

# BMDx2 Manual

## Contents

|                                                     |           |
|-----------------------------------------------------|-----------|
| <b>Introduction</b>                                 | <b>1</b>  |
| <b>Set mode</b>                                     | <b>2</b>  |
| <b>Format of data input</b>                         | <b>2</b>  |
| Phenodata . . . . .                                 | 2         |
| Expression data . . . . .                           | 4         |
| <b>Filtering</b>                                    | <b>5</b>  |
| ANOVA and Trend test parameters . . . . .           | 6         |
| ANOVA and Trend test filtering Results . . . . .    | 7         |
| Fold-change filtering parameters . . . . .          | 7         |
| <b>BMD Analysis</b>                                 | <b>7</b>  |
| Model Fitting . . . . .                             | 8         |
| Model comparison . . . . .                          | 10        |
| Model Variance . . . . .                            | 10        |
| Model Adverse Direction . . . . .                   | 11        |
| R <sup>2</sup> Calculation . . . . .                | 12        |
| BMR Selection and BMD estimation . . . . .          | 12        |
| BMDL/BMDU Estimation . . . . .                      | 13        |
| Selection of Parameters for bmd modelling . . . . . | 13        |
| <b>BMD Results</b>                                  | <b>14</b> |
| Filtering BMD Results . . . . .                     | 14        |
| Selection of the optimal model . . . . .            | 18        |
| Visual investigation of the models . . . . .        | 18        |
| <b>Gene frequency analysis</b>                      | <b>27</b> |
| <b>FunMappOne enrichment analysis</b>               | <b>28</b> |
| <b>Compare gene pairs profile across doses</b>      | <b>31</b> |
| Gene Sets . . . . .                                 | 37        |
| PPI analysis . . . . .                              | 41        |
| <b>AOP and KE enrichment</b>                        | <b>41</b> |
| Individual AOP visualization . . . . .              | 50        |

## Introduction

BMDx is an R Shiny graphical interface tool for the dose dependent analysis of transcriptomic data.

BMDx implements a multi-step pipeline that include the following steps: - Set mode - Load phenotype data - Load experimental data, e.g. normalized expression matrix and vst normalized expression matrix

with blind=False for microarray and RNA-seq downstream analysis - Filtering - Benchmark dose analysis - Functional enrichment with the FunMappOne tool, e.g. KEGG pathways - Gene pairs analysis - AOP and KE enrichment - Download report

BMDx not only allows the user to compare the results between multiple time points, but also multiple experiments at once. Results along the way are visualised as several types of plots and the output can be downloaded as Excel files or figures in pdf format at multiple steps of the analysis.

The final results include the estimation of effective doses (BMD/BMDL/BMDU) for each dose-dependent gene. These estimates can be used for further analysis to help construct a biological interpretation of the exposure of interest.

## Set mode

BMDx is implemented to work in a dual mode: GLP (Good Laboratory Practice) and non-GLP mode. If GLP mode is selected by the user, all the steps of the analysis will be recorded in a log file. Furthermore, at every step of the analysis the user will be prompted to add information about the analysis. First, the user will be asked to give a description about the aims of the analysis. Next, for every step that the user performs, an explanation for performing the step and the selected values for the parameters will be requested.

## Format of data input

BMDx takes as an input a phenotype file and an expression matrix, both provided as an Excel spreadsheet (xlsx). If multiple experiments are included, both files must contain separate sheets for each experiment in corresponding orders. Specific instructions for the file structures are provided below. Sample data can be found [here](#).

### Phenodata

The phenotype file is an Excel file containing separate sheets for each experiment. Each sheet contains information about the samples used in the specific experiment. In particular, BMDx requires the spreadsheets to have at least three columns that specify the following characteristics: 1) Unique sample IDs corresponding to the column names in the expression matrix, 2) condition (e.g., dose) and 3) the time points included in the experiment. Each sheet must have the columns (sample ID, condition, and time point) in the same positions.

The required parameters are the following:

- **Path to pheno data file:** Browse and select the file from your local system
- **Sample ID variable:** Column name of the phenodata of the samples IDs
- **Dose variable:** Column name of the phenodata of the dose levels
- **Time point variable:** Column name of the phenodata of the time points
- **Optional variable:** Column names of the phenodata that can be used during the analysis to group the results

Once all options are set, load the phenotype data and proceed to the next step.

|    | A         | B         | C         | D | E | F |
|----|-----------|-----------|-----------|---|---|---|
| 1  | sample_id | condition | timepoint |   |   |   |
| 2  | F_1       | 0         | 24        |   |   |   |
| 3  | F_2       | 0         | 24        |   |   |   |
| 4  | F_3       | 0         | 24        |   |   |   |
| 5  | F_4       | 0         | 24        |   |   |   |
| 6  | F_5       | 20        | 24        |   |   |   |
| 7  | F_6       | 20        | 24        |   |   |   |
| 8  | F_7       | 20        | 24        |   |   |   |
| 9  | F_8       | 20        | 24        |   |   |   |
| 10 | F_9       | 40        | 24        |   |   |   |
| 11 | F_10      | 40        | 24        |   |   |   |
| 12 | F_11      | 40        | 24        |   |   |   |
| 13 | F_12      | 40        | 24        |   |   |   |
| 14 | F_13      | 60        | 24        |   |   |   |
| 15 | F_14      | 60        | 24        |   |   |   |
| 16 | F_15      | 60        | 24        |   |   |   |
| 17 | F_16      | 60        | 24        |   |   |   |
| 18 | F_17      | 80        | 24        |   |   |   |
| 19 | F_18      | 80        | 24        |   |   |   |
| 20 | F_19      | 80        | 24        |   |   |   |

Figure 1: Sample pheno data

Import Phenotype Data

**Choose phenodata file**

Browse... phenodata\_all\_.csv  
Upload complete Load Phenodata File

| Variable      | Type   | Class     | Sample1 | Sample2 | Sample3 | Sample4 | Sample5 | Sample6 |
|---------------|--------|-----------|---------|---------|---------|---------|---------|---------|
| sample_id [1] | factor | character | F_1     | F_2     | F_3     | F_4     | F_5     | F_6     |
| condition [2] | vector | numeric   | 0       | 0       | 0       | 0       | 20      | 20      |
| timepoint [3] | vector | numeric   | 24      | 24      | 24      | 24      | 24      | 24      |

**Sample ID Variable**: sample\_id  
**Dose Variable**: condition  
**Time Point Variable**: timepoint

Set Optional Variables of Interest

Next Close

An overview of the metadata is presented in the “Metadata data” window.

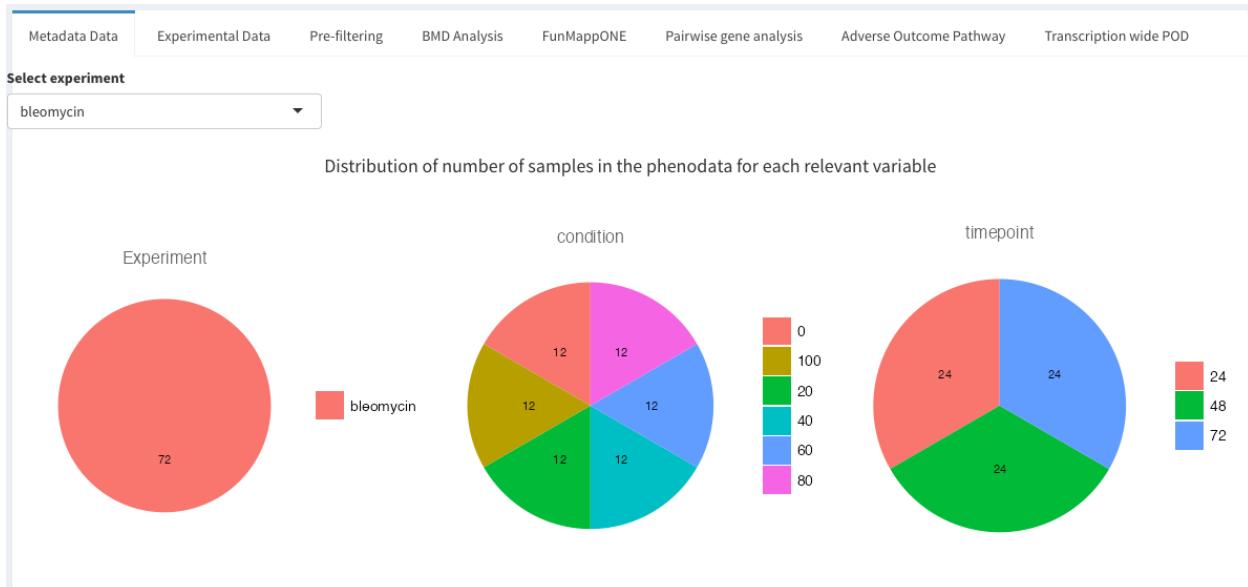


Figure 2: Overview metadata

## Expression data

The expression matrix is an Excel file with a separate sheet for each experiment. Each experiment involves a unique test system or materials being tested, such as different cell lines or treatments. Note that different doses or timepoints tested within the same system are not considered separate experiments. Sample columns

are named with unique sample IDs that match the sample IDs provided in the phenotype file. Gene names must be provided in the first column of each spreadsheet, and each following column specifies the expression values for those genes in each individual sample. The order of the sheets must match the order of the sheets in the phenotype file.

|    | A               | B        | C        | D        | E        | F        | G        | H        | I        | J        | K        | L        | M        | N        | I    |
|----|-----------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|------|
| 1  |                 | F_1      | F_2      | F_3      | F_4      | F_5      | F_6      | F_7      | F_8      | F_9      | F_10     | F_11     | F_12     | F_13     | F_14 |
| 2  | ENSG00000227232 | 7,497576 | 7,285982 | 7,479587 | 7,347935 | 7,649445 | 7,691155 | 7,632452 | 7,953827 | 7,646189 | 7,593268 | 7,016584 | 7,785649 | 7,666028 | 7,74 |
| 3  | ENSG00000278267 | 5,882885 | 5,662301 | 5,716926 | 5,618714 | 5,967409 | 5,925647 | 5,970899 | 6,245134 | 6,178336 | 5,876374 | 5,520354 | 5,822176 | 5,635772 | 6,06 |
| 4  | ENSG00000268903 | 6,310302 | 6,608071 | 6,580339 | 6,611976 | 6,465182 | 6,800613 | 6,569819 | 6,725432 | 6,647411 | 6,287911 | 6,478043 | 6,472526 | 6,524565 | 6,75 |
| 5  | ENSG00000269981 | 6,28737  | 6,291085 | 6,296516 | 6,575125 | 6,326449 | 6,681719 | 6,28363  | 6,463874 | 6,23644  | 6,513853 | 6,34644  | 6,409343 | 6,573787 | 6,86 |
| 6  | ENSG00000279457 | 7,966986 | 8,003673 | 7,936404 | 8,1063   | 8,223592 | 8,211427 | 8,073408 | 8,732532 | 8,040742 | 8,008473 | 8,213937 | 7,972674 | 7,90044  | 8,2  |
| 7  | ENSG00000237094 | 6,163593 | 6,16695  | 6,296516 | 6,195599 | 6,05785  | 6,192492 | 5,847446 | 5,992053 | 6,207845 | 6,287911 | 6,041172 | 5,777172 | 6,381048 | 6,12 |
| 8  | ENSG0000025630  | 13,36306 | 13,32023 | 13,46953 | 13,61588 | 13,22718 | 13,14753 | 13,32102 | 12,88174 | 13,02127 | 13,1815  | 13,4918  | 13,12413 | 13,57326 | 13,5 |
| 9  | ENSG00000237973 | 8,393396 | 8,32002  | 8,331296 | 8,205625 | 8,527186 | 8,15463  | 8,795058 | 8,578213 | 8,783336 | 8,461408 | 8,24107  | 8,239167 | 8,790011 | 8,76 |
| 10 | ENSG00000248527 | 13,12682 | 12,9998  | 13,13486 | 13,32896 | 12,84333 | 12,80933 | 12,97916 | 12,584   | 12,60079 | 12,69757 | 12,92658 | 12,67167 | 13,1422  | 13,2 |
| 11 | ENSG00000228327 | 7,734861 | 7,399523 | 7,444651 | 7,489642 | 7,425269 | 7,490713 | 7,388094 | 7,365817 | 7,430635 | 7,539611 | 7,65042  | 7,58279  | 7,63     |      |
| 12 | ENSG00000228794 | 8,596962 | 8,566519 | 8,493685 | 8,553211 | 8,836291 | 8,839323 | 8,839312 | 8,693706 | 8,821111 | 8,778592 | 8,901744 | 8,879992 | 8,989517 | 8,81 |
| 13 | ENSG00000223764 | 7,374965 | 7,255992 | 7,137034 | 7,200334 | 7,35619  | 7,124997 | 7,021326 | 6,898893 | 7,059882 | 6,919541 | 6,757471 | 6,893542 | 6,881495 | 7,03 |
| 14 | ENSG00000187634 | 8,155454 | 7,962326 | 7,930222 | 7,903526 | 7,538157 | 7,707263 | 7,3515   | 7,473308 | 7,500639 | 7,523406 | 6,887759 | 7,473486 | 7,439498 | 7,91 |
| 15 | ENSG00000188976 | 10,97395 | 10,93834 | 11,0076  | 10,91004 | 11,09699 | 11,07907 | 10,91618 | 11,38443 | 11,07682 | 11,1069  | 10,82725 | 11,03439 | 11,20153 | 11,3 |
| 16 | ENSG00000187961 | 8,453406 | 8,422967 | 8,403819 | 8,691959 | 8,984087 | 9,090867 | 8,326823 | 8,960509 | 8,537405 | 8,424475 | 8,377558 | 8,656328 | 8,144321 | 8,87 |
| 17 | ENSG00000187583 | 5,987833 | 6,053641 | 6,272864 | 6,139431 | 6,028821 | 6,069725 | 6,200798 | 6,644841 | 6,291161 | 6,235878 | 6,012622 | 6,244552 | 6,009169 | 6,1  |
| 18 | ENSG00000188290 | 5,358854 | 5,112694 | 5,360872 | 5,618714 | 6,113047 | 6,274753 | 6,156431 | 6,628018 | 6,608794 | 5,916671 | 6,478043 | 6,342037 | 6,437189 | 6,58 |
| 19 | ENSG00000187608 | 8,638078 | 8,766322 | 8,911036 | 8,797479 | 8,796035 | 8,935941 | 8,850164 | 9,096298 | 8,935677 | 9,097382 | 8,722256 | 9,283779 | 9,02321  | 8,99 |
| 20 | ENSG00000188157 | 13,22761 | 13,18897 | 13,13008 | 13,04776 | 13,06808 | 13,17172 | 12,89612 | 13,58633 | 13,03571 | 13,15349 | 12,84076 | 13,10974 | 13,3797  | 13,3 |
| 21 | ENSG00000217801 | 7,166575 | 6,935994 | 7,181057 | 7,337927 | 8,080037 | 8,039921 | 7,782857 | 8,378477 | 7,960761 | 7,683797 | 7,796589 | 8,015735 | 7,956031 | 7,76 |
| 22 | ENSG00000273443 | 5,95446  | 5,604865 | 5,46324  | 5,867    | 6,085898 | 6,248123 | 6,171463 | 6,463874 | 5,937435 | 5,83361  | 5,851359 | 6,044458 | 6,065104 | 6,22 |
| 23 | ENSG00000131591 | 8,348916 | 8,465488 | 8,43873  | 8,606341 | 8,496499 | 8,810354 | 8,131486 | 8,742939 | 8,347767 | 8,424475 | 8,200171 | 8,577283 | 8,525629 | 8,93 |
| 24 | ENSG00000186891 | 5,803175 | 5,604865 | 5,664942 | 5,824739 | 5,864795 | 6,036073 | 6,043262 | 6,292903 | 6,049363 | 6,208746 | 5,814201 | 5,822176 | 5,949    | 5,8  |
| 25 | ENSG00000078808 | 12,45494 | 12,51943 | 12,54273 | 12,53627 | 12,46757 | 12,76185 | 12,36801 | 12,84262 | 12,52802 | 12,55682 | 12,39484 | 12,65578 | 12,49576 | 12,8 |

The required parameters are the following:

- **Path to expression data file:** Browse and select the file from your local system
- **x:** Name of dose variable (default = “dose”)
- **y:** Name of expression variable (default = “exp”)

Import Gene Expression Table

---

|                                                              |                                                               |                                                                   |
|--------------------------------------------------------------|---------------------------------------------------------------|-------------------------------------------------------------------|
| <b>File</b>                                                  | <b>add name for x variable</b>                                | <b>add name for y variable</b>                                    |
| <input type="button" value="Browse..."/> vst_counts_all_.csv | dose                                                          | expr                                                              |
| <b>Upload complete</b>                                       |                                                               |                                                                   |
| <input type="checkbox"/> RNASEq raw counts                   | <input checked="" type="checkbox"/> First columns as rownames | <input checked="" type="checkbox"/> Force columns to be numerical |
| <b>Import</b>                                                |                                                               |                                                                   |
| <b>Close</b>                                                 |                                                               |                                                                   |

Figure 3: Load expression data parameters

Once all options are set, import the normalized expression matrix and proceed to the next step. An overview of the gene expression matrix is presented in the “Experimental data” window.

