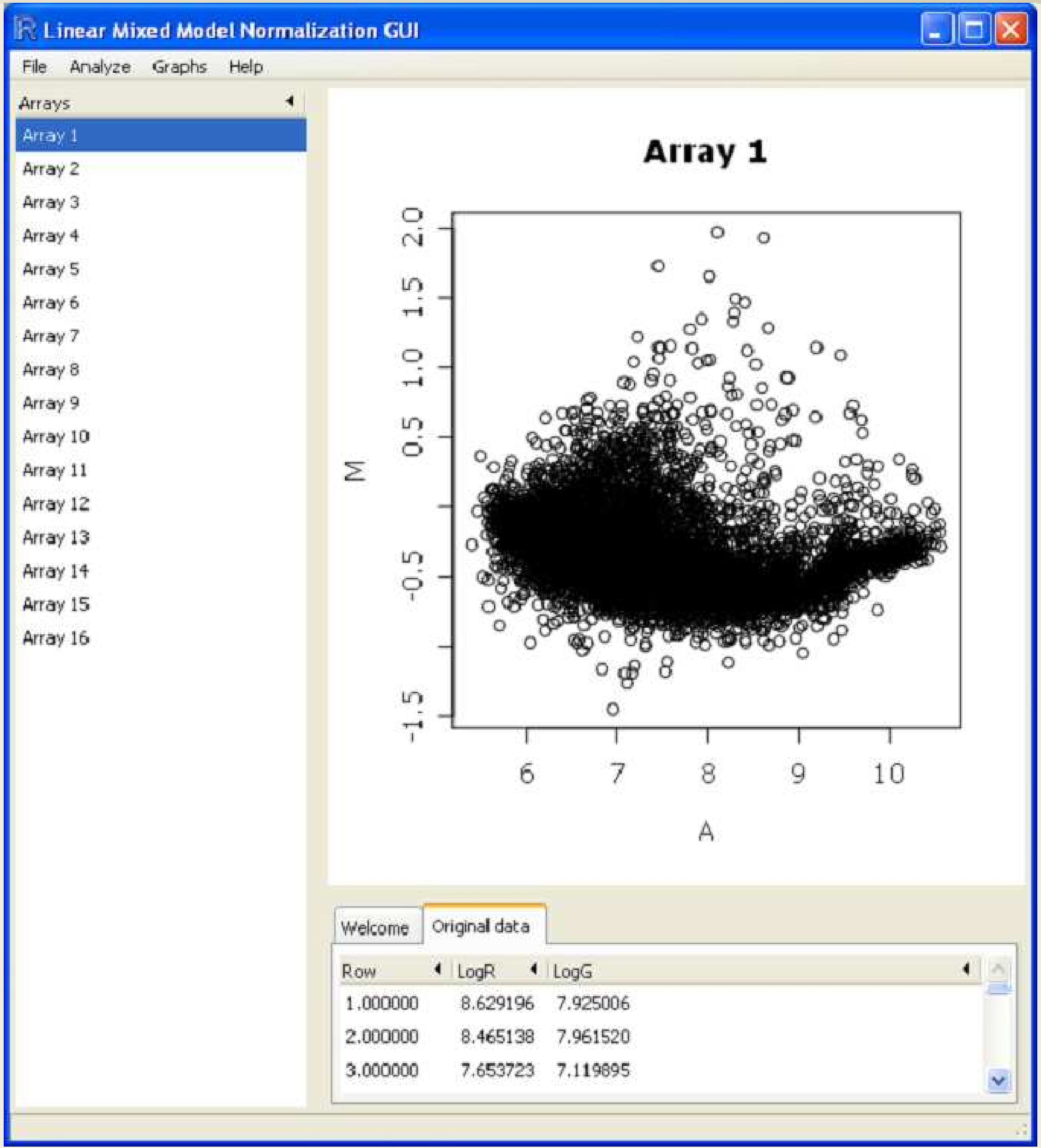
Column: apoai$column

LMMNormGUI: a demo



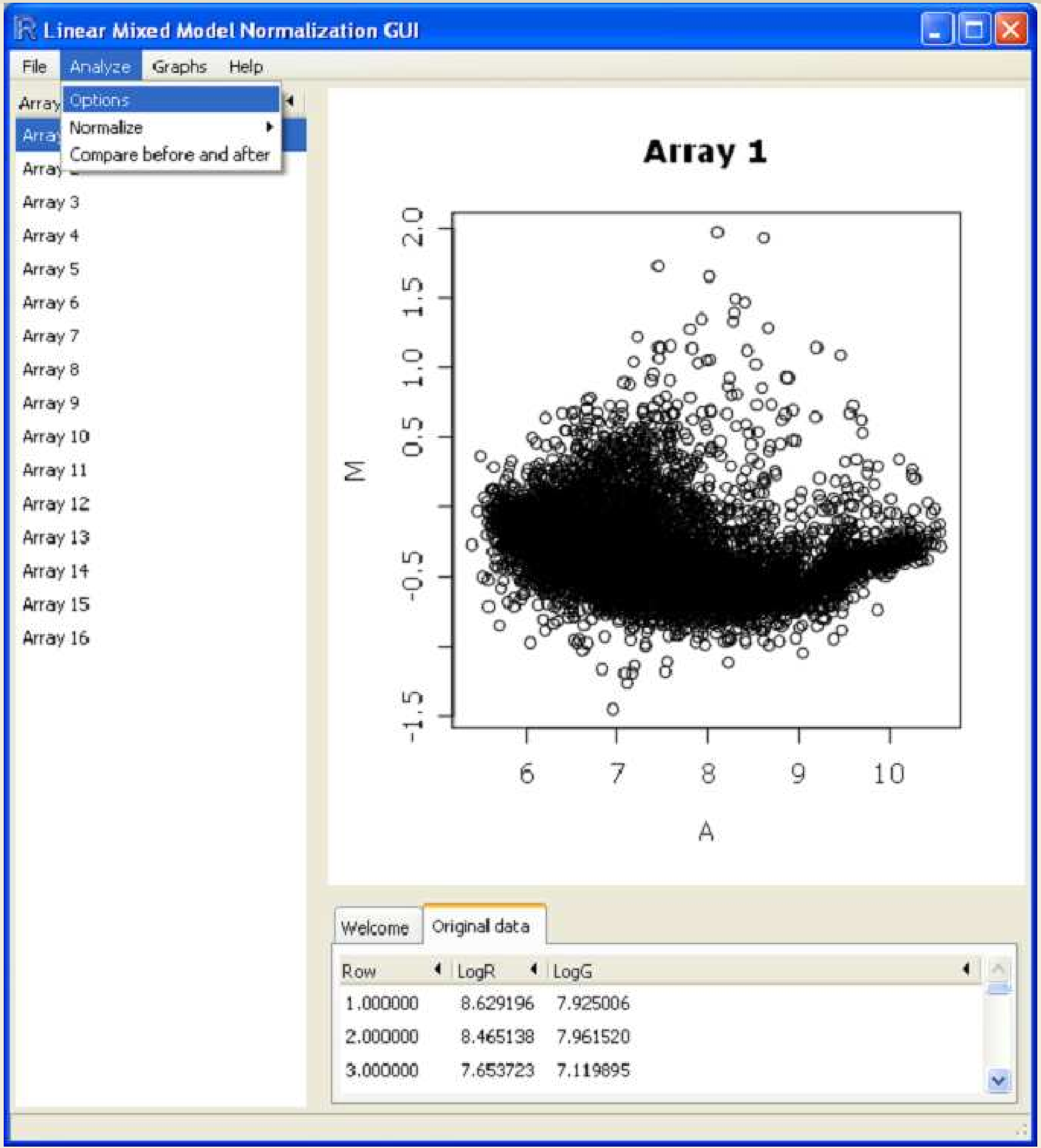
Welcome to the LMMNorm GUI

LMMNormGUI: a demo



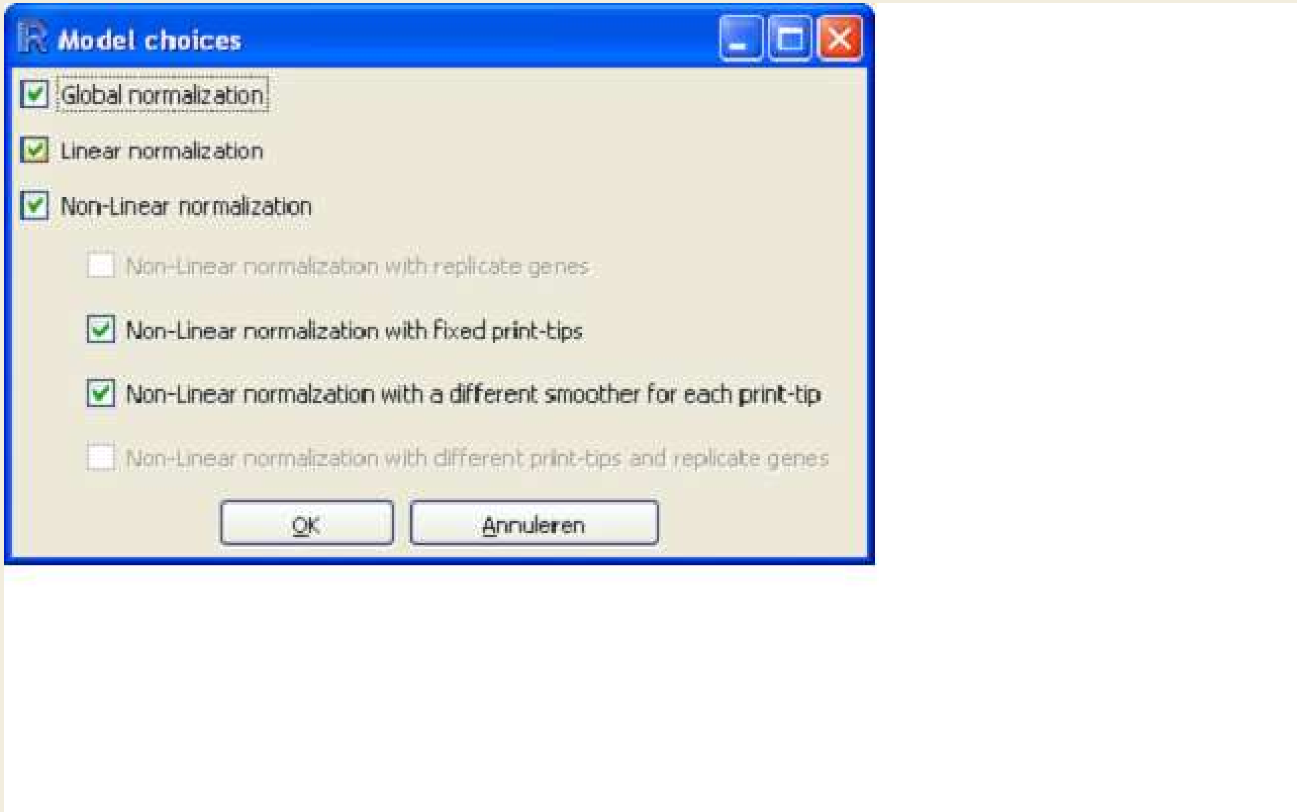
Row	LogR	LogG
1.000000	8.629196	7.925006
2.000000	8.465138	7.961520
3.000000	7.653723	7.119895

LMMNormGUI: a demo



Row	LogR	LogG
1.000000	8.629196	7.925006
2.000000	8.465138	7.961520
3.000000	7.653723	7.119895

