

DRomics tool tutorial

Last update: 29/08/2018

Welcome in the DRomics tool tutorial. This tool aims to build dose-response curves from OMICs dataset, classify the molecular items (e.g probes, metabolites) based on their typology, and derive benchmark dose. In the available version, data are supposed to be microarray data transformed in log2 or another type of data (such as metabolomic data) that can be directly fitted by least-square regression without any normalization step. In the future, this tool will also be able to process RNA-seq data.

The DRomics tool contains 4 main steps, which take place on 4 different and specific tabs.

Before to start:

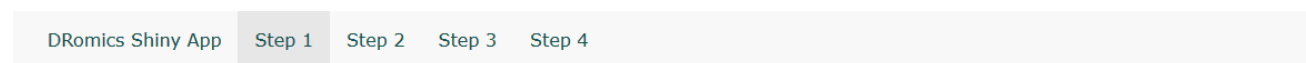
Please consider that the dataset must present the samples in columns and the molecular items in rows. The name of the columns (first row) must be the concentration value and the name of the rows (first column) must be the name/code of the items (e.g. probes, metabolites). The required format is “.txt”. As the DRomics application run on R software, please avoid spaces and special characters by naming the items.

In this tutorial, the functionalities of the application are illustrated with an example dataset, available for download in the app. This dataset encompasses the fluorescence value of 1000 probes for 6 concentrations in 5 replicates, obtained via microarray analysis, and were previously log2-transformed.

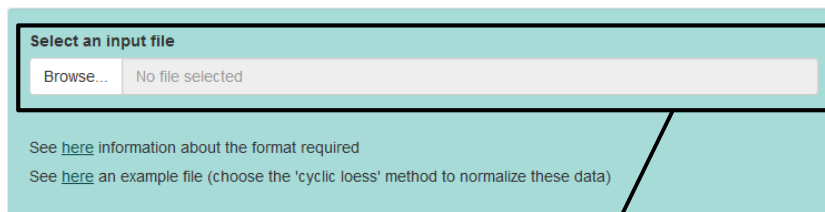
1. TAB STEP 1: IMPORT, CHECK AND NORMALIZATION OF OMIC DATA

The first step consists to import the dataset, check automatically the format and if the format is good, to proceed to the normalization of the dataset (this last one is optional). For more information about the suggested methods of normalization, please check the dedicated help section.

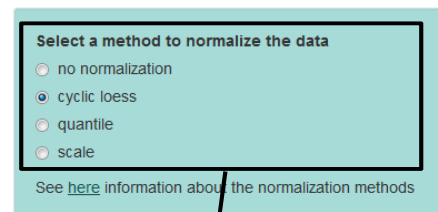
- How to proceed?



IMPORT, CHECK AND NORMALIZATION OF OMIC DATA



1.1 To import the dataset, click on “Browse” and select your file



1.2 Select the normalization method

- **Results of step 1**

The summary of the imported dataset is then provided as below.

```

Just wait, the normalization using cyclicloess may take a few minutes.
Elements of the experimental design in order to check the coding of the data :
Tested doses and number of replicates for each dose:

```

0	0.69	1.223	2.148	3.774	6.631
5	5	5	5	5	5

Number of items: 1000

```

Identifiers of the first 20 items:
[1] "1" "2" "3" "4" "5.1" "5.2"
[7] "6.1" "6.2" "7.1" "7.2" "8.1" "8.2"
[13] "9.1" "9.2" "10.1" "10.2" "11.1" "11.2"
[19] "12.1" "12.2"

```

Data were normalized between arrays using the following method: `cyclicloess`

Annotations:

- Tested concentrations (specified in the first row of the dataset)
- Number of replicates per concentration
- Names of the items (first column of the dataset)
- Normalization method

Once the normalization is done, a graph appears which compares the distribution of the values of each molecular item (y axis) in each sample (x axis) before and after normalization to visually check the normalization effect on the data.