## Filtering

The number of genes (or molecules) profiled by transcriptomic data is usually high. Thus, in order to reduce the time of the BMDx analysis and focus on relevant genes, different filtering criteria are available:



Figure 4: Overview expression data

- **ANOVA:** ANOVA is used to test the null hypothesis that the mean responses (e.g. expression of the genes) at the different doses are the same. The alternative hypothesis is that the response is different in at least one of the doses. If the p-value is below a certain threshold (e.g. 0.05) then we reject the null hypothesis that each dose mean response is the same.
- **Mann-Kendall Trend test:** The Mann-Kendall Trend tests is performed under the null hypothesis that the data come from a population with independent realizations and are identically distributed. The alternative hypothesis is that the data follow a monotonic trend. If the p-value is below a certain threshold (e.g. 0.05) then the null hypothesis is reject and the response is considered to follow a monotonic trend.
- **Filter by fold-change:** The filter by fold-change is performed by performing a limma analysis between the treated samples (all the different doses) against the controls of each individual time-point. Details on the limma package can be found [here](#).
- **Trend filtering:** It allows for filtering genes that do not exhibit a linear trend. For each data point and the next, you compute the derivative and ensure they all have the same sign. This method is less stringent than monotonic filtering, as monotonic filtering requires a consistent increase or decrease between consecutive points.

The user can also decide to skip the filtering.

## ANOVA and Trend test parameters

- **Pvalue adjustment method:** The adjustment method for the p-value. Possible values are “nominal”, “fdr”, “bonferroni” and “BH”. See [this link](#) for more information.
- **P-value threshold:** The threshold on the p-value used to reject the null hypothesis. Default = 0.05
- **Number of cores:** Number of cores used to perform the filtering. The number of scores affect the processing speed, as the quantity of cores directly impact the efficiency of the analysis.



Figure 5: Anova and trend test parameters

## ANOVA and Trend test filtering Results

The results of the Anova and Trend test analysis will be displayed in a table with the following columns:

- Exp:** the name of the experiment
- Time:** the time point of the experiment
- Feature:** the name of the gene
- Pvalue:** the p-value obtained from the analysis
- adjPVal:** the adjusted p-values
- usedFilteringPVal:** the p-value used to filter the results. If the adjustment is performed, usedFilteringPVal will be equal to adjPVal, otherwise will be equal to Pvalue.

## Fold-change filtering parameters

- Pvalue adjustment method:** The adjustment method for the p-value. Possible values are “nominal”, “fdr”, “bonferroni” and “BH”. Nominal means no adjustment will be performed.
- Pvalue threshold:** The threshold on the p-value used to reject the null hypothesis. Default = 0.05
- Fold-change threshold::** The minimum fold-change to consider a gene differentially expressed (default 1.5)
- Number of cores:** Number of cores used to perform the filtering



## Fold-change filtering results

The results of the filtering by fold-change will be displayed in a table with the following columns:

- Exp:** the name of the experiment
- Time:** the time point of the experiment
- Feature:** the name of the gene
- logFC:** the name of the gene
- P.value:** the pvalue obtained from the analysis
- adj.P.Val:** the adjusted pvalues
- usedddFilteringPVal:** the pvalue used to filter the results. If the adjustment is performed, usedddFilteringPVal will be equal to adj.P.Val, otherwise will be equal to P.value

## BMD Analysis

The benchmark dose (BMD) is the dose or concentration of a substance that corresponds to a specified level of response above or below that observed in a control or background population. The specified level of response within this definition is referred to as the benchmark response (BMR), while the statistical lower confidence bound of the BMD (referred to as BMDL) and the statistical upper confidence bound on the BMD (BMDU) have been typically used by regulatory agencies to set safe levels of exposure. BMD modelling involves fitting the experimental data, in this case, the gene expression values, to a selection of mathematical models, such as linear, second- or third-degree polynomial, an exponential model, hill model, asymptotic regression model, and Michaelis-Menten model. Additionally, when multiple models can be fitted to the same gene, their average consensus model can be also produced. The best model is selected by using a goodness of fit criteria, such as the Akaike information and the goodness-of-fit p-value. Alternatively, the

average consensus model can be also used. A predefined response level of interest, the BMR, is identified and the optimal model is used to predict the corresponding dose (BMD) (Abraham et al. 2012). Moreover, the European Food Safety Authority (EFSA) suggest reporting both the lower and upper 95% confidence limit on the BMD called BMDL and BMDU respectively (EFSA Scientific Committee et al. 2017).

## Model Fitting

### Linear Model

$$f(\text{dose}) = \beta_0 + \beta_1 \text{dose}$$

Where  $\beta_0$  is the control response (intercept), and  $\beta_1$  is the slope.

### Polynomial Model

$$f(\text{dose}) = \beta_0 + \beta_1 \text{dose} + \beta_2 \text{dose}^2 + \dots + \beta_n \text{dose}^n$$

Where  $\beta_0$  is the control response (intercept),  $\beta_1, \dots, \beta_n$  are the polynomial coefficients, and  $n$  is the degree of the polynomial.

### Power Model

$$f(\text{dose}) = \beta_0 + \beta_1 \times (\text{dose})^\delta$$

Where  $\beta_0$  is the control response (intercept),  $\beta_1$  is the slope, and  $\delta$  is the power.

### Hill Model

$$f(\text{dose}) = \beta_0 + \frac{v \times \text{dose}^n}{K^n + \text{dose}^n}$$

Where  $\beta_0$  is the control response,  $v$  is the maximum response,  $K$  is the dose at which half the maximum response is reached, and  $n$  is the Hill coefficient.

### Exponential Model

#### Exp2:

$$f(\text{dose}) = a \times e^{\pm b \times \text{dose}}$$

#### Exp3:

$$f(\text{dose}) = a \times e^{\pm(b \times \text{dose})^d}$$

#### Exp4:

$$f(\text{dose}) = a \times (c - (c - 1) \times e^{\pm b \times \text{dose}})$$

#### Exp5:

$$f(\text{dose}) = a \times (c - (c - 1) \times e^{\pm(b \times \text{dose})^d})$$

Where  $a$  is the control response (intercept),  $b$  is the slope,  $c$  is the asymptote term, and  $d$  is the power.

### Log-Logistic Model

#### Log-Logistic 5:

$$f(\text{dose}, (b, c, d, e, f)) = c + \frac{d - c}{(1 + \exp(b(\log(\text{dose}) - \log(e))))^f}$$

If  $f \neq 1$ , the function is asymmetric; otherwise, it is symmetric (on a log scale).

**Log-Logistic 4:**

$$f(\text{dose}, (b, c, d, e)) = c + \frac{d - c}{1 + \exp(b(\log(\text{dose}) - \log(e)))}$$

The function is symmetric about the inflection point  $e$ .

**Log-Logistic 3:**

$$f(\text{dose}, (b, c, e)) = c + \frac{1 - c}{1 + \exp(b(\log(\text{dose}) - \log(e)))}$$

The function is symmetric about the inflection point  $e$ .

**Log-Logistic 2:**

$$f(\text{dose}, (b, e)) = \frac{1}{1 + \exp(b(\log(\text{dose}) - \log(e)))}$$

The function is symmetric about the inflection point  $e$ .

**Weibull Models****Weibull 1.2:**

$$f(\text{dose}, (b, e)) = 1 - \exp(-b \cdot \text{dose}^e)$$

This model describes an asymmetric sigmoidal growth curve.

**Weibull 1.3:**

$$f(\text{dose}, (b, d, e)) = d - d \cdot \exp(-b \cdot \text{dose}^e)$$

Extends Weibull 1.2 by adding a scaling parameter  $d$  that adjusts the upper asymptote.

**Weibull 1.4:**

$$f(\text{dose}, (b, c, d, e)) = c + (d - c) \cdot (1 - \exp(-b \cdot \text{dose}^e))$$

Allows both upper ( $d$ ) and lower ( $c$ ) asymptotes, providing greater flexibility.

**Weibull 2.2:**

$$f(\text{dose}, (b, e)) = \exp(-\exp(b \cdot (\log(\text{dose}) - e)))$$

A sigmoidal decay function where  $e$  is the inflection point.

**Weibull 2.3:**

$$f(\text{dose}, (b, d, e)) = d \cdot \exp(-\exp(b \cdot (\log(\text{dose}) - e)))$$

Extends Weibull 2.2 with a scaling parameter  $d$ , modifying the maximum response.

**Weibull 2.4:**

$$f(\text{dose}, (b, c, d, e)) = c + (d - c) \cdot \exp(-\exp(b \cdot (\log(\text{dose}) - e)))$$

Generalizes Weibull 2.3 by incorporating both lower ( $c$ ) and upper ( $d$ ) asymptotes.

**Michaelis-Menten Model****MM.3 Model:**

$$f(\text{dose}, (c, d, e)) = c + \frac{d - c}{1 + (e/\text{dose})}$$

This model increases as a function of dose, attaining the lower limit  $c$  at dose 0 and the upper limit  $d$  for infinitely large doses. The parameter  $e$  corresponds to the dose yielding a response halfway between  $c$  and  $d$ .

**MM.2 Model:** A two-parameter version of the Michaelis-Menten model is obtained by setting  $c = 0$ .

## Average model

BMDx implements model averaging through an AIC-based combination of models fitted to the same gene's dose-response data. By calculating Akaike weights from the AIC values of each model, it assigns relative probabilities to the models, reflecting their support from the data. Predictions for benchmark dose (BMD), lower bounds (BMDL), and upper bounds (BMDU) are generated for each model and combined using these weights to produce averaged estimates. This approach ensures that the final predictions account for model uncertainty, integrating information from multiple models while penalizing overly complex or poorly fitting ones.

**Model Averaging Methodology** For every gene with multiple available models, the **average model** is obtained by weighting the predictions of the models by their AIC values. This ensures that models with a better fit contribute more to the final prediction while models with poorer fits have less influence.

The following steps are followed:

1. **Fitting individual models:** Each model is fitted separately to the dose-response data.
2. **Computing AIC values:** The Akaike Information Criterion (AIC) values for each model are extracted.
3. **Calculating Akaike Weights:**
  - The difference between each model's AIC and the minimum AIC is computed.
  - These differences are transformed into Akaike weights using the softmax function:

$$w_i = \frac{\exp(-\frac{AIC_i - AIC_{min}}{2})}{\sum \exp(-\frac{AIC_i - AIC_{min}}{2})}$$

where  $w_i$  represents the weight assigned to each model.

4. **Computing Weighted Predictions:**

- Each model's predictions for BMD, BMDL, and BMDU are obtained.
- The final averaged prediction is computed as:

$$\hat{Y} = \sum w_i \cdot Y_i$$

where  $Y_i$  is the prediction from each model.

## Model comparison

Model comparison is performed by means of the Akaike Information Criterion (AIC) implemented in the R stats package as follows:

$$-2x\log - likelihood + kxnpar$$

where  $npar$  represents the number of parameters in the fitted model, and  $k = 2$  for the usual AIC.

## Model Variance

BMDx provides multiple options for handling variance assumptions when fitting dose-response models. Users can specify whether variance should be treated as **constant**, **nonconstant**, or **model-dependent**, or they can allow the tool to infer the appropriate variance assumption based on statistical testing.

## Variance Selection Options

BMDx supports four methods for variance estimation:

- “**infer**” (default): The tool automatically determines whether variance is constant or nonconstant using the **Levene test** for homogeneity of variance. If the test returns a p-value below the chosen significance threshold (e.g., 0.05), variance is classified as **nonconstant**. Otherwise, variance is assumed to be **constant**.
- “**constant**”: The model assumes homoscedasticity (equal variance across all observations). Variance is estimated using the **Mean Squared Error (MSE)** of the residuals, calculated as:

$$\sigma^2 = \frac{1}{n} \sum (r_i^2)$$

where  $r_i$  represents the model residuals.

- “**nonconstant**”: Variance is assumed to be heteroscedastic (changing across dose levels). It is estimated using only the residuals from control samples (dose = 0), following the **Root Mean Squared Error (RMSE)** approach for the control group.

## Implementation Details

If `control_variance` is set to "infer", BMDx applies the Levene test to assess variance homogeneity across groups. The test evaluates whether population variances are equal (null hypothesis). If the test result (p-value) is below the selected significance level, variance is classified as **nonconstant**; otherwise, it is considered **constant**.

For **constant variance**, the model estimates variance using the residuals from all samples. For **nonconstant variance**, only residuals from control samples (dose = 0) are used for variance estimation. If **model-based variance** is selected, a variance function is fitted to the residuals, and variance predictions are incorporated into a weighted regression model to refine the estimates dynamically.

By default, if no variance option is specified, BMDx will automatically infer the variance type based on statistical testing.

## Model Adverse Direction

The adverse direction of the models is determined based on the sign of specific model parameters. This estimation follows different approaches depending on the model type:

- **Linear Model:** The sign of the slope coefficient is used to determine the direction of the adverse response. A positive slope indicates a **positive (+1)** adverse reaction, while a negative slope indicates a **negative (-1)** adverse reaction.
- **Power Model:** The sign of the **b** coefficient is used to assess the adverse direction. A positive **b** coefficient corresponds to a **positive (+1)** adverse reaction, whereas a negative **b** coefficient corresponds to a **negative (-1)** adverse reaction.
- **Hill Model:** The sign of the **v** coefficient is used to estimate the adverse reaction direction. A positive **v** coefficient indicates a **positive (+1)** adverse reaction, while a negative **v** coefficient indicates a **negative (-1)** adverse reaction.
- **Polynomial, Exponential, and Models from the drc Package:** For these models, a linear regression is performed, and the sign of the slope coefficient is used to estimate the adverse reaction direction. A positive slope results in a **positive (+1)** adverse reaction, whereas a negative slope results in a **negative (-1)** adverse reaction.

## R<sup>2</sup> Calculation

The coefficient of determination (**R<sup>2</sup>**) is used to evaluate the goodness of fit of the models. It quantifies the proportion of variance in the observed data that is explained by the fitted model. Different model types use distinct approaches to compute **R<sup>2</sup>**, as described below:

- **Models from the drc package:**

The **R<sup>2</sup>** value is computed using the variance-based formula:

$$R^2 = 1 - \frac{\text{Var}(\text{residuals})}{\text{Var}(\text{observed values})}$$

where the residual variance is divided by the variance of the observed response values, providing a measure of how well the model accounts for the data variability.

- **Exponential Model:**

The **R<sup>2</sup>** value is computed using the `modelr::rsquare()` function, which applies a standard squared correlation approach between observed and predicted values, ensuring consistency in nonlinear regression evaluation.

- **Hill Model:**

The **R<sup>2</sup>** value is computed similarly to the exponential model, using `modelr::rsquare()`. This function assesses the proportion of explained variance by comparing fitted values with observed responses.