LMMNormGUI: a demo



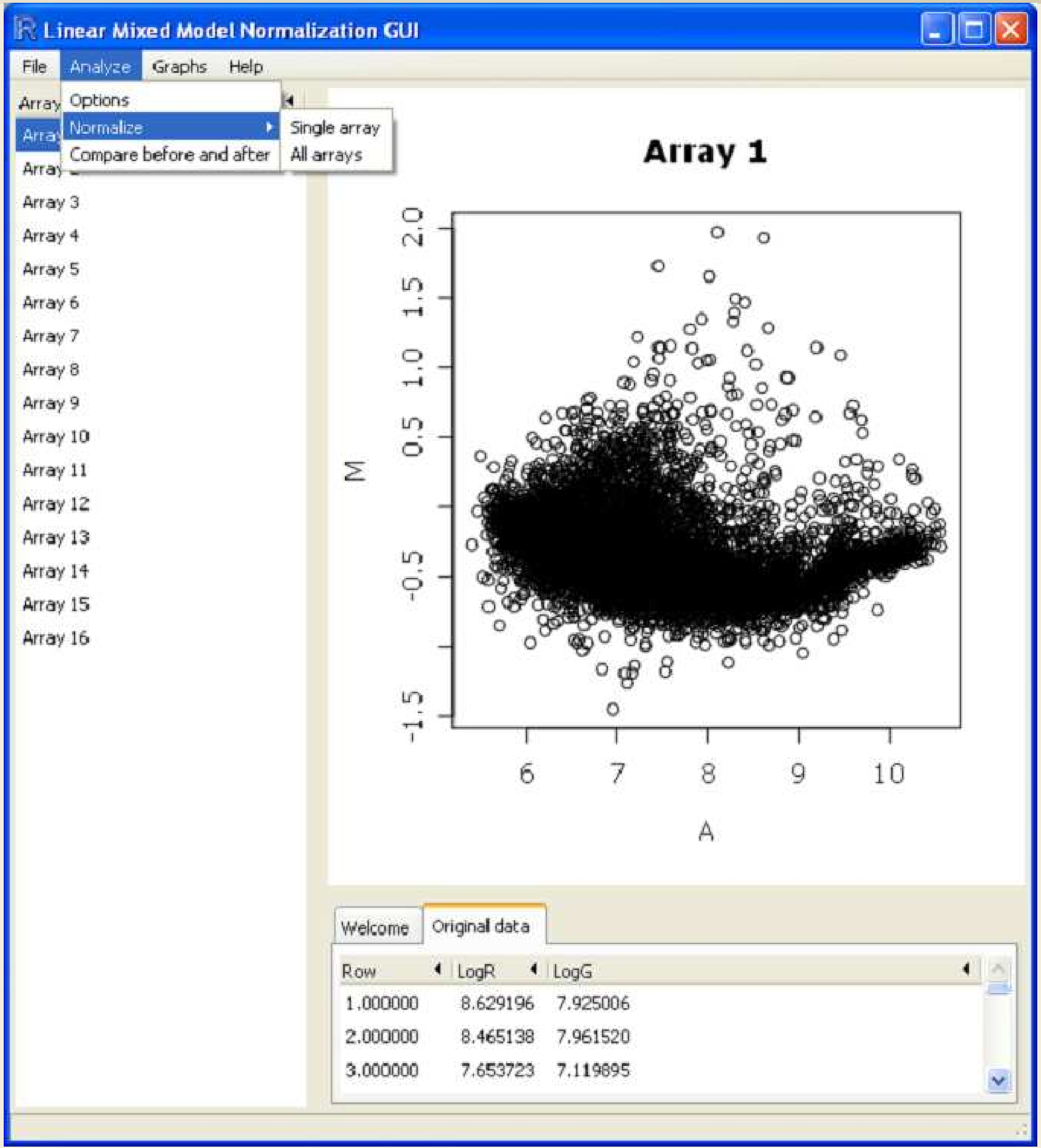
☑ Global normalization

☑ Linear normalization

☑ Non-Linear normalization

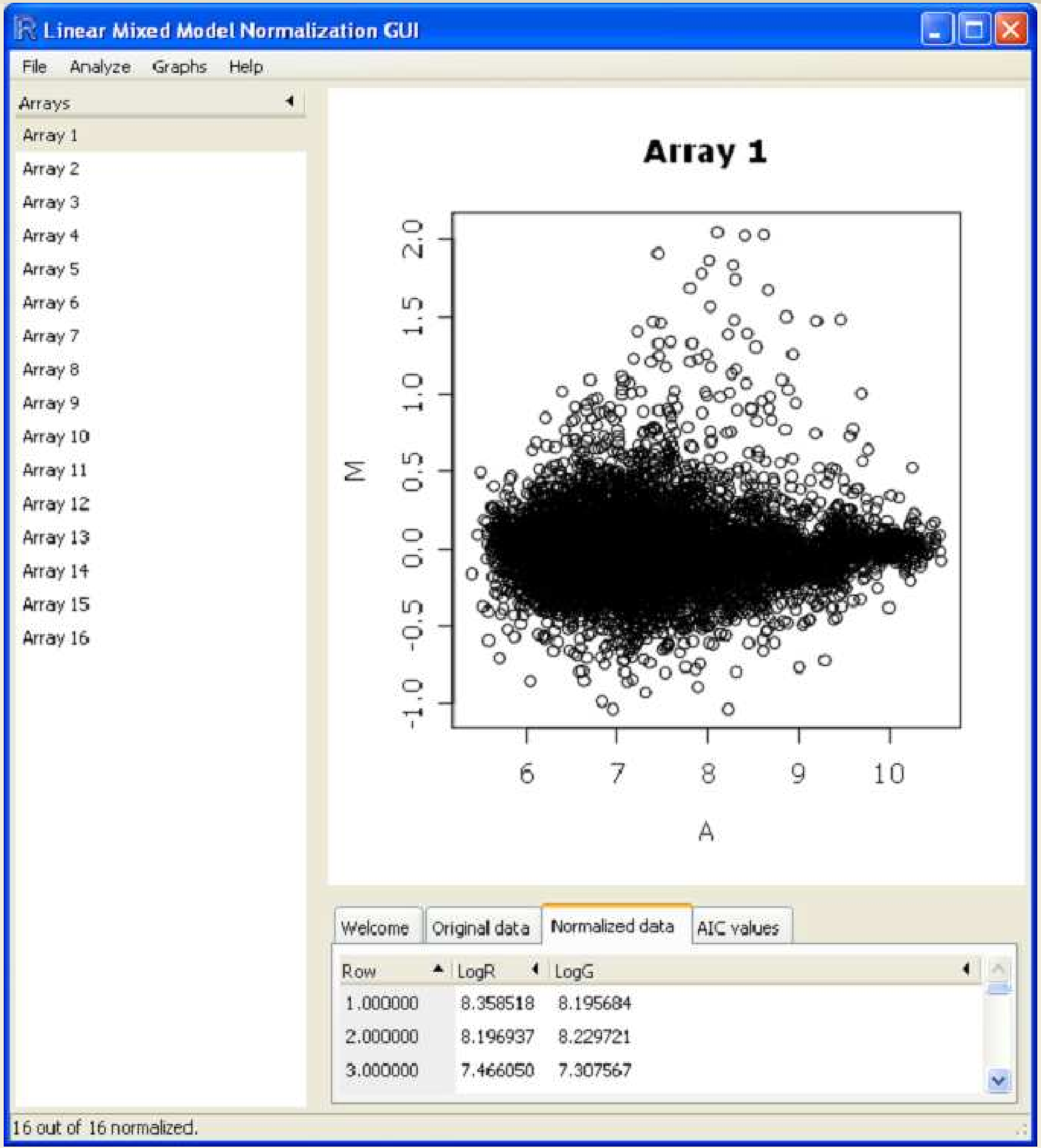
- ☐ Non-Linear normalization with replicate genes
- ☑ Non-Linear normalization with fixed print-tips
- ☑ Non-Linear normalzation with a different smoother for each print-tip
- ☐ Non-Linear normalization with different print-tips and replicate genes