Then, click on the tab “Step 2”.

2. TAB STEP 2: SELECTION OF SIGNIFICANTLY RESPONDING ITEMS

The second step aims to identify the significantly responding molecular items to the gradient. These items will then be selected to proceed to the next steps. For that, three different methods can be used: the quadratic trend test, the linear trend test and the ANOVA test. We recommend the use of the quadratic trend test for a typically dose-response design. Please check the help section to choose the most appropriate one for your design. The tool also offers the possibility to control the False Discovery Rate (FDR).

- **How to proceed?**

DRomics Shiny App Step 1 **Step 2** Step 3 Step 4

SELECTION OF SIGNIFICANTLY RESPONSIVE ITEMS

Select a method

quadratic trend test
 linear trend test
 ANOVA test

See [here](#) information about the selection methods

False Discovery Rate (FDR) for the Benjamini-Hochberg correction of p-values

0.05

Annotations:

- 2.1 Choose one of the three methods to select the responding items
- 2.2 A FDR can also be applied to the data and its value can be set by the user. A value set at 0.05 is proposed by default.

- **Results of step 2**

The number of selected items (the responsive ones) and the names of the 20 most responsive ones (identified as the ones with the lowest p-values obtained from the selection test) are finally presented in a box.

Number of selected items using a quadratic trend test with an FDR of 0.05 : 318 → **Number of selected items**

Identifiers of the first 20 most responsive items: → **Name of the 20 most responsive items (with the lowest p-values)**

```
[1] "384.2" "383.1" "383.2" "384.1" "381.1"
[6] "363.1" "380.2" "364.2" "364.1" "363.2"
[11] "381.2" "380.1" "351.1" "350.2" "239.1"
[16] "240.1" "240.2" "378" "15" "350.1"
```

Then, click on the tab “Step 3”.

3. TAB STEP 3: MODEL SELECTION AND DOSE-RESPONSE MODELLING

The third tab consists in the selection of the best-fit model for each previously selected item (those which respond significantly according to the chosen test and the FDR chosen value) and their respective dose-response curve building. For information about the best model selection please check the dedicated help section.

- **How to proceed?**



DOSE RESPONSE MODELLING FOR RESPONSIVE ITEMS

Click this button each time you update a setting in previous steps

See [here](#) information about the dose response modelling procedure

Download all the fitted dose-response plots

3.1 Click on “Fit” to start the procedure (model fitting and best model selection)

- **Results of step 3**

First, a summary of the analysis is provided in a box:

42 dose-response curves out of 318 previously selected were removed because no model could be fitted reliably. → **Number of items for which no model adequately fit the data**

Distribution of the chosen models among the 276 fitted dose-response curves :

Hill	linear
6	67
exponential	Gauss-probit
68	107
log-Gauss-probit	
28	

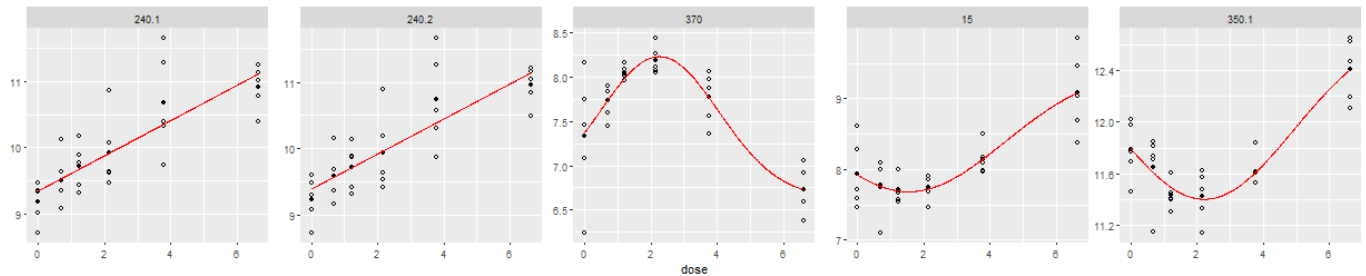
→ **Number of items per model (e.g. the Hill model best fitted the response of 6 items)**