- **Linear Model:**

For linear regression models, **R<sup>2</sup>** is derived from the standard formula used in `lm()` results:

$$R^2 = 1 - \frac{\sum(y_{\text{obs}} - y_{\text{pred}})^2}{\sum(y_{\text{obs}} - \bar{y})^2}$$

where  $y_{\text{obs}}$  represents the observed response,  $y_{\text{pred}}$  the predicted values, and  $\bar{y}$  the mean of the observed responses. The **R<sup>2</sup>** value is extracted directly from `summary(lm())$r.squared`.

- **Polynomial Models:**

Similar to the linear model, polynomial models use the standard **R<sup>2</sup>** computation from `summary(lm())$r.squared`. Since polynomials are fitted using linear regression techniques, their goodness-of-fit evaluation follows the same approach.

- **Power Model:**

The **R<sup>2</sup>** value for power models is computed using `modelr::rsquare()`, as in the exponential and Hill models. This function provides a robust measure of the proportion of variance explained by the fitted model.

## Interpretation

An **R<sup>2</sup>** value close to 1 indicates that the model explains a high proportion of the variability in the observed data, while an **R<sup>2</sup>** value close to 0 suggests that the model does not provide a good fit. It is important to note that **R<sup>2</sup>** does not inherently account for model complexity, so it should be interpreted alongside other model selection criteria, such as AIC.

## BMR Selection and BMD estimation

The BMR can be computed by means of three different methods: standard deviation, relative and absolute. The standard deviation method for normal distributed responses it is defined as

$$\frac{|m(BMD) - m(0)|}{s(0)} = BMRF$$

Where  $s(0)$  is the standard deviation of the controls, and  $m(0)$  is the predicted value of the response at the control level. The BMRF is the multiple of the standard deviation, whose default value 1.349 corresponds to

a 10% of difference with respect to the controls. Thus, the BMD (benchmark doses) is the estimated dose, whose predicted value  $m(BMD)$  corresponds to the desired BMRF.

## BMDL/BMDU Estimation

The **Benchmark Dose Lower Bound (BMDL)** and **Benchmark Dose Upper Bound (BMDU)** values are estimated using confidence interval methods as suggested by Gaylor et al. (1998). This approach provides an uncertainty range for the Benchmark Dose (BMD), ensuring that the estimated dose-response relationship accounts for statistical variability.

### Methodology

The estimation of BMDL and BMDU is based on the **confidence interval of the predicted dose-response curve**. Given a predefined confidence level (e.g., 95%), the **lower (BMDL)** and **upper (BMDU)** bounds are determined by interpolating the benchmark response (BMR) within the confidence interval of the fitted model.

#### 1. Computation of the Confidence Interval

- A confidence interval is constructed around the predicted response using the model's standard error estimates.
- The confidence level (e.g., 95%) defines the range within which the true dose-response relationship is expected to lie.

#### 2. Interpolation of BMDL and BMDU

- If the adverse direction of the response is **negative (-1)**, BMDL is computed as the **lower bound** of the confidence interval, while BMDU is the **upper bound**.
- If the adverse direction is **positive (+1)**, BMDL is interpolated from the **upper bound**, while BMDU is obtained from the **lower bound**.

#### 3. Handling Errors and Warnings

- If the confidence interval estimation fails due to numerical issues or model constraints, BMDL and BMDU are assigned **NA values**, indicating that the bounds could not be determined.

## Implementation in BMDx

BMDL and BMDU are computed using the `predict()` function on a range of dose values spanning the observed data. The interpolation is performed using the `stats::approx()` function to estimate the benchmark dose bounds at the benchmark response (BMR) level. The estimated values are then stored as **BMDL** and **BMDU** in the model output.

This methodology is aligned with the **U.S. Environmental Protection Agency (EPA) recommendations** for benchmark dose estimation, following the framework outlined by Gaylor et al. (1998) ([link](#)).

## Selection of Parameters for bmd modelling

The BMD analysis requires the following parameters:

- **Number of cores:** The number of cores that are used during the analysis
- **Maximum Iterations:** Is the maximum number of iterations allowed as a convergence criteria
- **Data type:** Is a string character describing the type of response. Possible values are continuous (default) and binomial
- **Deviation type:** Is a string character describing the type of deviation computed. Possible values are standard (default) and relative and absolute.

- **BMR factor:** Is the number of standard deviation at which the BMD is defined. The default value is 1.349, which corresponds to a change of 10% with respect to the controls. For further information refer to the following papers: Thomas, Russell S, Bruce C Allen, Andy Nong, Longlong Yang, Edilberto Bermudez, Harvey J Clewell, and Melvin E Andersen. 2007. “A Method to Integrate Benchmark Dose Estimates with Genomic Data to Assess the Functional Effects of Chemical Exposure.” *Toxicological Sciences* 98 (1): 240–48. <https://doi.org/10.1093/toxsci/kfm092>; Filipsson, Agneta Falk, Salomon Sand, John Nilsson, and Katarina Victorin. 2003. “The Benchmark Dose Method—Review of Available Models, and Recommendations for Application in Health Risk Assessment.” *Critical Reviews in Toxicology* 33 (5): 505–42. <https://doi.org/10.1080/10408440390242360>.
- **Confidence interval:** The statistical confidence limit applied to the BMD estimated from the models.
- **Variance type:** Is a string character describing the type of variance of the data. Possible values are constant (default), non constant, model and inferred.
- **Significance threshold for levene test:** In case of variance type equal to inferred, a levene test is used to test the assumption of constant or non constant variance. This parameter specify the threshold for the significant p-value for the test. Default value is 0.05.
- **Lack-of-fit p-value:** Threshold used to identify good fitted models. Default 0.1
- **Select models:** Predefined subselection of the available models. Possible values are custom, regulatory, degree of freedom, all. The user can use any of the available subset of models or select the models himself from the checkbox list. Regulatory models include models of interest for regulatory agencies. Degree of freedom models comprise models with a degree of freedom less than n-1, considering n is the number of doses tested. Custom allows for manual selection of desired models.

## BMD Results

The results of the BMD analysis will be reported in a tabular format with the following column names:

- **Experiment:** The name of the experiment
- **time\_point:** The time point of the experiment
- **Feature:** The name of the gene
- **Model:** The model fitted
- **BMR:** The expression of the gene at the response level
- **BMDL:** The estimated BMDL
- **BMD:** The estimated BMD
- **BMDU:** The estimated BMDU
- **AC50:** The estimated AC50
- **Adverse\_direction:** Indicate if the model is increasing along the doses (+1) or decreasing (-1)
- **AIC:** The Akaike information criterion computed for the model
- **Lack\_of\_fit:** The lack of fit p-value
- **Log\_likelihood:** The log-likelihood of the data given the model
- **R2:** The R2 square associated with the model
- **Residual\_variance:** The variance of the residual represent differences between observed values and the values predicted by the model. It represents how much of the variability of response is not explained by the model.

The table reports all the models that can be estimated for every gene.

## Filtering BMD Results

The user can discard the model fitted and displayed in the BMD analysis result table by means of the following criteria:

## Compute BMD Value

**Number of cores:** 1      **Maximum iterations:** 1024

**Data type:** continuous      **Deviation type:** standard      **Variance type:** constant

**BMR factor:** 10 (1.349)

**Significance threshold for levene test:** 0.05 (0.05 to 1.0)

**Confidece interval:** 0.05 (0.0 to 1.0)

**Lack-of-fit p-value:** 0.1 (0.0 to 1.0)

**Select models:** Custom

| Models available                           |                                           |                                   |                                   |
|--------------------------------------------|-------------------------------------------|-----------------------------------|-----------------------------------|
| <input checked="" type="checkbox"/> linear | <input checked="" type="checkbox"/> hill  | <input type="checkbox"/> llog2    | <input type="checkbox"/> weibul13 |
| <input checked="" type="checkbox"/> poly2  | <input checked="" type="checkbox"/> power | <input type="checkbox"/> llog3    | <input type="checkbox"/> weibul14 |
| <input type="checkbox"/> poly3             | <input checked="" type="checkbox"/> exp2  | <input type="checkbox"/> llog4    | <input type="checkbox"/> weibul22 |
| <input type="checkbox"/> poly4             | <input type="checkbox"/> exp3             | <input type="checkbox"/> llog5    | <input type="checkbox"/> weibul23 |
| <input type="checkbox"/> poly5             | <input type="checkbox"/> exp4             | <input type="checkbox"/> mm2      | <input type="checkbox"/> weibul24 |
|                                            | <input type="checkbox"/> exp5             | <input type="checkbox"/> weibul12 |                                   |

**Run BMD**

Figure 6: BMD parameters.

Experiment  Time Points

Click on the rows of the table to show the corresponding fitted models

Show 10 entries Search:

|    | Experiment | timepoint | Feature         | Model  | BMR       | BMDL      | BMD       | BMDU      | AC50       | Adverse_direction | AIC       | Lack_of_fit_p_value |
|----|------------|-----------|-----------------|--------|-----------|-----------|-----------|-----------|------------|-------------------|-----------|---------------------|
|    | All        | All       | All             | /      | /         | /         | /         | /         | AI         | All               | /         | All                 |
| 1  | bleomycin  | 24        | ENSG00000188290 | linear | 6.097404  | 18.095102 | 32.805239 | 43.734522 | 50         | 1                 | 14.233876 | 0.004               |
| 2  | bleomycin  | 24        | ENSG00000188290 | poly2  | 5.887715  | 5.630519  | 15.39189  | 23.434859 | 28.2362    | 1                 | 11.012088 | 0.012               |
| 3  | bleomycin  | 24        | ENSG00000188290 | hill   | 5.920416  | 3.27572   | 21.193367 | 34.671517 | 220.403205 | 1                 | 15.45735  | 0.002               |
| 4  | bleomycin  | 24        | ENSG00000188290 | power  | 5.651302  | 0.029005  | 0.568764  | 4.96573   | 9.988693   | 1                 | 1.165727  | 0.353               |
| 5  | bleomycin  | 24        | ENSG00000217801 | linear | 7.873942  | 30.411676 | 48.873847 | 66.701306 | 50         | 1                 | 16.171739 | 0.001               |
| 6  | bleomycin  | 24        | ENSG00000217801 | poly2  | 7.777058  | 14.669109 | 30.816783 | 52.586038 | 32.602375  | 1                 | 17.002038 | 0.001               |
| 7  | bleomycin  | 24        | ENSG00000217801 | hill   | 7.48913   | 4.705655  | 0.948243  | 16.694121 | 1.611422   | 1                 | 11.510631 | 0.007               |
| 8  | bleomycin  | 24        | ENSG00000217801 | power  | 7.485844  | 0.009794  | 0.11275   | 8.944299  | 1.237353   | 1                 | 8.188425  | 0.029               |
| 9  | bleomycin  | 24        | ENSG00000142583 | linear | 11.289318 | 27.387539 | 45.118891 | 60.599389 | 50         | -1                | 10.204342 | 0.000               |
| 10 | bleomycin  | 24        | ENSG00000142583 | poly2  | 11.612115 | 3.724548  | 11.408118 | 17.521388 | 21.247459  | -1                | -1.625609 | 0.03                |

Showing 1 to 10 of 4,670 entries Previous  2 3 4 5 ... 467 Next

Figure 7: BMD results

- **Filter NAs:** If selected, every model for which any of the estimated doses were not computed will be removed
- **Filter by lack-of-fit:** If selected, every model for which any of the estimated doses were not computed will be removed
  - **Threshold for the lack-of-fit p-value** All models with lack-of-fit p-value lower than this threshold will be removed. Default = 0.1
- **Apply ratio filters:** If selected, all models with BMD/BMDL or BMDU/BMD or BMDU/BMDL ration greater than the specified threshold will be removed
  - **BMD/BMDL ratio:** Default parameter set to 20
  - **BMDU/BMD ratio:** Default parameter set to 20
  - **BMDU/BMDL ratio:** Default parameter set to 40
- **Filter by lower boundary:** If selected, models with estimated doses lower than a certain percentage of the lowest tested doses will be removed
  - **Lowest dose filter:** Default parameter set to 0.1, meaning that the model will be removed if it's effective doses are predicted to be lower than 10% of the lowest tested dose
- **Filter by upper boundary:** If selected, models with estimated doses higher than a certain percentage of the highest tested doses will be removed
  - **Lowest dose filter:** Default parameter set to 0.1, meaning that the model will be removed if it's effective doses are predicted to be higher than 10% of the highest tested dose
- **Filter by R2:** If selected, models with R2 lower than a certain threshold will be removed
  - **Threshold R2:** Default parameter set to 0.6.
- **only strictly monotonic models allowed** If selected, only models that are strictly monotonic will be allowed.
- **No negative effective doses** If selected, only models that do not have negative effective doses will be allowed.
- **No unordered effective doses** If selected, only models that have effective doses ordered will be allowed.

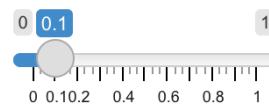
**Number of cores:**

1

Filter NAs

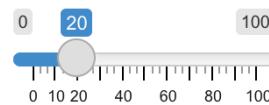
Filter by Lack of Fit

**Threshold for the Lack  
of Fit Pvalue:**

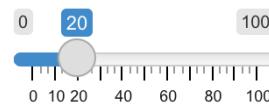


Apply Ratio filters

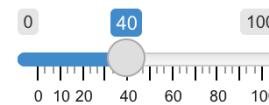
**BMD/BMDL ratio:**



**BMDU/BMD ratio:**



**BMDU/BMDL ratio:**



Filter by Lower  
Boundary

**Lowest dose filter:**



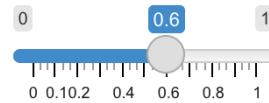
Filter by Upper  
Boundary

**Highest dose filter:**



Filter models with  
R2  |

**Threshold R2:**



Only strictly  
monotonic models  
allowed

**Apply Filters**

Figure 8: Filtering of BMD results.

## Selection of the optimal model

The selection of the final optimal model for every gene can be selected as the one with the lowest AIC value, or as the average consensus model. The consensus model should be computed only merging the models that pass the filtering criteria.

## Visual investigation of the models

When clicking on a row of the BMD result table, the model fitted to the response data of the corresponding gene is visualized along with the corresponding effective doses. Furthermore the model formula with the estimated parameters are also visualised.

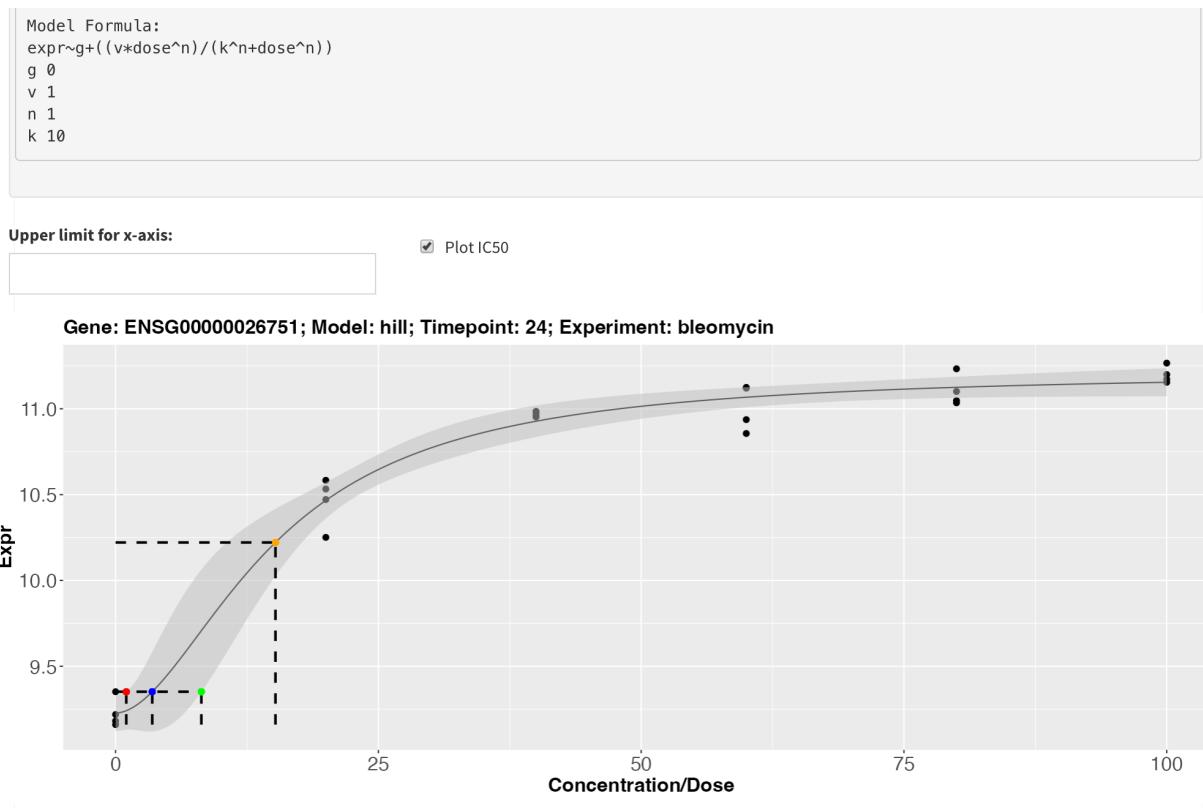


Figure 9: BMD model fitting of a specific gene.