LMMNormGUI: a demo



Row	LogR	LogG
1.000000	8.629196	7.925006
2.000000	8.465138	7.961520
3.000000	7.653723	7.119895

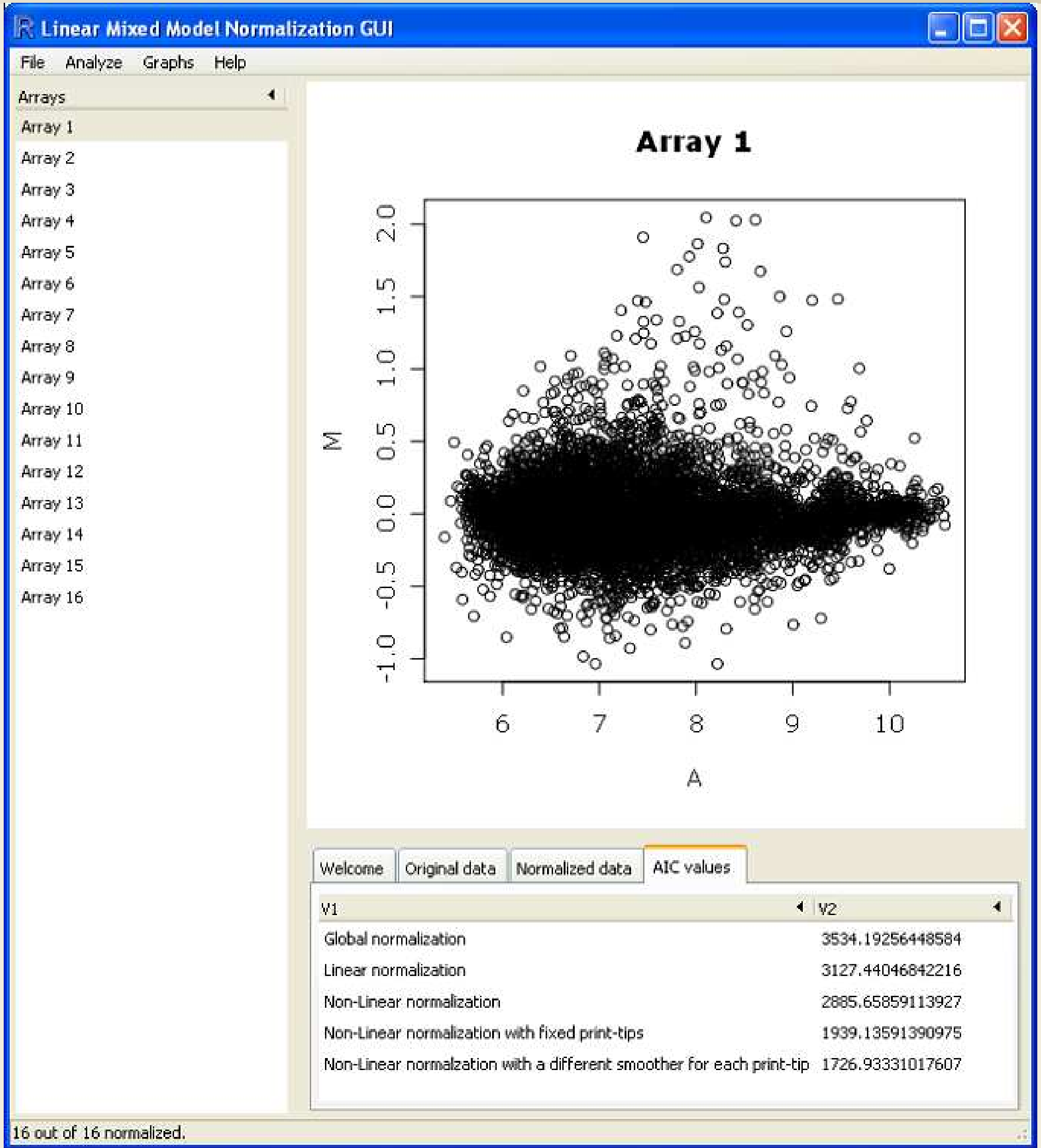
LMMNormGUI: a demo



Row	LogR	LogG
1.000000	8.358518	8.195684
2.000000	8.196937	8.229721
3.000000	7.466050	7.307567

16 out of 16 normalized.