Distribution of the typology of the 276 fitted dose-response curves :

H.inc	H.dec	L.inc
3	3	46
L.dec	E.inc.convex	E.dec.concave
21	20	25
E.inc.concave	E.dec.convex	GP.U
17	6	57
GP.bell	lGP.U	lGP.bell
50	17	11

→ **Number of items per typology (e.g. 46 items were modelled by a linear increasing curve)**
For more information about the full meaning of each typology name, please check the dedicated help section.

The graphical presentation of the dose-response curve of the first 20 items is also presented. Those Graphs present the value of each replicate (white dots) per sample as well as the mean value (black dots). The x axis presents the concentration values and the y axis the signal values.



Finally, the user can export the totality of the best-fit dose-response curves by clicking on “Download all the fitted dose-response plots”.

Then, click on the tab “Step 4”.

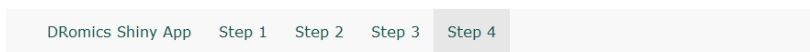
4. TAB STEP 4: COMPUTATION OF BENCHMARK DOSES FOR RESPONSIVE ITEMS

The last step aims to derive a Benchmark Dose (BMD) according to the EFSA report (Hardy et al., 2017). Two kinds of BMD can be derived by the tool:

- BMD-zSD: the BMD based on z times the standard deviation of the data along the whole curve
- BMD-xfold: the BMD based on a x-fold change of the signal compared to the control

For more details about how are calculated both kind of BMDs, please check the dedicated help section. For both BMDs, the value of z (factor multiplying the standard deviation value) and x (fold change value compared to the control) are modifiable by the user. A value of 1 and 10 (for z and x, respectively) are proposed by default.

- **How to proceed**



COMPUTATION OF BENCHMARK DOSES FOR RESPONSIVE ITEMS

- **Results of step 4**

This step provides two kinds of results.

First, a visualization of the 10 first rows of a dataframe summarizing the properties of the curves, and the BMD values with following columns:

- **id**: The name of each item
- **irow**: Their line number in the initial dataset
- **adjpvalue**: The p-value resulting from the selection test (step2)
- **model**: The best model associated to the item (step3)
- **npar**: the number of parameters of the model
- **b, c, d, e, f**: the value of the parameters of the model
- **SDres**:the residual standard deviation of the best model
- **typology**: the typology associated to the item
- **trend**: the main trend of the response
- **BMD.zSD**: the value of the BMD.SD for the selected z value
- **BMD.xfold**: the value of the BMD.xfold for the selected x value

```

10 BMD-xfold values and 0 BMD_zSD values are not defined
(coded NaN as the BMR stands outside the range of response values
defined by the model).
149 BMD-xfold values and 6 BMD_zSD values could not be calculated
(coded NA as the BMR stands within the range of response values defined by the model
but outside the range of tested doses).

  id irow  adjpvalue      model npar
1 384.2  727 2.519524e-07 Gauss-probit  4
2 383.1  724 6.558388e-07 Gauss-probit  4
3 383.2  725 8.234946e-07 Gauss-probit  4
4 384.1  726 2.804671e-06 Gauss-probit  4
5 301.1  569 6.932747e-06 exponential  3
6 363.1  686 7.084800e-06 exponential  3
7 300.2  568 7.569186e-06 exponential  3
8 364.2  689 8.165905e-06 exponential  3
9 364.1  688 1.162731e-05 exponential  3
10 363.2  687 1.171204e-05 exponential  3

      b      c      d      e
1  8.39029981  6.160057  6.160057  1.538639
2  3.81611516 10.480252 10.480252  1.833579
3  6.27821811  8.505651  8.505651  1.751051
4  8.59615317  5.683611  5.683611  1.874867
5  2.02444945      NA 12.846110 -1.404110
6 -0.06029687      NA  9.026630  2.064798
7  2.23832801      NA 11.215114 -2.105123
8 -0.08074800      NA  8.946640  2.238033
9 -0.04949102      NA  8.981422  1.939483
10 -0.03105373      NA  9.031729  1.728463

      f      SDres      typology trend
1  6.077678 0.1126233      GP.bell bell
2  1.861386 0.1411563      GP.bell bell
3  3.683446 0.1335847      GP.bell bell
4  6.569230 0.1379656      GP.bell bell
5      NA 0.4905313  E.dec.convex dec
6      NA 0.2526946  E.dec.concave dec
7      NA 0.5308073  E.dec.convex dec
8      NA 0.2726621  E.dec.concave dec
9      NA 0.2630138  E.dec.concave dec
10     NA 0.2427759  E.dec.concave dec

      BMD.zSD BMD.xfold
1  3.7849580      NA
2  0.8421349      NA
3  1.3656752      NA
4  1.2663420      NA
5  0.3896017  1.413406
6  3.4005072  5.721002
7  0.5698289  1.463580
8  3.3039995  5.576120
9  3.5741348  5.725695
10 3.7624970  5.883690