#Compare different experiments

Once the models have been fitted their distribution across different experiments and different time-points can be explored by different visualizations.

## BMD histograms

The density distribution of the BMD values can be investigated by means of histograms. The user can select the relevant variable of interest and different grouping variables. Filtering can be also performed to plot only a subset of the results.

The following parameters are required:

- **Color by:** how the bmd values are going to be colored. This should be a factor variable whose levels will be used to fill the legend
- **Group by X:** the bmd values can be grouped by a factor variable

- **Group by Y (optional):** the bmd values can be grouped by another optional factor variable to perform a grid plot
- **Filter by Column:** Can be used to investigate the bmd distribution of specific variable of interest. Default is None which means that no filtering is applied. However, if one of the column names is selected the **Filter by Column Values** window is populated by the possible column values. When one of those is selected, only the bmd corresponding to the genes characterized by those values are investigated (e.g. only the bmd produced by a specific model).

|                                                   |                                |                              |
|---------------------------------------------------|--------------------------------|------------------------------|
| <b>Color By</b>                                   | <b>Group By X</b>              | <b>Group By Y (optional)</b> |
| Model                                             | Experiment                     | None                         |
| <b>Filter By Column</b>                           | <b>Filter By Column Values</b> | Plot                         |
| <input type="button" value="None"/><br>Experiment |                                |                              |
| Model                                             |                                |                              |
| time_point                                        |                                |                              |

Figure 10: Parameters for plotting BMD histogram.

### Fitting statistics

The distribution of the lack-of-fit p-values, AIC, and  $R^2$  can be investigated by means of histograms. The user can select the relevant variable of interest and different grouping variables. Filtering can be also performed to plot only a subset of the results.

The following parameters are required:

- **Color by:** how the bmd values are going to be colored. This should be a factor variable whose levels will be used to fill the legend
- **Group by X:** the bmd values can be grouped by a factor variable
- **Group by Y (optional):** the bmd values can be grouped by another optional factor variable to perform a grid plot
- **Filter by Column:** Can be used to investigate the bmd distribution of specific variable of interest. Default is None which means that no filtering is applied. However, if one of the column names is selected the **Filter by Column Values** window is populated by the possible column values. When one of those is selected, only the bmd corresponding to the genes characterized by those values are investigated (e.g. only the bmd produced by a specific model).

### BMD/BMDL

The relationship between the BMD/BMDL values can be investigated by means of scatterplots. The user can select the relevant variable of interest and different grouping variables. Filtering can be also performed to plot only a subset of the results. The following parameters are required:

- **Color by:** how the bmd values are going to be colored. This should be a factor variable whose levels will be used to fill the legend
- **Group by X:** the bmd values can be grouped by a factor variable

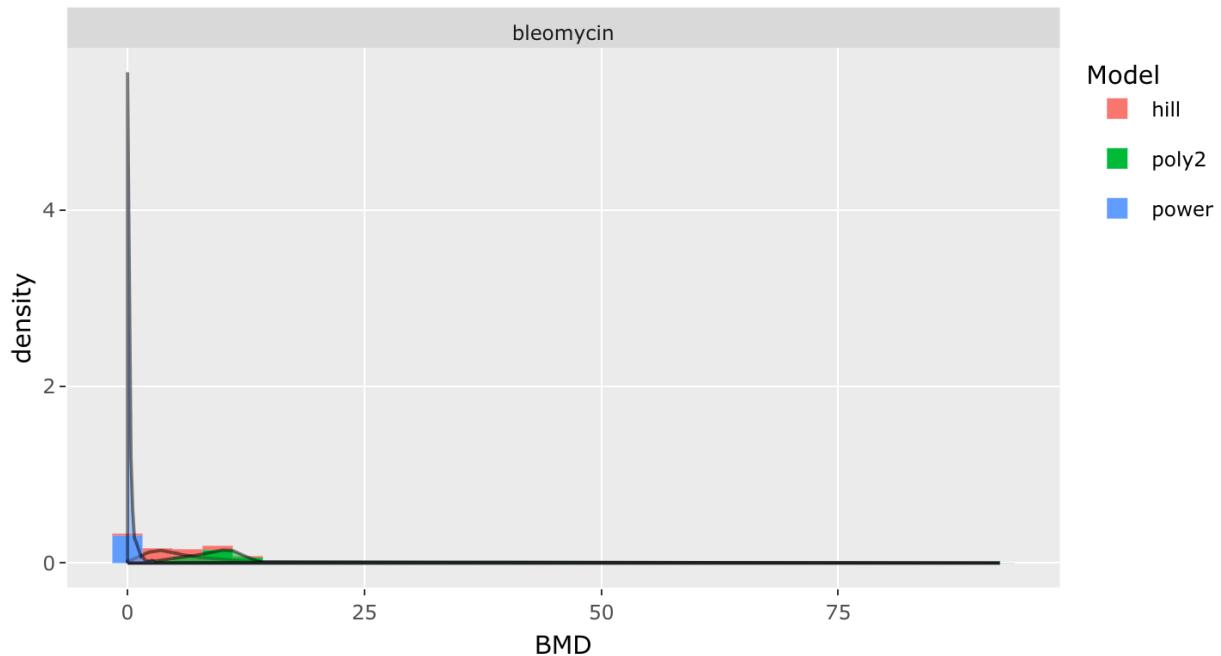


Figure 11: BMD density

|                                                                        |                                |                              |
|------------------------------------------------------------------------|--------------------------------|------------------------------|
| <b>Color By</b>                                                        | <b>Group By X</b>              | <b>Group By Y (optional)</b> |
| Model                                                                  | Experiment                     | None                         |
| <b>Filter By Column</b>                                                | <b>Filter By Column Values</b> | <b>Plot</b>                  |
| Experiment<br>Model<br>Adverse_direction<br>is_monotonic<br>timenpoint | bleomycin                      |                              |

Figure 12: Parameters for plotting the distribution of the lack-of-fit p-values

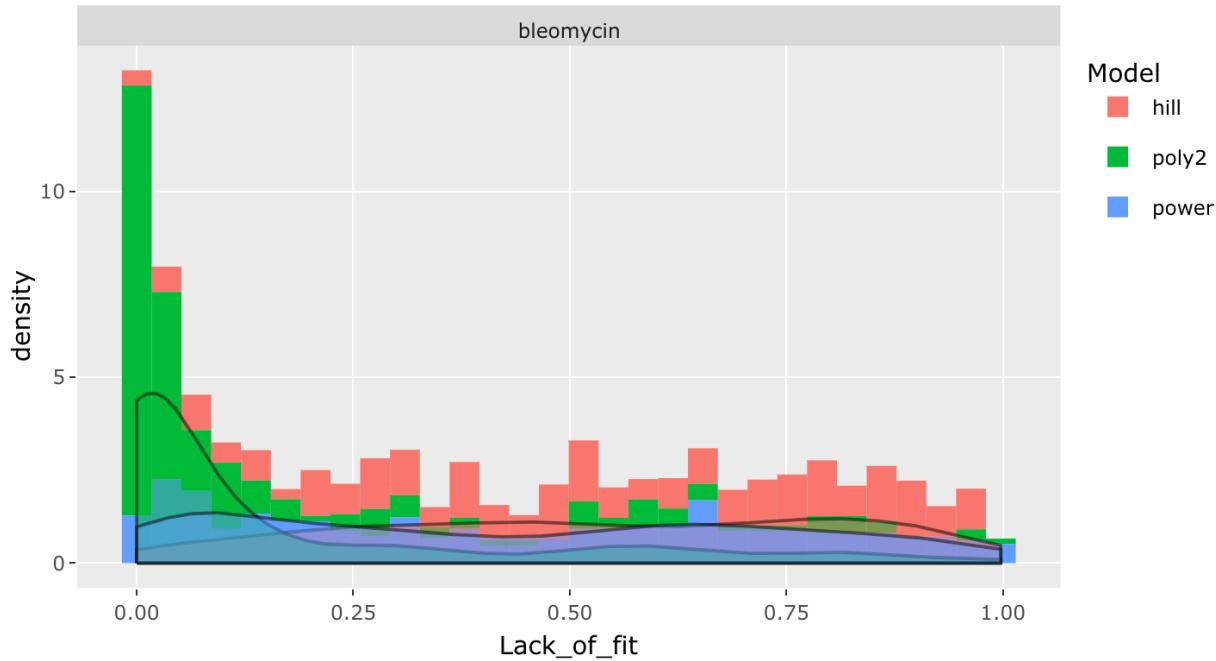


Figure 13: Lack-of-fit.

- **Group by Y (optional):** the bmd values can be grouped by another optional factor variable to perform a grid plot
- **Filter by Column:** Can be used to investigate the bmd distribution of specific variable of interest. Default is None which means that no filtering is applied. However, if one of the column names is selected the **Filter by Column Values** window is populated by the possible column values. When one of those is selected, only the bmd corresponding to the genes characterized by those values are investigated (e.g. only the bmd produced by a specific model).

|                                                  |                                |                              |
|--------------------------------------------------|--------------------------------|------------------------------|
| <b>Color By</b>                                  | <b>Group By X</b>              | <b>Group By Y (optional)</b> |
| Model                                            | Experiment                     | None                         |
| <b>Filter By Column</b>                          | <b>Filter By Column Values</b> | Plot                         |
| None<br>Experiment<br>Model<br>Adverse_direction | bleomycin                      |                              |

Figure 14: Parameters for plotting the relationship between BMD/BMDL.

### Fitted models

The proportion of fitted models can be investigated by means of pie charts. The user can select different grouping variables. Filtering can be also performed to plot only a subset of the results. In this plot the variable of interest is fixed to the models and cannot be changed by the user.

The following parameters are required:

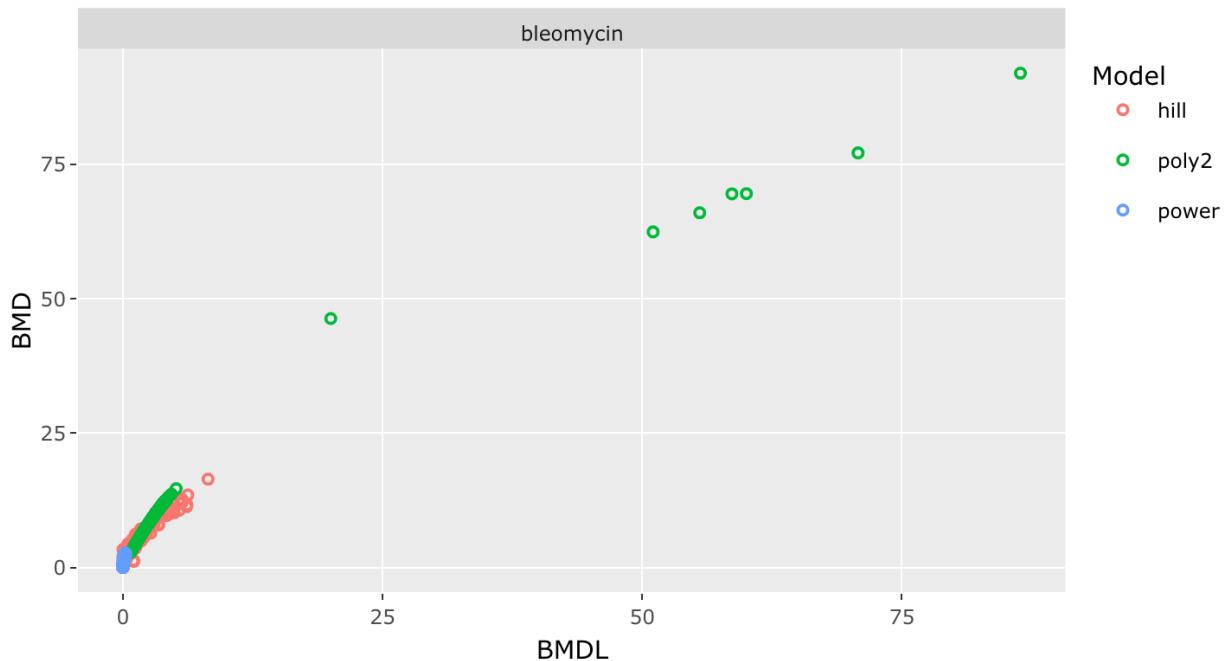


Figure 15: Relationship between the BMD/BMDL values

- **Group by X:** the bmd values can be grouped by a factor variable
- **Group by Y (optional):** the bmd values can be grouped by another optional factor variable to perform a grid plot
- **Filter by Column:** Can be used to investigate the bmd distribution of specific variable of interest. Default is None which means that no filtering is applied. However, if one of the column names is selected the **Filter by Column Values** window is populated by the possible column values. When one of those is selected, only the bmd corresponding to the genes characterized by those values are investigated (e.g. only the bmd produced by a specific model).

### Number of dose-dependent genes

The proportion of dose-dependent genes by time points can be investigated by means of barplots. The user can select the grouping variable. Filtering can be also performed to plot only a subset of the results. In this plot the variable of interest is fixed to the time point and cannot be changed by the user.

The following parameters are required:

- **Group by X:** the bmd values can be grouped by a factor variable
- **Filter by Column:** Can be used to investigate the bmd distribution of specific variable of interest. Default is None which means that no filtering is applied. However, if one of the column names is selected the **Filter by Column Values** window is populated by the possible column values. When one of those is selected, only the bmd corresponding to the genes characterized by those values are investigated (e.g. only the bmd produced by a specific model).

### ECDF

The distribution of the bmd values at the whole transcriptome level can be investigated by means of the ECDF plot. This plot is showing the percentage of genes (y-axis) with a bmd lower than a certain value (x-axis). The user can select the relevant variable of interest and different grouping variables. Filtering can be also performed to plot only a subset of the results. The following parameters are required:

**Group By X**

Experiment ▾

**Group By Y (optional)**

None ▾

**Filter By Column**

None

Experiment  
Model  
time\_point

**Filter By Column Values**

Plot

Figure 16: Parameters to compute the proportion of fitted models.

- **Relevant variable:** how the bmd values are going to be colored. This should be a factor variable whose levels will be used to fill the legend
- **Group by X:** the bmd values can be grouped by a factor variable
- **Numeric grouping variable:** If true the variable is considered numerical and ordered accordingly
- **Other variables (optional):** the bmd values can be grouped by another optional factor variable to perform a grid plot
- **Filter by Column:** Can be used to investigate the bmd distribution of specific variable of interest. Default is None which means that no filtering is applied. However, if one of the column names is selected the **Filter by Column Values** window is populated by the possible column values. When one of those is selected, only the bmd corresponding to the genes characterized by those values are investigated (e.g. only the bmd produced by a specific model).
- **Scale bmd:** If selected, the bmd of different experiments (that might be in different ranges) are scaled in the range [0-1] to make them comparable.