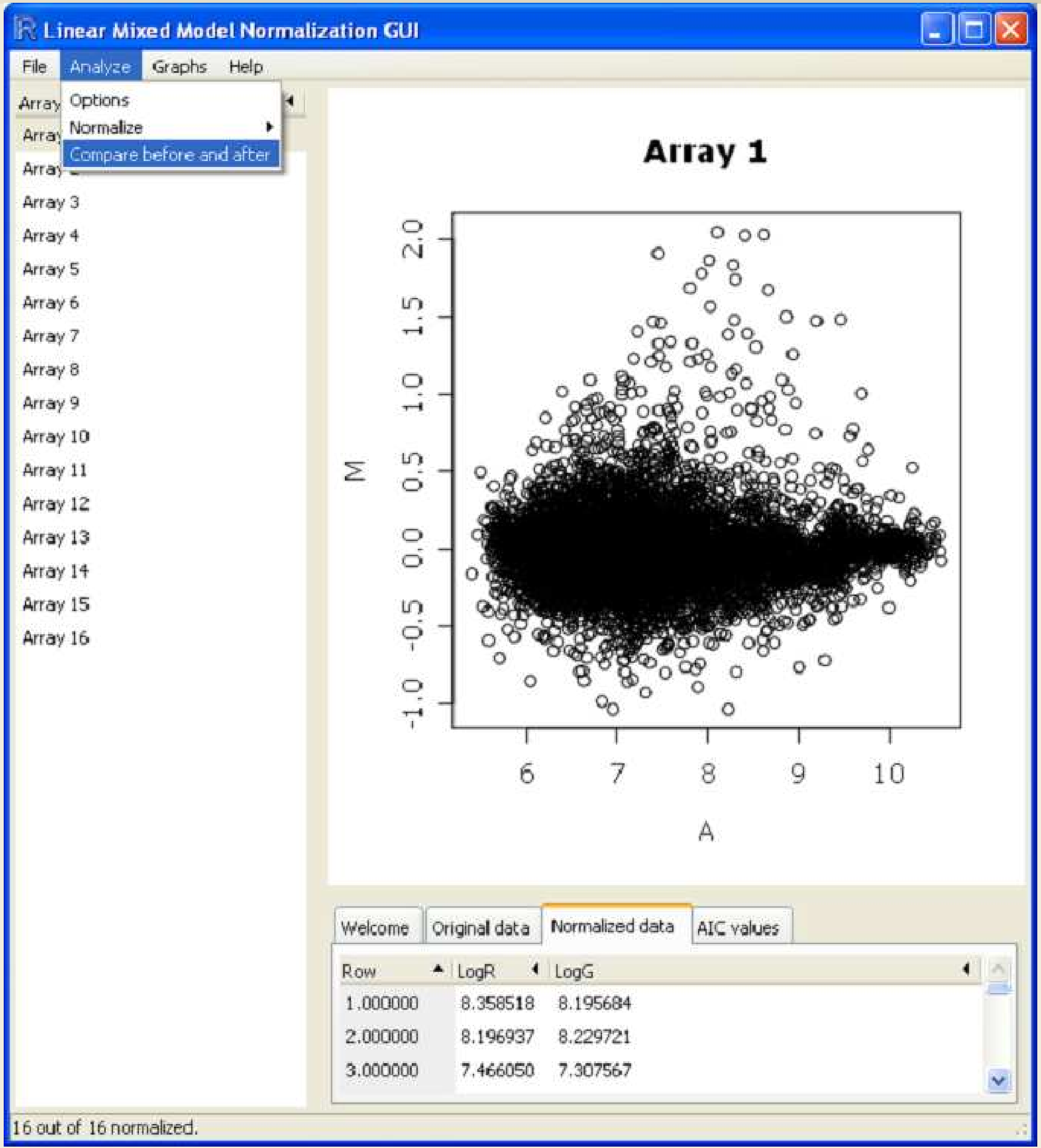
LMMNormGUI: a demo



V1	V2
Global normalization	3534.19256448584
Linear normalization	3127.44046842216
Non-Linear normalization	2885.65859113927
Non-Linear normalization with fixed print-tips	1939.13591390975
Non-Linear normalzation with a different smoother for each print-tip	1726.93331017607

16 out of 16 normalized.