```

To load these data for all of the items for which a dose-response curve was build, click on the “Download results” button.

Second, the BMD value results are also restituted *via* a graphical output which can be modulated according the user wishes:

BMD type
 zSD
 xfold

plot type
 empirical cumulative distribution
 histogram
 density

by typology ?
 TRUE
 FALSE
 See [here](#) information about typologies

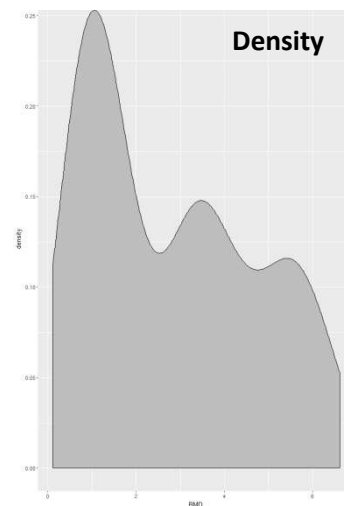
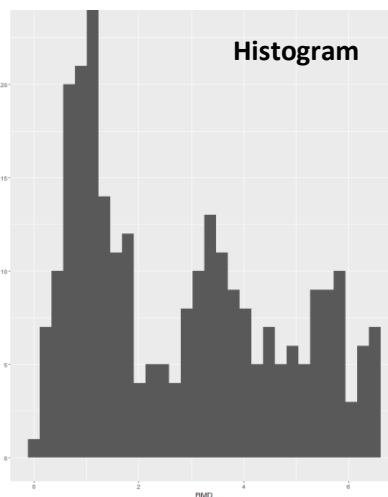
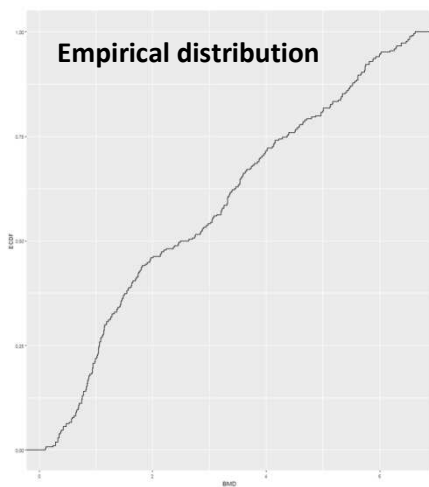
4.3 Select the BMD value of interest (zSD or xfold)

4.4 select a kind of plot

4.5 The plot can be done for the 12 typologies proposed by the tool or all typologies together (see below some examples)

4.6 Load the current plot

If by typology=FALSE:



If by typology=TRUE