### Upset Plot

The number of genes shared by the different experiments can be investigated by means of an upset plot. The user can select the relevant variable of interest and different grouping variables. Filtering can be also performed to plot only a subset of the results. The following parameters are required:

- **Relevant variable:** how the bmd values are going to be colored. This should be a factor variable whose levels will be used to fill the legend
- **Group by X:** the bmd values can be grouped by a factor variable
- **Numeric grouping variable:** If true the variable is considered numerical and ordered accordingly
- **Other variables (optional):** the bmd values can be grouped by another optional factor variable to perform a grid plot
- **Filter by Column:** Can be used to investigate the bmd distribution of specific variable of interest. Default is None which means that no filtering is applied. However, if one of the column names is selected the **Filter by Column Values** window is populated by the possible column values. When one of those is selected, only the bmd corresponding to the genes characterized by those values are investigated (e.g. only the bmd produced by a specific model).
- **Maximum intersection:** Maximum number of intersection to show

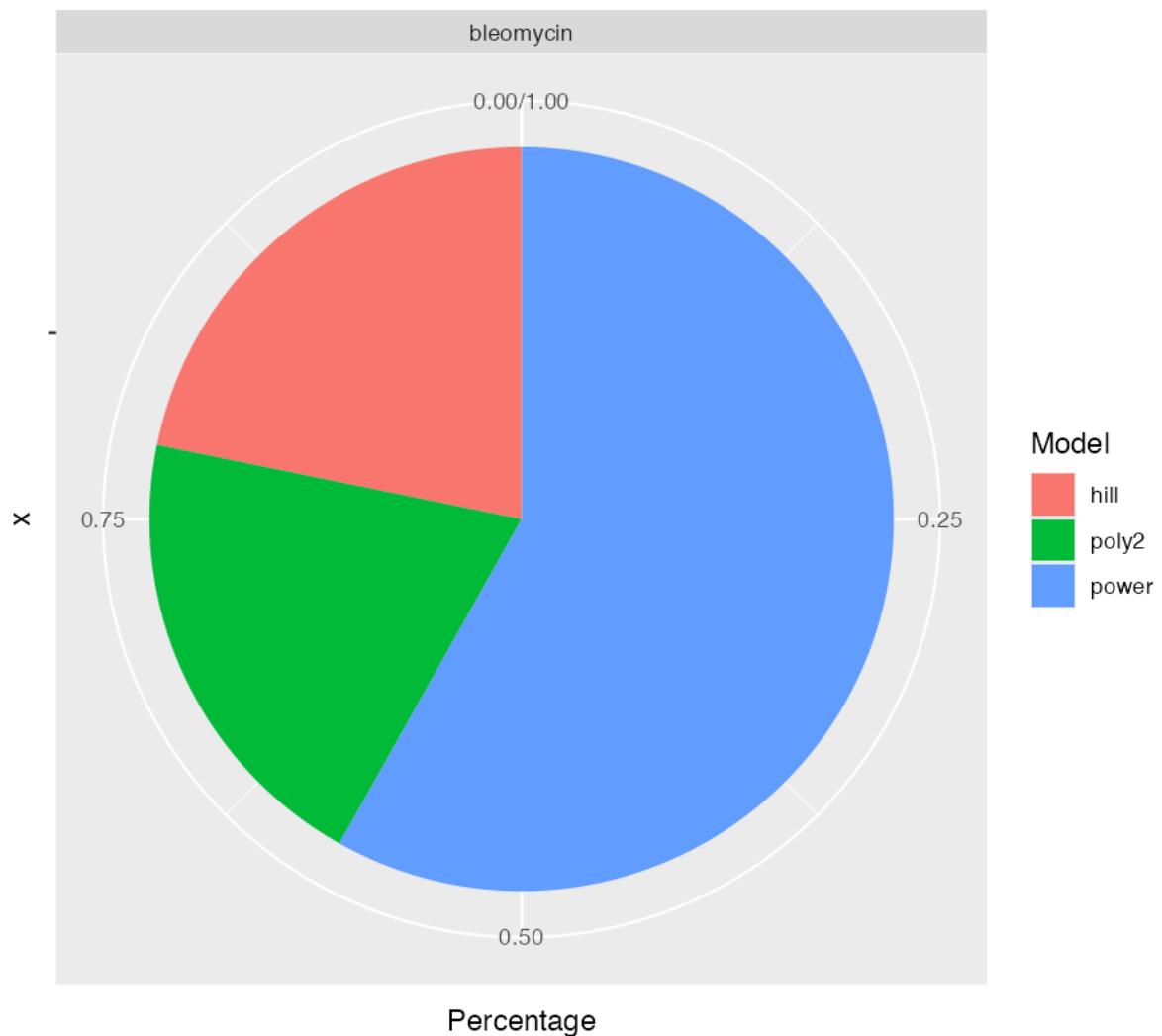


Figure 17: Proportion of fitted models to compute BMD.

**Group By X**

Experiment ▼

**Plot**

**Filter By Column**

None

Experiment

Model

time\_point

**Filter By Column Values**

Figure 18: Parameters for plotting dose-dependent genes

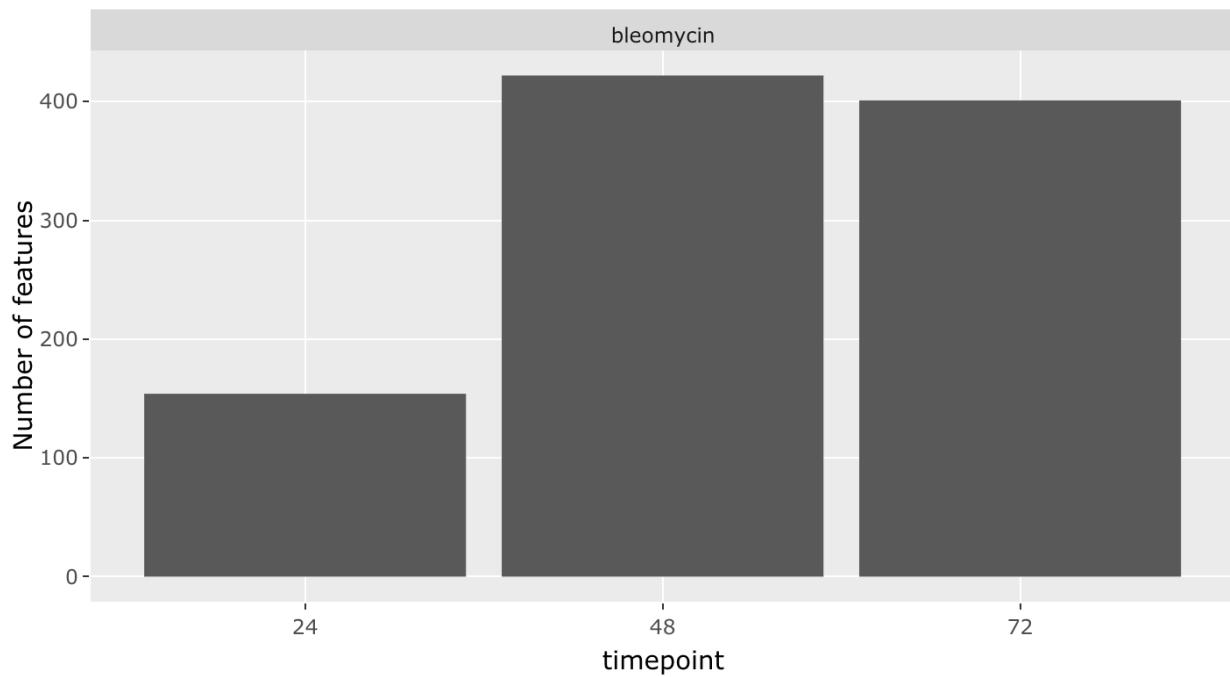


Figure 19: Proportion of dose-dependent genes by time points

**Relevant Variable** time\_point ▼    **Group By** None ▼     numeric grouping variable

**Other Variables**

|            |
|------------|
| None       |
| Experiment |
| Model      |
| time_point |

**Filter By Column**

|            |
|------------|
| None       |
| Experiment |
| Model      |
| time_point |

**Filter By Column Values**

Scale bmd

**Plot**

Figure 20: Parameters for plotting ECDF.

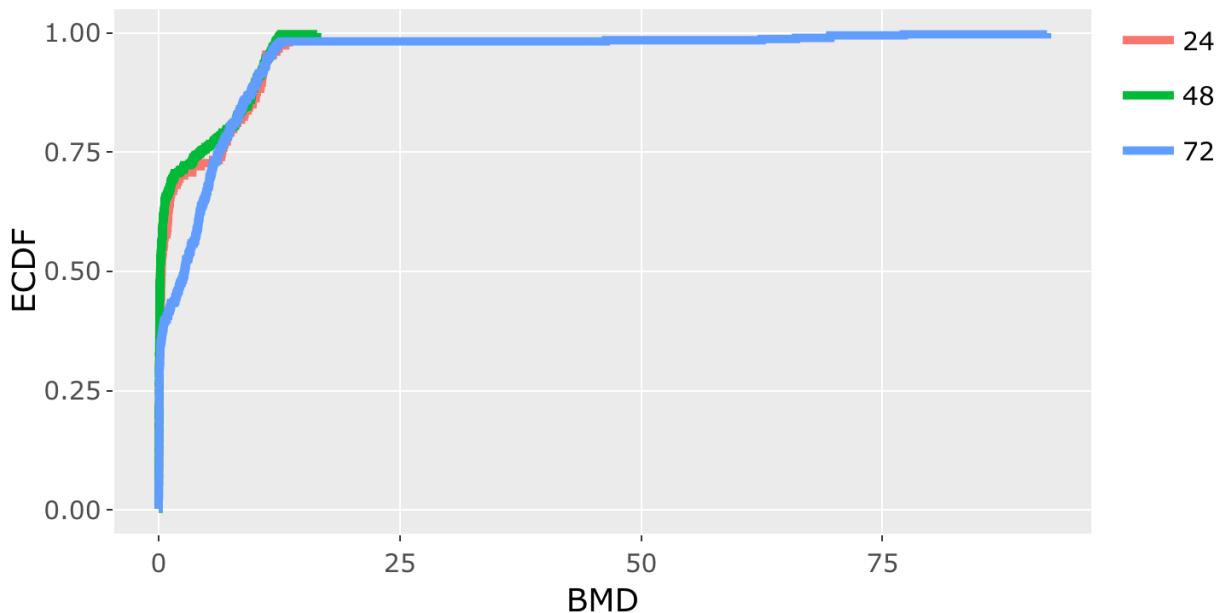


Figure 21: Distribution of the bmd values at the whole transcriptome level.

- **GroupBy:** UpsetPlot parameter
- **OrderBy:** UpsetPlot parameter

**Relevant Variable**

time\_point

**Group By**

None

**Other Variables**

None

**Filter By Column**

None

**Filter By Column Values**

**Maximum intersection:**

1
10
100

**GroupBy:**

degree

**OrderBy:**

frequency

Plot

Figure 22: Parameters for plotting upset plot.

## Gene frequency analysis

The aim of this analysis is to identify genes who are more often dose-dependent across different experiments or time points. The genes are ranked according to their occurrence frequency across the different experiments. Genes with a frequency below a certain threshold can be removed from the analysis. The tool compute the frequencies of the genes and plots them using a horizontal barplot, where genes are ranked based on their frequency. A GSEA is performed by means of the gprofiler package on the ranked list of genes.

The following parameters are required:

- **Main grouping variable:** Use to group the samples into categories. Gene frequency will be computed accordingly.
- **Second grouping variable:** Use to group the samples into categories. Gene frequency will be computed accordingly.
- **Threshold:** Used to filter out genes with low frequency. E.g. if 50% is selected, the genes with frequency below this number will not be included in further analysis
- **Organism:** to be specified for the GSEA analysis. Available organisms are human, mouse and rat.
- **Correction method:** Is the correction method applied to the GSEA p-values
- **Choose split:** If a split is performed this parameter can be used to show the GSEA results of the

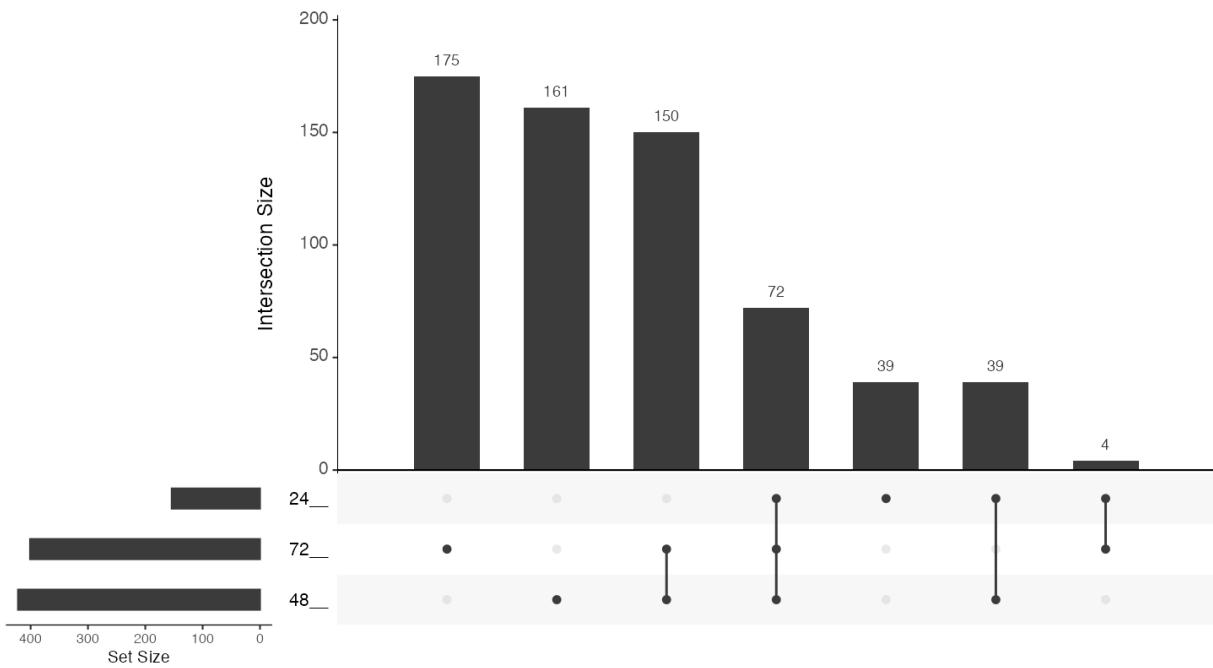


Figure 23: Number of genes shared by the different experiments.

specific subset of samples

## FunMappOne enrichment analysis

FunMappOne is a user friendly tool for the comparison of functional enrichment analysis af multiple experiments. The tool is publicly available at. Please refer to the FunMappOne paper for more details.

The FunMappOne tools is embedded into the BMDx as a module for the functional characterization of the dose-dependent genes.

### Heatmap visualization

Once the FunMappOne analysis is performed a heatmap with the enrichment results for each experiment is visualized. The enrichment heatmap shows each time point in each experiment as a separate column. Pathways are shown in the rows, and coloured boxes indicate enrichment of the pathway at the specific condition. When plotted values are shown in chromatic scale, the colour of the box changes according to the value. For example, in the figure below, the colour indicates the mean BMD value of the genes contributing to the enrichment of the pathway, showing the difference in the BMD values between the two experiments.