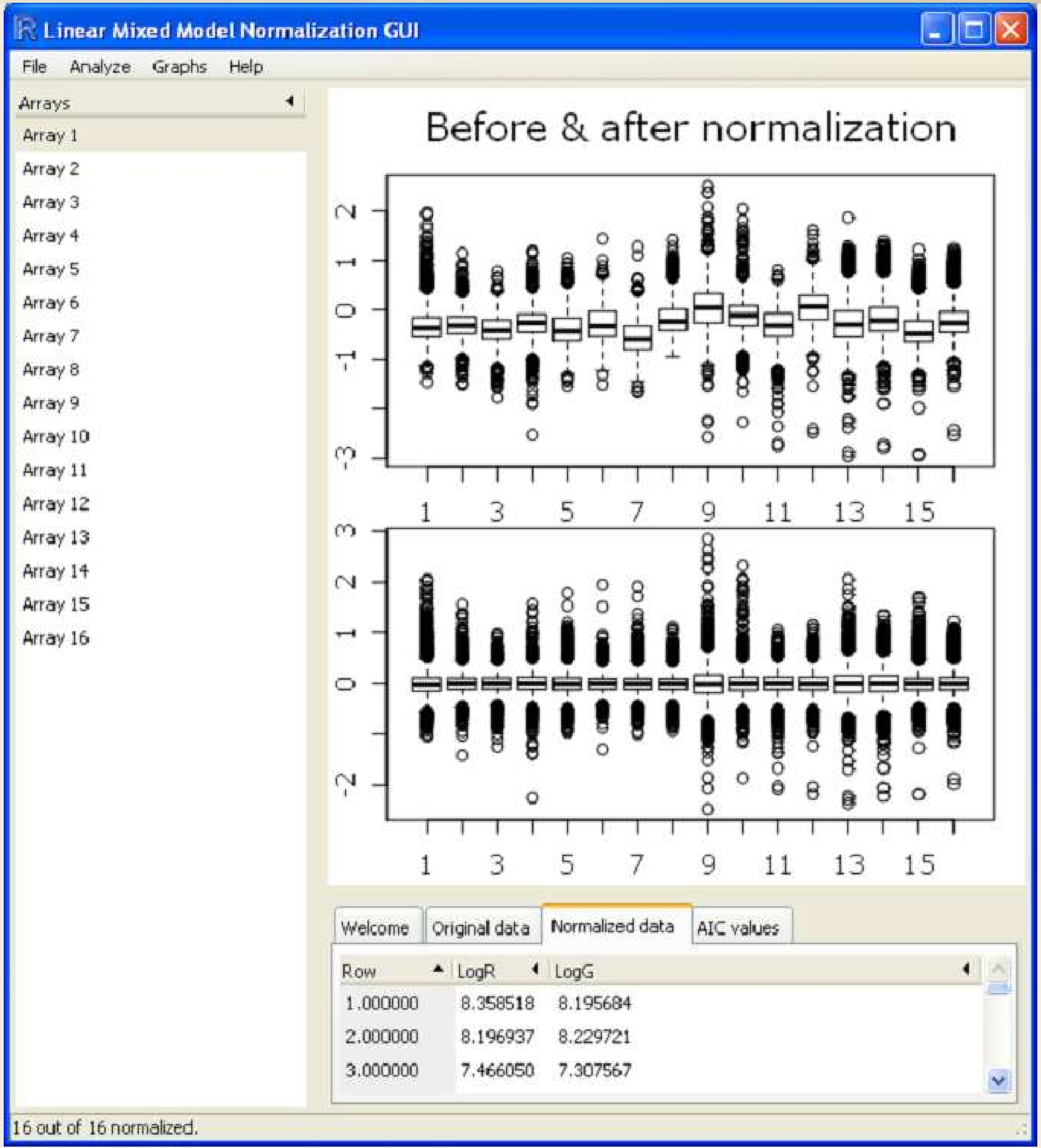
LMMNormGUI: a demo



Row	LogR	LogG
1.000000	8.358518	8.195684
2.000000	8.196937	8.229721
3.000000	7.466050	7.307567

16 out of 16 normalized.

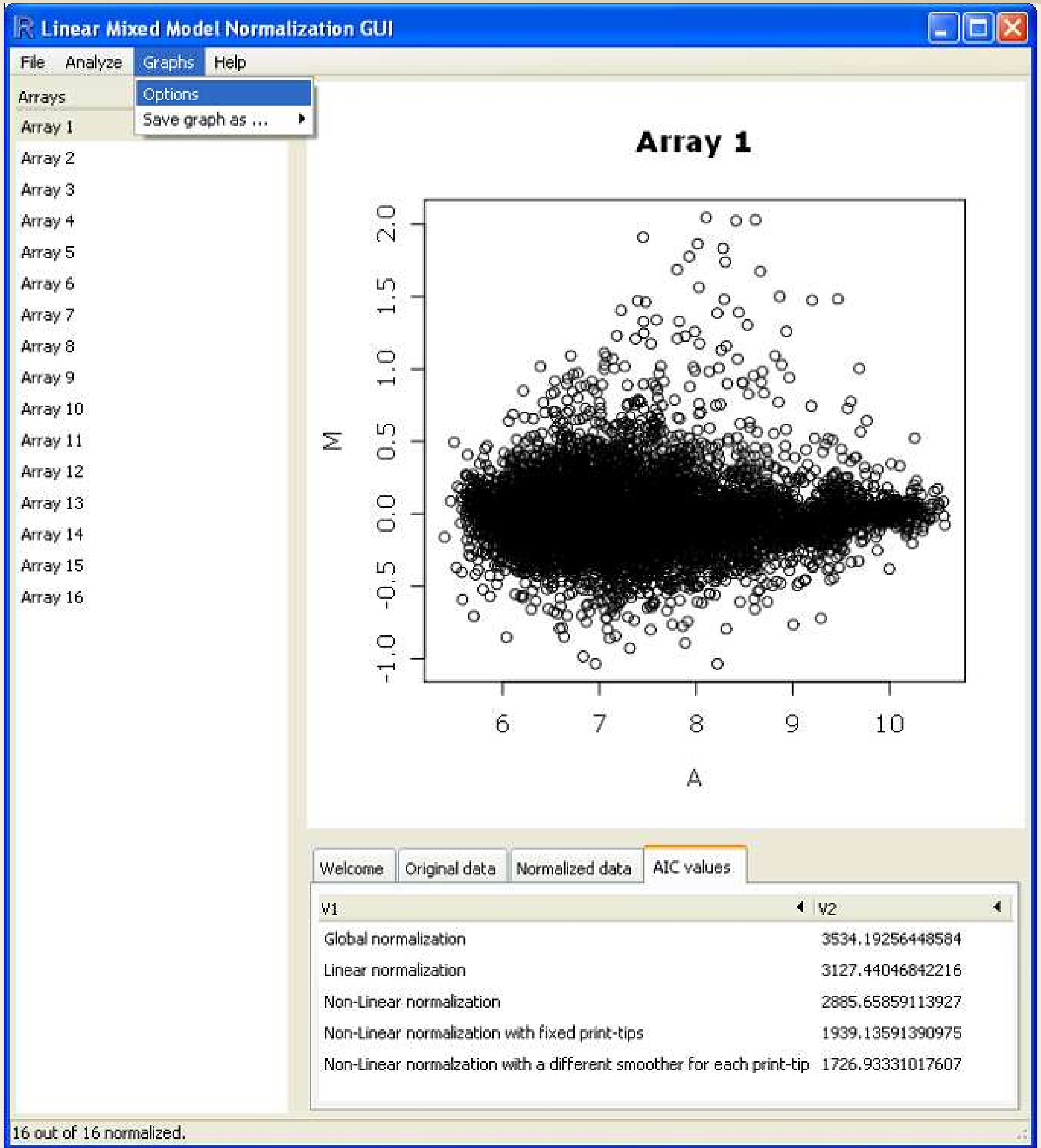
LMMNormGUI: a demo



Row	LogR	LogG
1.000000	8.358518	8.195684
2.000000	8.196937	8.229721
3.000000	7.466050	7.307567

16 out of 16 normalized.

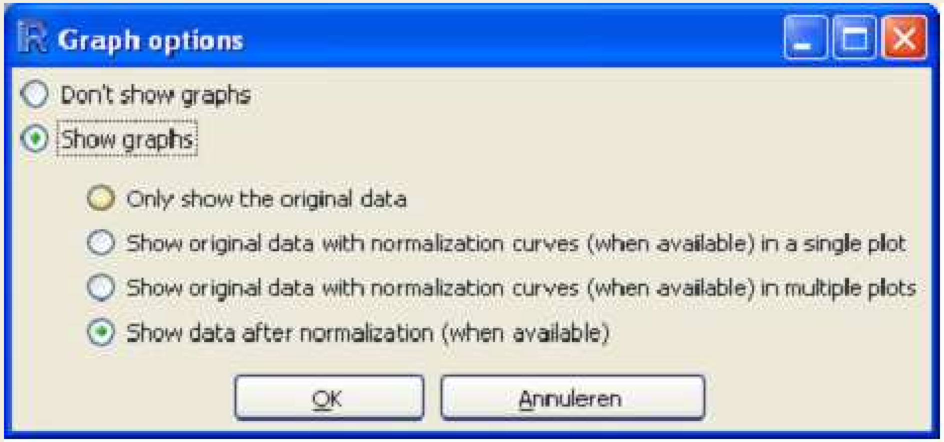
LMMNormGUI: a demo



V1	V2
Global normalization	3534.19256448584
Linear normalization	3127.44046842216
Non-Linear normalization	2885.65859113927
Non-Linear normalization with fixed print-tips	1939.13591390975
Non-Linear normalzation with a different smoother for each print-tip	1726.93331017607

16 out of 16 normalized.