### Range plot

A range plot can be visualised to show the span of values of the BMDL, BMD and BMDU of the different enriched pathways. The following parameters are required:

- **Groupd by X:** the pathways can be grouped by a factor variable
- **Numeric grouping variable:** If true the variable is considered numerical and ordered accordingly
- **Other variables (optional):** the pathways can be grouped by another optional factor variable to perform a grid plot
- **Filter by Column:** Can be used to investigate the bmd distribution of specific variable of interest. Default is None which means that no filtering is applied. However, if one of the column names is

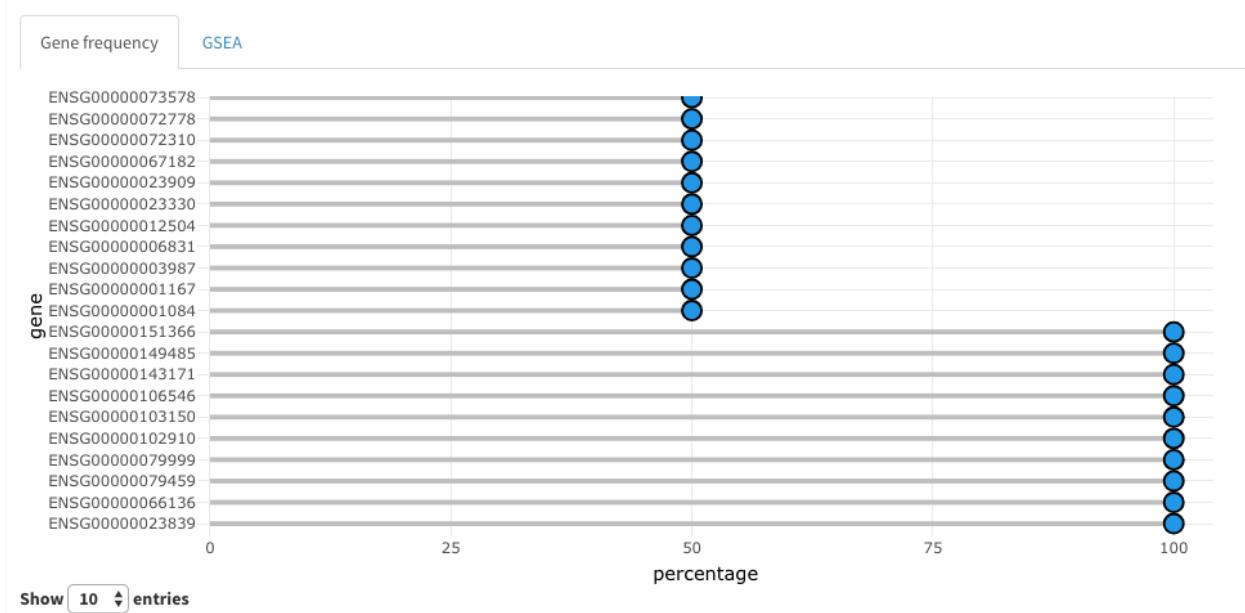


Figure 24: Lollipop gene frequency

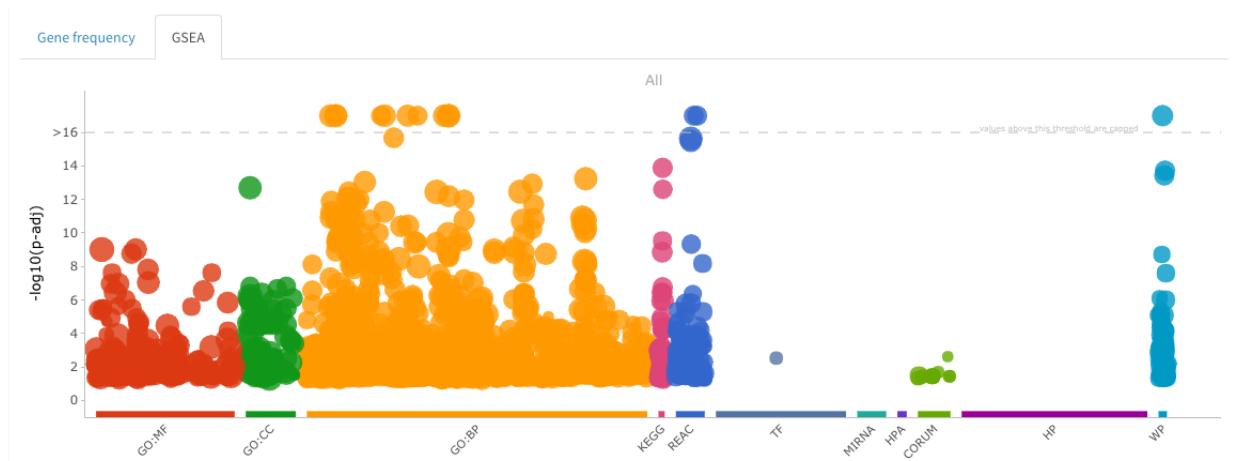


Figure 25: gsea

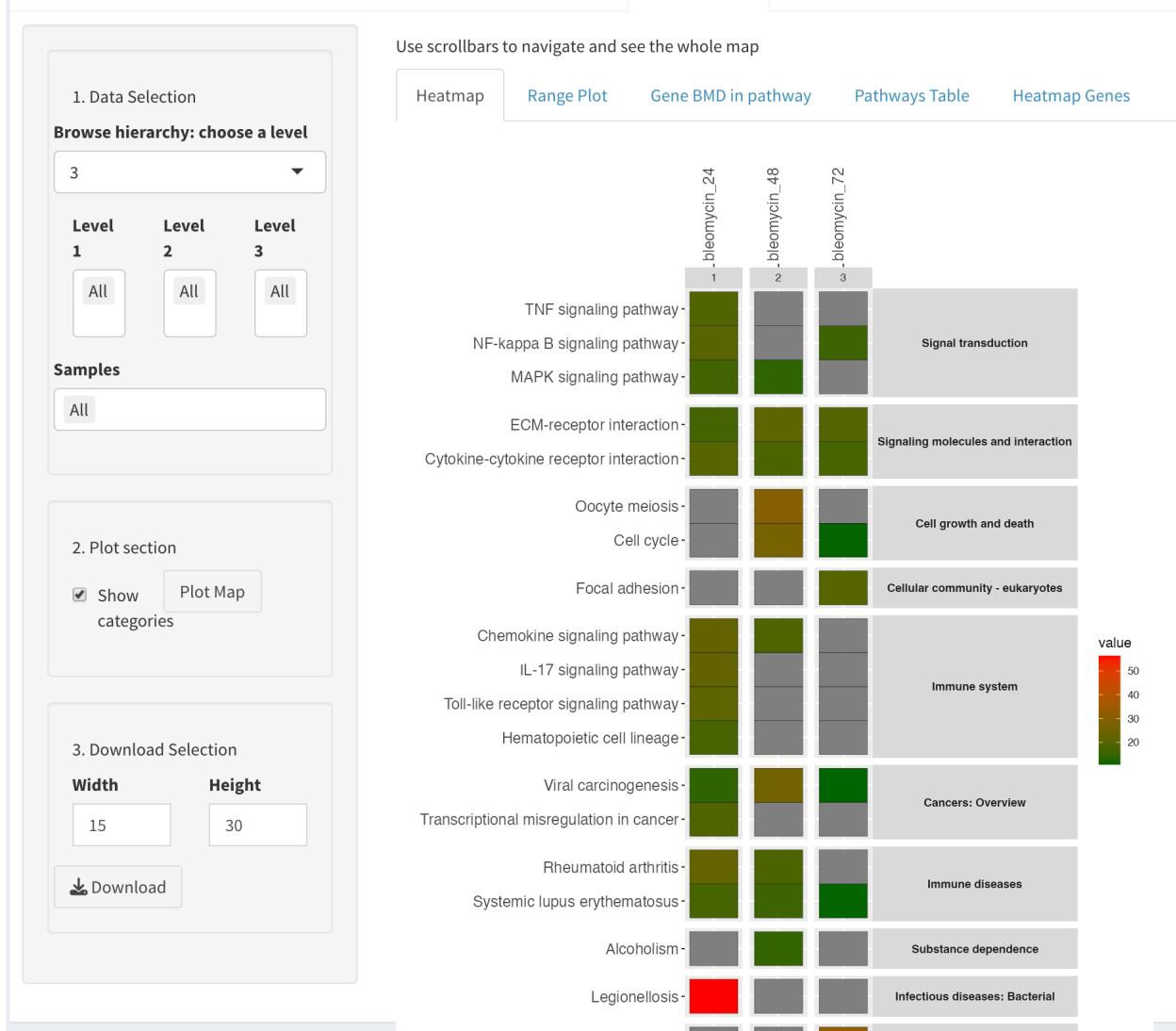


Figure 26: Heatmap of enrichment results from FunMappOne

selected the **Filter by Column Values** window is populated by the possible column values. When one of those is selected, only the bmd corresponding to the genes characterized by those values are investigated (e.g. only the bmd produced by a specific model).

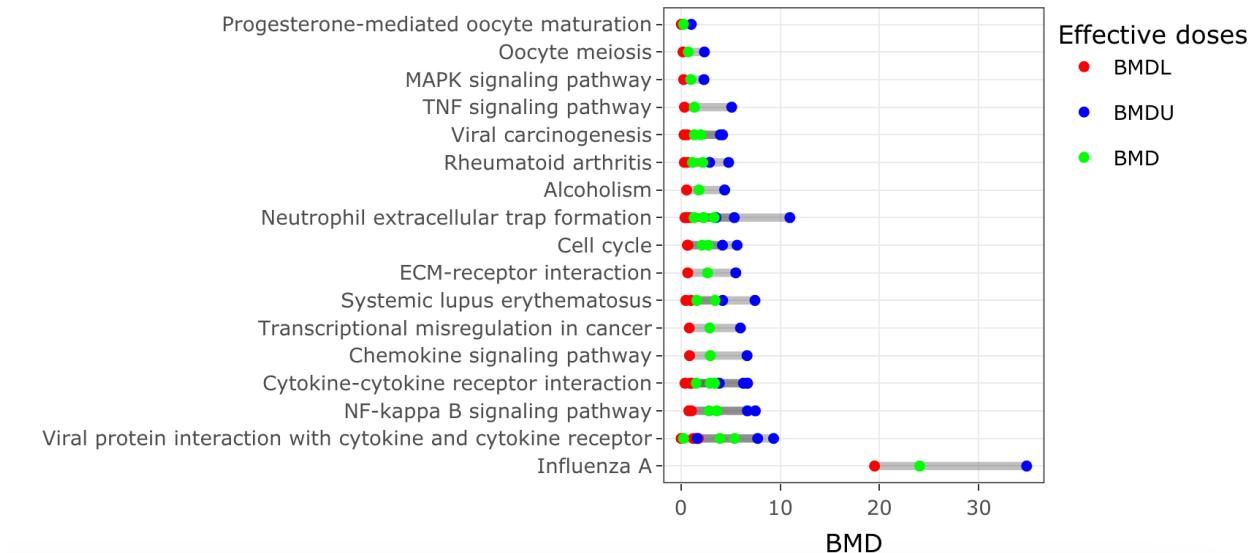


Figure 27: Range plot BMD, BMDL, and BMDU values

### Gene BMD in pathway

Gene BMD in a Pathway tab shows all pathways with their enrichment p-values. Selecting a row of the table plots a graph under the table with all the genes in the pathway, their BMD values as well as the lower and upper confidence bound BMD. Note! Deselect the row before selecting the next row to only include the genes in the desired pathway.

### Pathways Table

The genes mapped to each term are shown as a table in Pathways table tab. Time point for which data is shown can be changed from the time point drop menu. The tables can be downloaded as a single Excel file by clicking download.

### Heatmap Genes

Finally, the genes in different pathways can be explored in the form of a heatmap. The drop menus allow for the specification of the hierarchy level, the terms belonging to that level and specification of values shown on the heatmap. The heatmap then shows the genes mapped to the selected term on rows with the experiments and time points as columns, and coloured boxes indicating the value specified in the Show values drop menu.

## Compare gene pairs profile across doses

This analysis is used to compare the expression profiles of pairs of genes across the doses. Users can filter by specific columns, experiments, and timepoints. Given the optimal model fitted to every gene, N predicted expression values are computed. These are used to compute the Pearson correlation between each pair of genes. The parameters required for the analysis are the following:

|   |       |                         |                         |                                                               |
|---|-------|-------------------------|-------------------------|---------------------------------------------------------------|
| 1 | 04061 | 0.000001303325443192978 | 0.000001303325443192978 | Viral protein interaction with cytokine and cytokine receptor |
| 2 | 04060 | 0.0001036589419989639   | 0.0001036589419989639   | Cytokine-cytokine receptor interaction                        |
| 3 | 05323 | 0.0006677558007263812   | 0.0006677558007263812   | Rheumatoid arthritis                                          |
| 4 | 04064 | 0.001992577273634806    | 0.001992577273634806    | NF-kappa B signaling pathway                                  |
| 5 | 04668 | 0.004185330650154644    | 0.004185330650154644    | TNF signaling pathway                                         |
| 6 | 04062 | 0.0060020328701073      | 0.0060020328701073      | Chemokine signaling pathway                                   |
| 7 | 04613 | 0.02864597470622191     | 0.02864597470622191     | Neutrophil extracellular trap formation                       |
| 8 | 05202 | 0.03345663463035942     | 0.03345663463035942     | Transcriptional misregulation in cancer                       |
| 9 | 04512 | 0.04999665135818998     | 0.04999665135818998     | ECM-receptor interaction                                      |

Showing 1 to 9 of 9 entries

Previous 1 Next

**Experiment: bleomycin; Timepoint: 24  
ECM-receptor interaction**

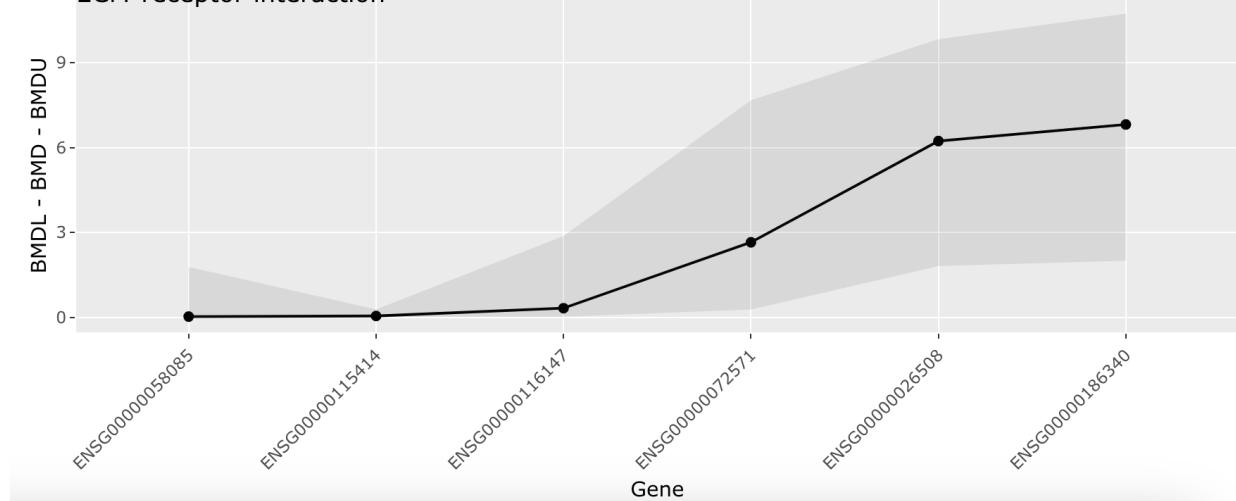


Figure 28: BMD values of genes belonging to a selected pathway

[Heatmap](#)
[Range Plot](#)
[Gene BMD in pathway](#)
Pathways Table
[Heatmap Genes](#)

**Time Points**

bleomycin\_24 ▾

[Download](#)

Show 10 entries

Search:

|   | Description                                                   | annID | gID                                                                                              |
|---|---------------------------------------------------------------|-------|--------------------------------------------------------------------------------------------------|
|   | All                                                           |       | All                                                                                              |
| 1 | Viral protein interaction with cytokine and cytokine receptor | 04061 | ENSG00000107562,ENSG00000126353,ENSG00000137077,ENSG00000144476,ENSG00000163739,ENSG00000168329, |
| 2 | Cytokine-cytokine receptor interaction                        | 04060 | ENSG00000099985,ENSG00000107562,ENSG00000125378,ENSG00000125538,ENSG00000126353,ENSG00000128342, |
| 3 | Rheumatoid arthritis                                          | 05323 | ENSG00000107562,ENSG00000112715,ENSG00000125538,ENSG00000163739,ENSG00000169429,ENSG00000184371, |
| 4 | NF-kappa B signaling pathway                                  | 04064 | ENSG00000056558,ENSG00000104856,ENSG00000107562,ENSG00000125538,ENSG00000137077,ENSG00000163739, |
| 5 | TNF signaling pathway                                         | 04668 | ENSG00000056558,ENSG00000069399,ENSG00000125538,ENSG00000128342,ENSG00000146592,ENSG00000163739, |
| 6 | Chemokine signaling pathway                                   | 04062 | ENSG00000107562,ENSG00000126353,ENSG00000137077,ENSG00000158517,ENSG00000163739,ENSG00000168329, |