LMMNormGUI: a demo



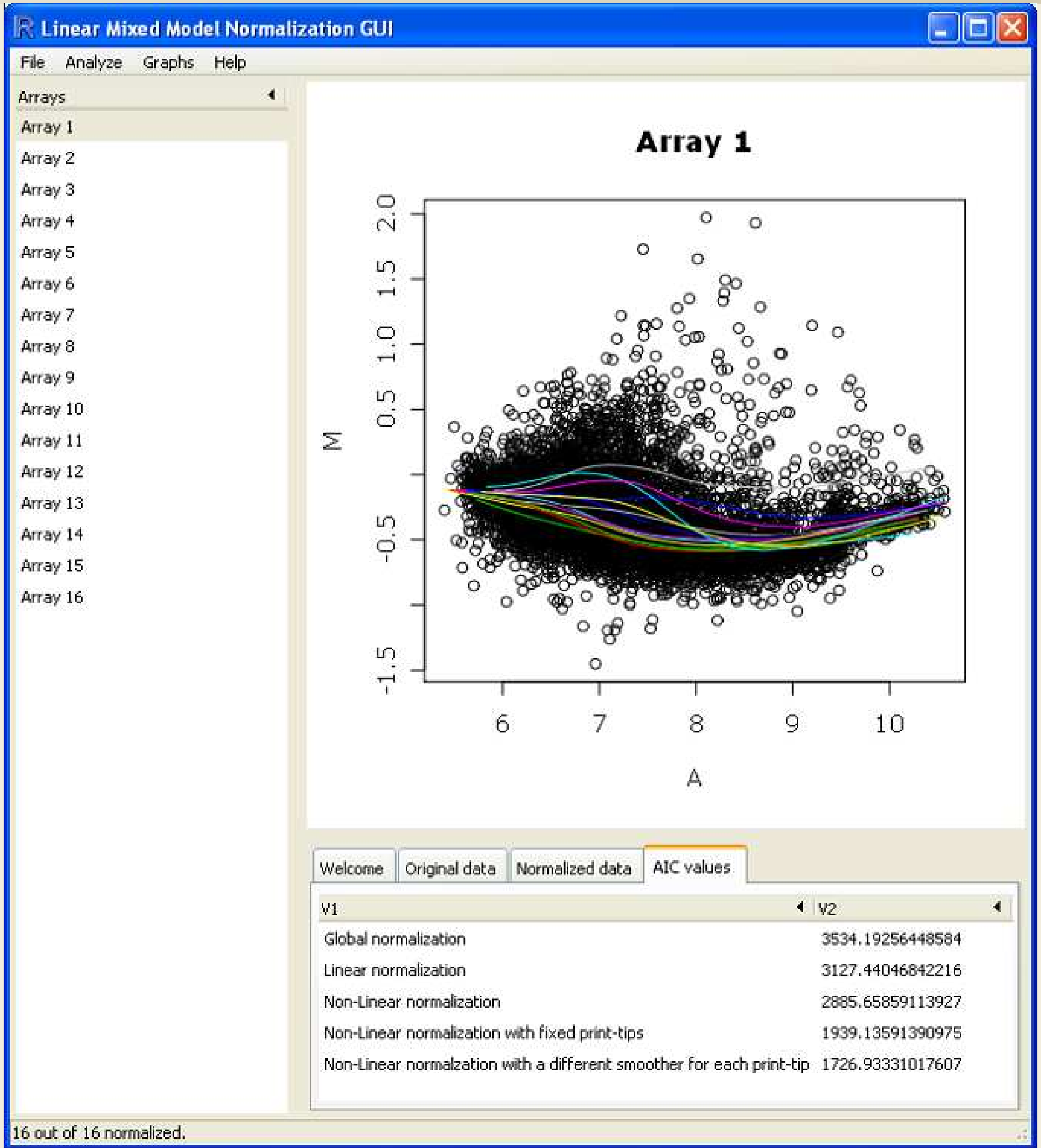
Graph options

○ Don't show graphs

● Show graphs

- ○ Only show the original data
- ○ Show original data with normalization curves (when available) in a single plot
- ○ Show original data with normalization curves (when available) in multiple plots
- ● Show data after normalization (when available)

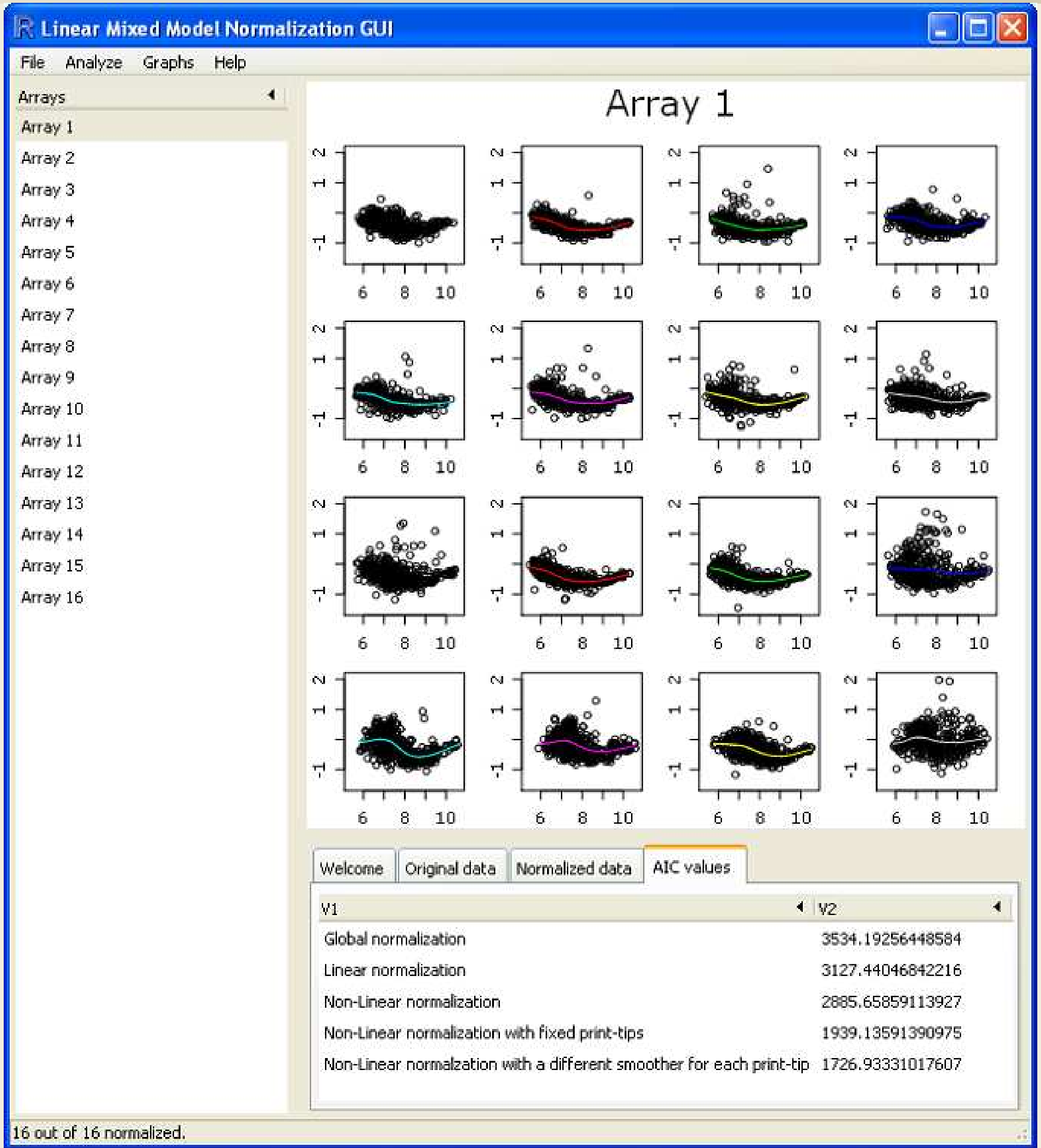
LMMNormGUI: a demo



V1	V2
Global normalization	3534.19256448584
Linear normalization	3127.44046842216
Non-Linear normalization	2885.65859113927
Non-Linear normalization with fixed print-tips	1939.13591390975
Non-Linear normalzation with a different smoother for each print-tip	1726.93331017607

16 out of 16 normalized.

LMMNormGUI: a demo



V1	V2
Global normalization	3534.19256448584
Linear normalization	3127.44046842216
Non-Linear normalization	2885.65859113927
Non-Linear normalization with fixed print-tips	1939.13591390975
Non-Linear normalzation with a different smoother for each print-tip	1726.93331017607

16 out of 16 normalized.

Future work

- Non constant variance
- Multiple arrays
- ...

References

- `http://www.censtat.be/software`
- Haldermans, P. Shkedy, Z., Van Sanden S., Burzykowksi T. and Aerts M. (2007). *Using Linear Mixed Models for Normalization of cDNA microarrays.* Statistical Applications in Genetics and Molecular Biology Vol. 6, Article 19.