Figure 29: Table including enriched pathways

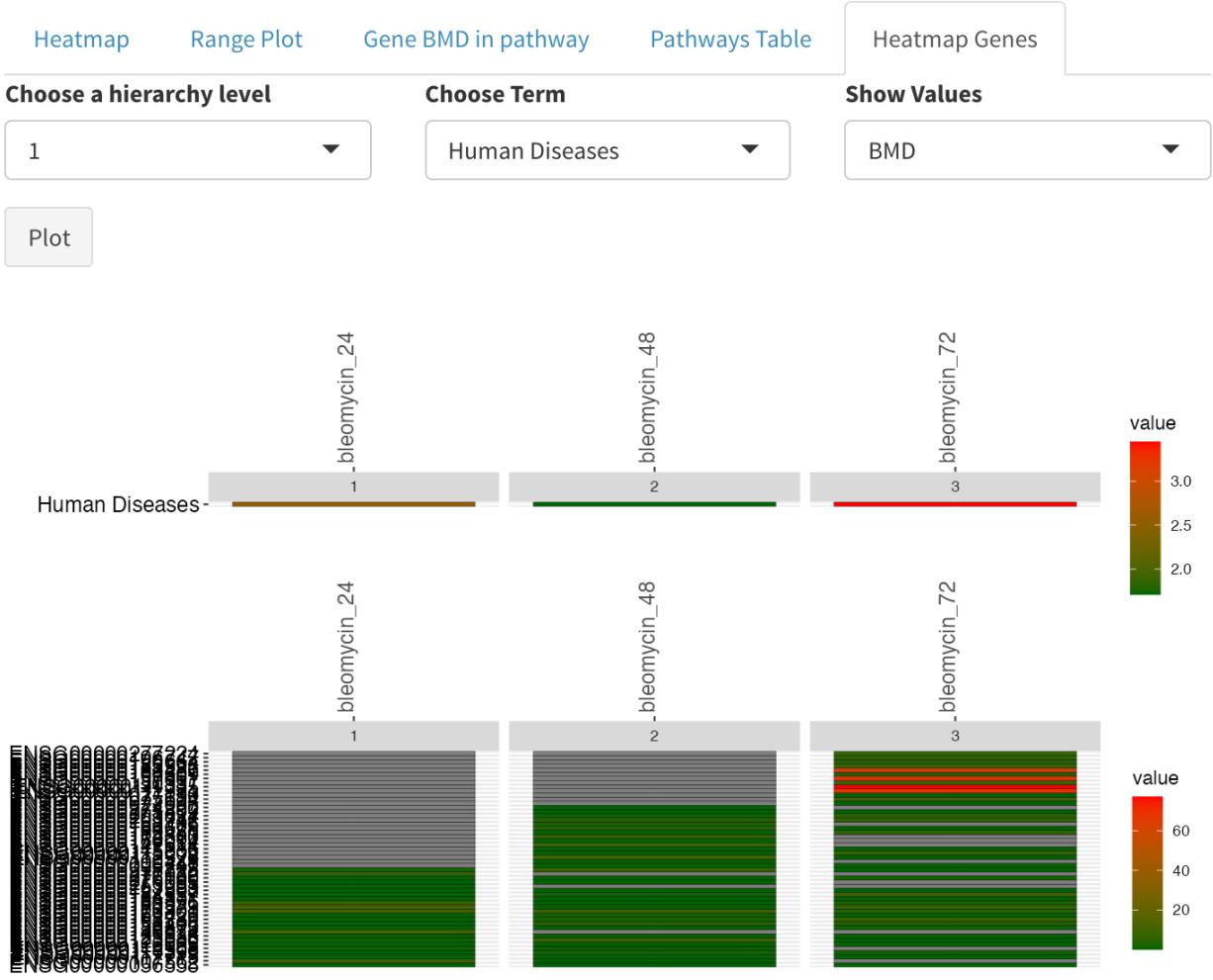


Figure 30: BMD of altered genes across experiments

Compare genes X

---

**Filter By Column**

**Filter By Experiment**

**Filter By Column**

**Filter By Time Point**

**organism:**

**Number of predictions:**

**Number of cores:**

---

- **Number of predictions:** The number of predicted values for each gene. Default = 1000 - **Number of cores:** The number of cores used for the analysis - **Organism:** The organism for which the expression values have been computed (available values are human, mouse and rat)

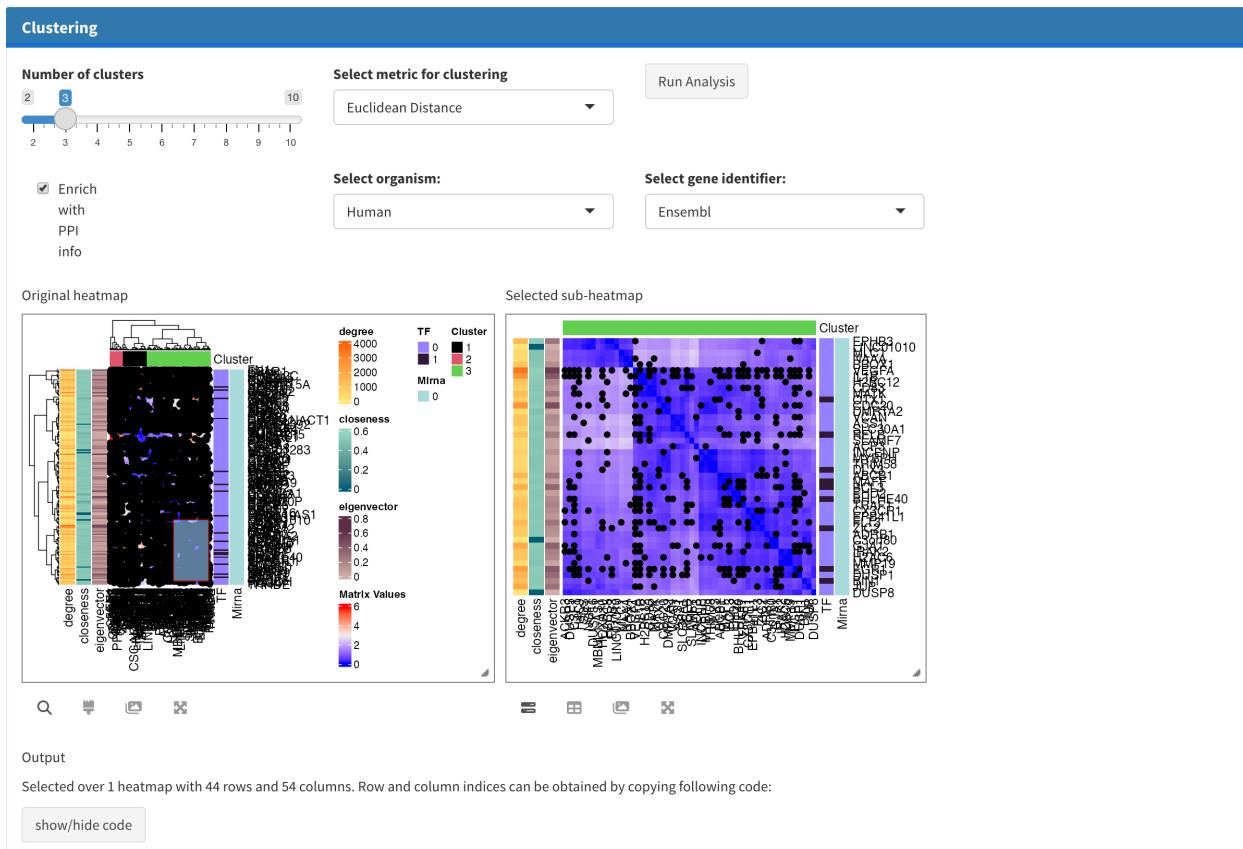
The result of the analysis are visualized in a tabular format with the following columns:

- **Filter by Column:** Fixed to “Experiment”
- **Filter by Experiment:** Used to select the experiment for which the correlation will be computed.
- **Filter by Time Point:** Used to select the time point for which the correlation will be computed.
- **Feature 1:** Name of the first gene
- **Model 1:** Optimal fitted model for the fist gene
- **Feature 2:** Name of the second gene
- **Model 2:** Optimal fitted model for the second gene
- **Pearson Correlation:** Correlation between the pairs of gene

Furthermore the genes will be clustered by performing a hierarchical clustering on top of the correlation matrix. For this the **Number of clusters** parameter is required. Default value = 2. The clustering also includes annotations related to network properties of those genes within the PPI, such as degree, closeness, and eigenvector, as well as information on whether a gene is a transcription factor (TF) or a microRNA.



Figure 31: Gene pairs correlation



Genes belonging to each cluster can be visualized as a table.

For each cluster, an enrichment analysis will be performed by means of the R package gprofiler2. For this the **Correction method** parameter is required. Default value = "fdr". Enrichment results can be visualized in a

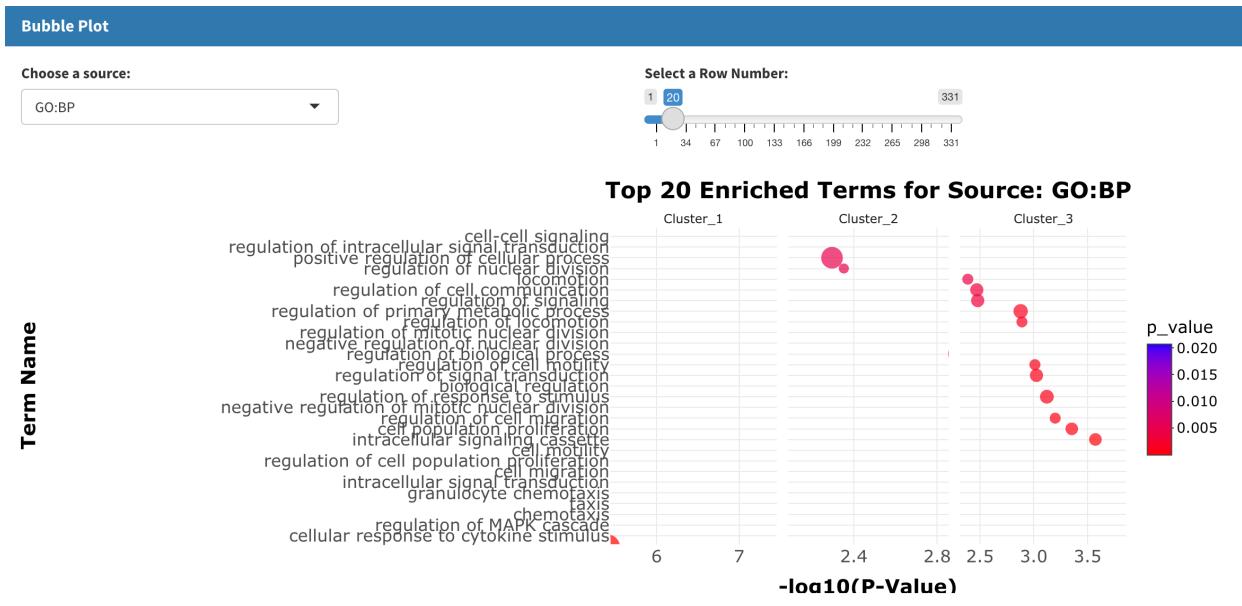
| Genes in Cluster |         |
|------------------|---------|
| Show 10 entries  | Search: |
| cluster          | feature |
| All              | All     |
| Cluster_1        | HES4    |
| Cluster_1        | ALPL    |
| Cluster_1        | KCNN3   |
| Cluster_1        | TNR     |
| Cluster_1        | LAMC2   |
| Cluster_1        | USH2A   |
| Cluster_1        | GABRA4  |
| Cluster_1        | KIT     |
| Cluster_1        | CXCL1   |
| Cluster_1        | AREG    |

Showing 1 to 10 of 151 entries

Previous 1 2 3 4 5 ... 16 Next

Figure 32: Genes in each cluster

gost plot or bubble plot.



Enrichment results can be visualized and exported as a table.

## Gene Sets

For all the statistically significant pathways identified with the FunMappOne tool or KE enrichment for the selected experiment and time point, the gene expression profile of the genes annotated in the pathways can be plotted and compared.

User can also visualize how these genes are associated to each other within the PPI, taking into account their properties (degree and closeness) and characteristics, such as whether they are TF or miRNA.

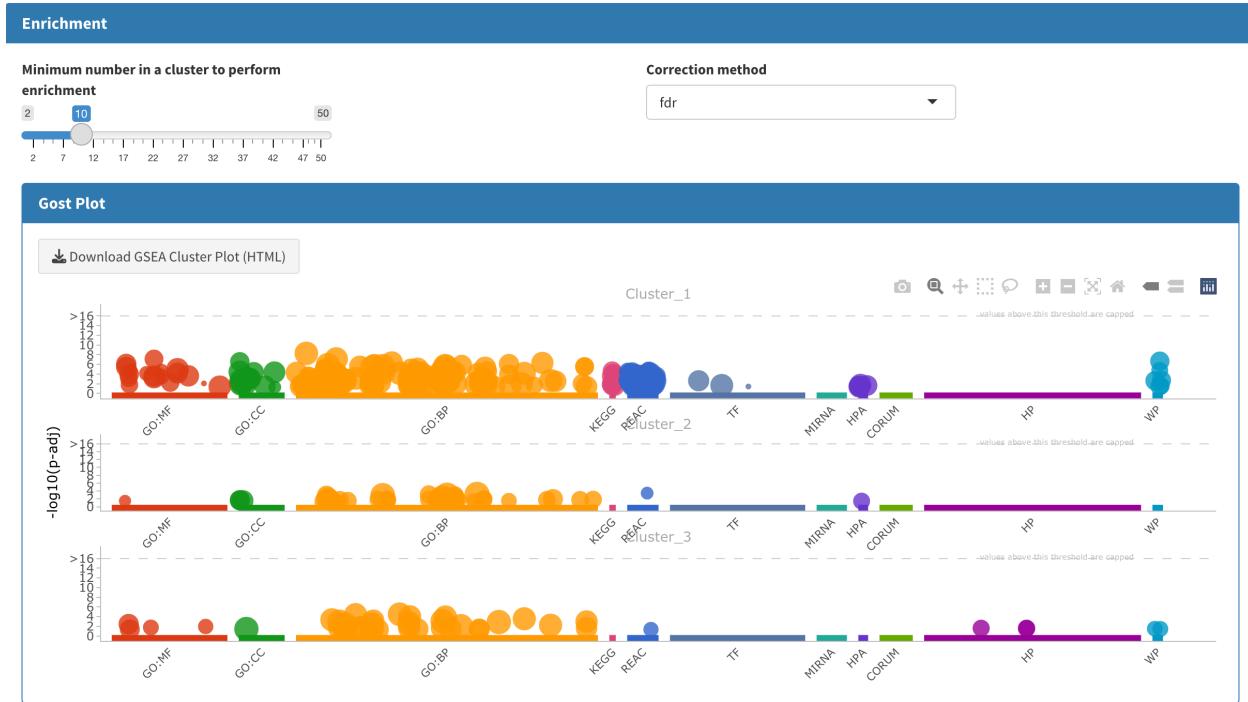


Figure 33: Gost enrichment plot

**Gost Table**

Download GSEA Cluster Table

Show 10 entries

Search:

| query       | significant | p_value               | term_size | query_size | intersection_size | precision          | recall               | term_id    | source |
|-------------|-------------|-----------------------|-----------|------------|-------------------|--------------------|----------------------|------------|--------|
| 1 Cluster_1 | true        | 0.0003790923414426457 | 1530      | 33         | 13                | 0.3939393939393939 | 0.00849673202614379  | GO:0007155 | GC     |
| 2 Cluster_1 | true        | 0.001622135395856767  | 1736      | 33         | 13                | 0.3939393939393939 | 0.007488479262672811 | GO:0042592 | GC     |
| 3 Cluster_1 | true        | 0.002919868130201777  | 238       | 33         | 6                 | 0.1818181818181818 | 0.02521008403361345  | GO:0097529 | GC     |
| 4 Cluster_1 | true        | 0.002919868130201777  | 238       | 33         | 6                 | 0.1818181818181818 | 0.02521008403361345  | GO:0030595 | GC     |
| 5 Cluster_1 | true        | 0.003144954364142211  | 1534      | 33         | 12                | 0.3636363636363636 | 0.007822685788787484 | GO:0016477 | GC     |
| 6 Cluster_1 | true        | 0.006005217725871892  | 825       | 33         | 9                 | 0.2727272727272727 | 0.01090909090909091  | GO:0048871 | GC     |

Figure 34: Gost enrichment table

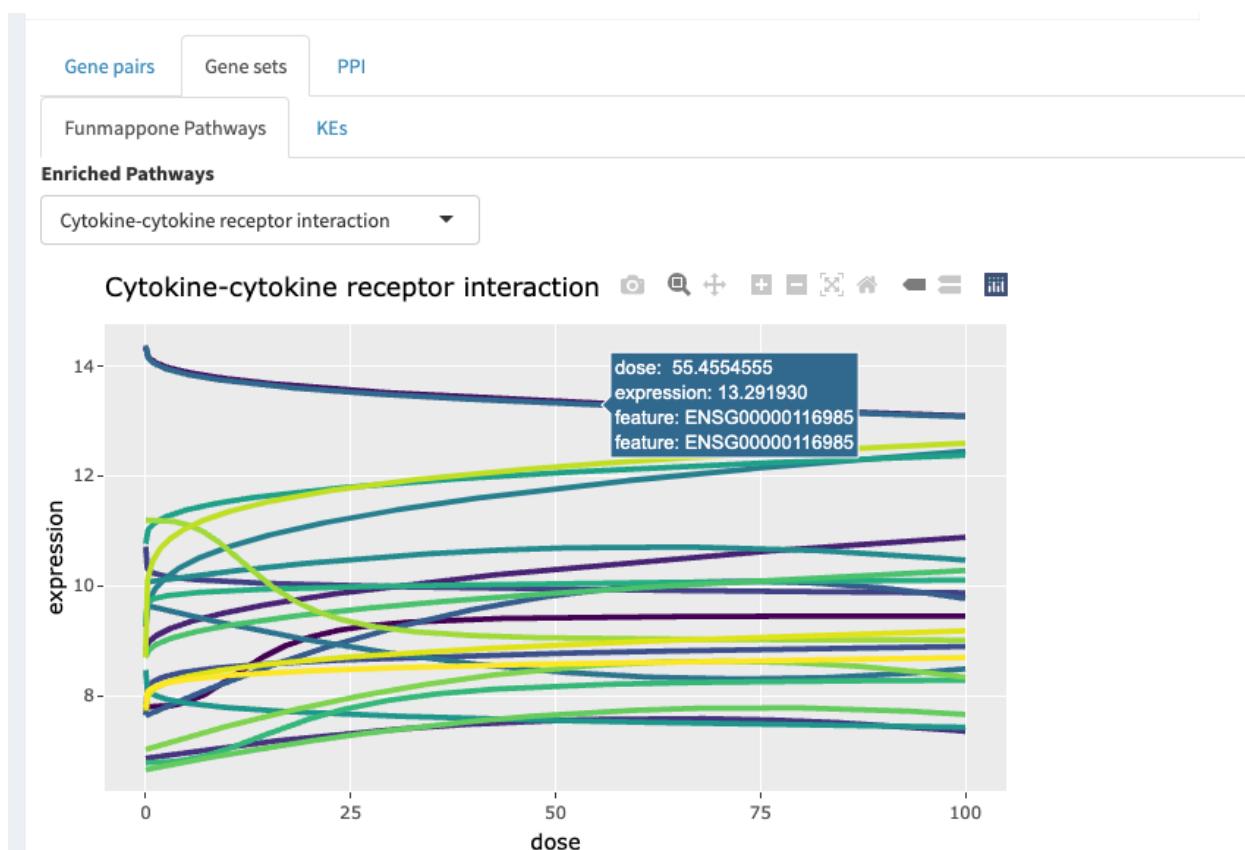


Figure 35: Gene expression profile of genes belonging to the same pathway

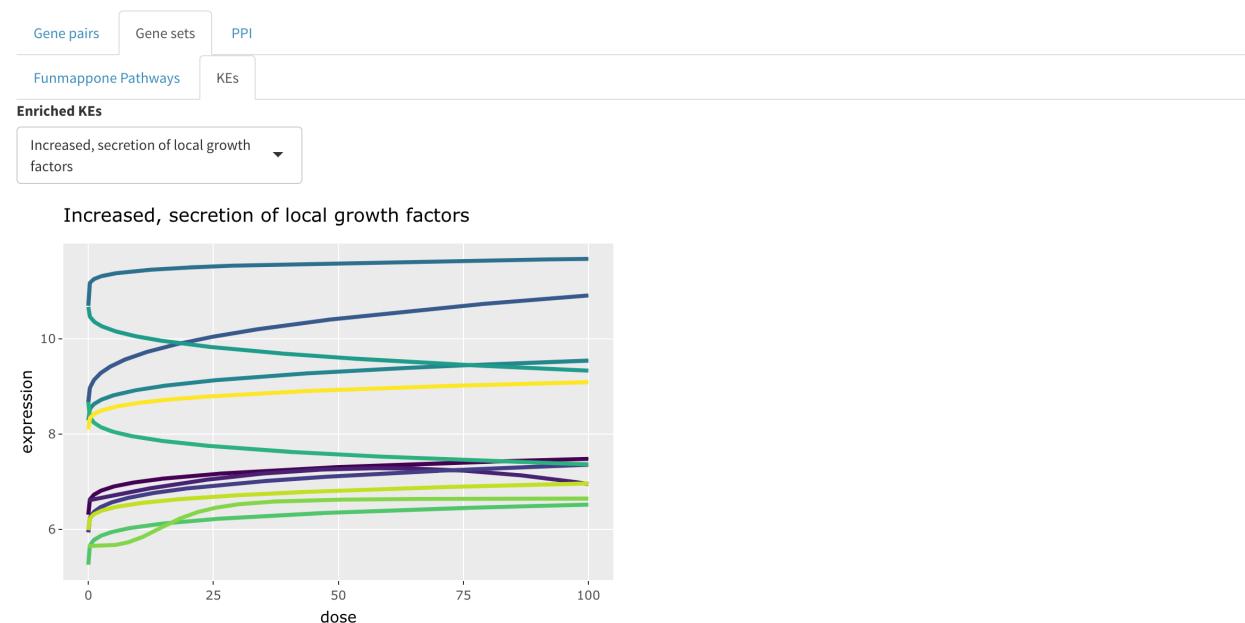


Figure 36: Gene expression profile of genes belonging to the same KE

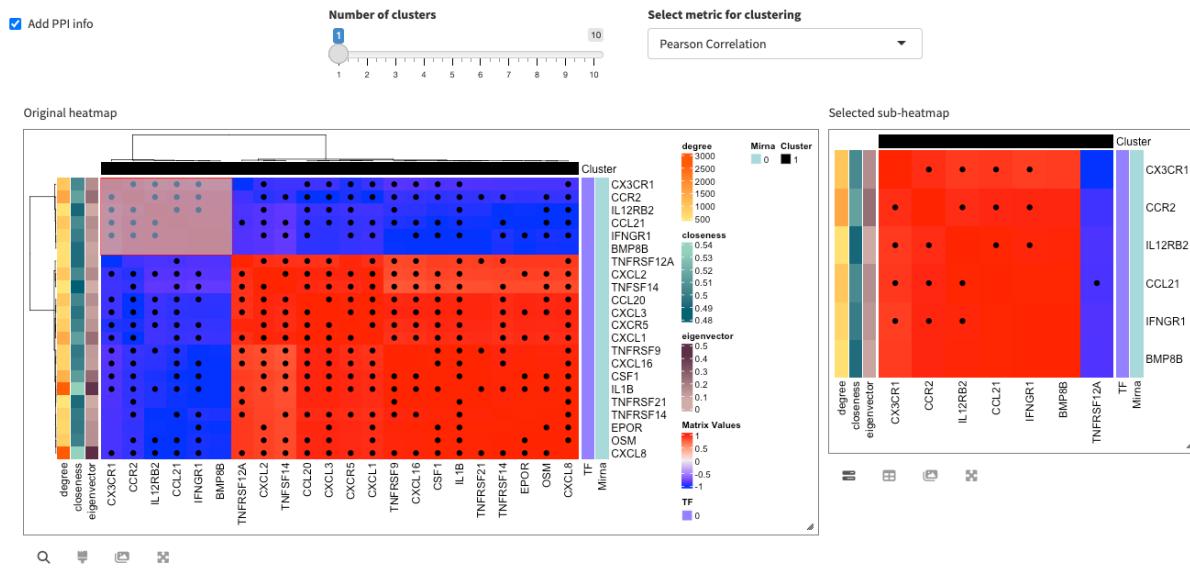


Figure 37: Clustering of genes correlated in the cytokine-cytokine pathway

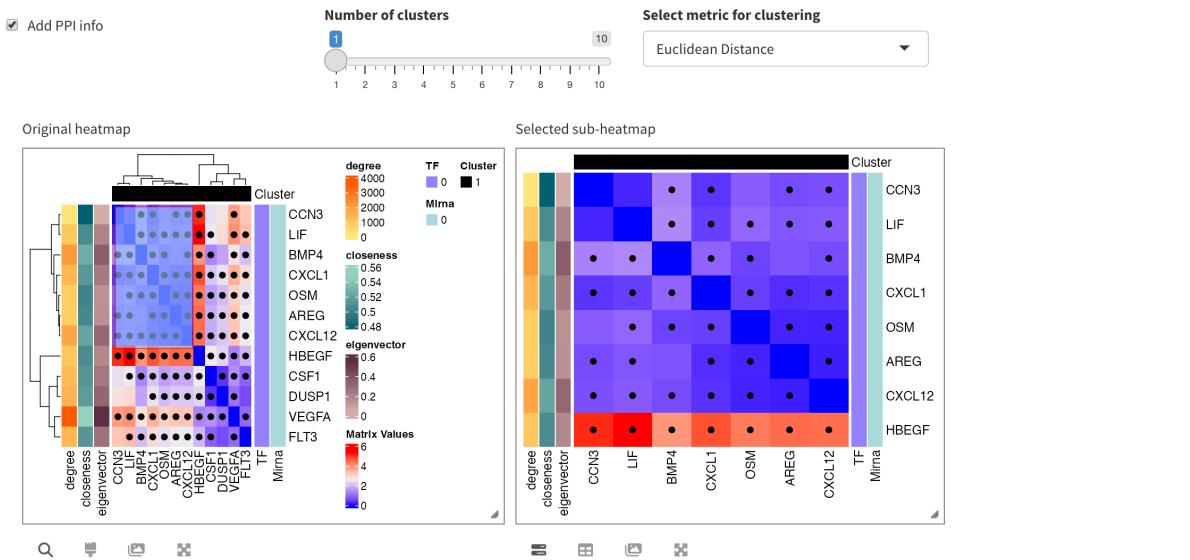


Figure 38: Clustering of genes correlated in the same KE

## PPI analysis

The PPI network of the selected organism (human, mouse, or rat) will be used to analyse the pairs of correlated genes. A subgraph of the PPI network is created by selecting all the edges incident on pairs of correlated genes. The genes in the subgraph are ranked according to their degree (number of edges incident on the node). The gprofiler2 package is used to perform a GSEA analysis over the genes ranked by degree centrality.

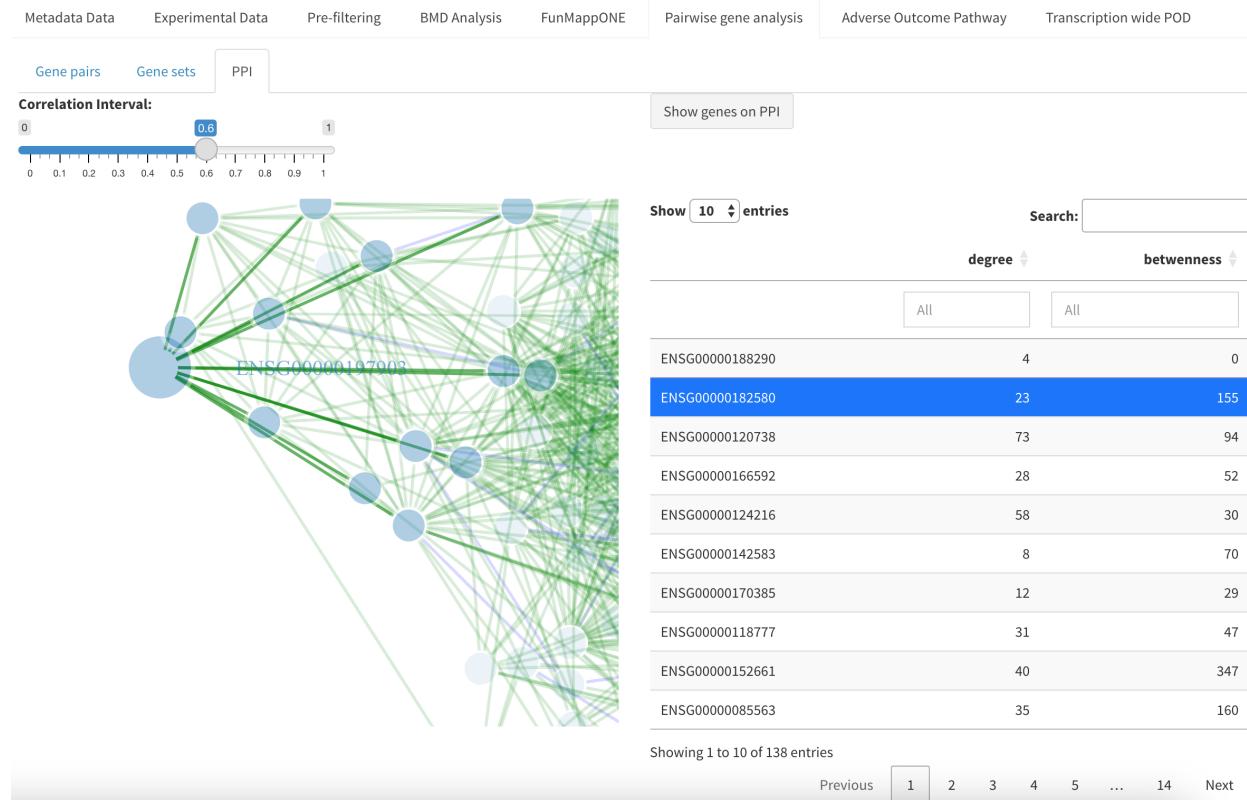


Figure 39: Gene pairs correlation

The parameters used for this analysis are the following:

- **Correlation interval:** Used to filter pairs of genes with correlation under this threshold
- **Correction method:** Method used to correct the p-value of the GSEA analysis

## AOP and KE enrichment

Enrichment analysis is performed using Fisher's exact considering recent annotation of genes to key events (KE) and adverse outcome pathways (AOPs) from AOP-Wiki (<https://aopwiki.org/>), for more details. The following parameters are required to run the enrichment:

- **Select organism:** Choose a species related to humans, such as human, mouse, or rat.
- **Select gene identifier:** Select a gene identifier type, which can be Ensembl, Entrez, or symbol.
- **(KEs) Enrichment Pval:** Specify the p-value threshold to consider a Key Event (KE) as significantly enriched.
- **(AOPs) Enrichment Pval:** Specify the p-value threshold to consider an Adverse Outcome Pathway (AOP) as significantly enriched.
- **Minimum length AOP:** Define the minimum size of an AOP to be considered.

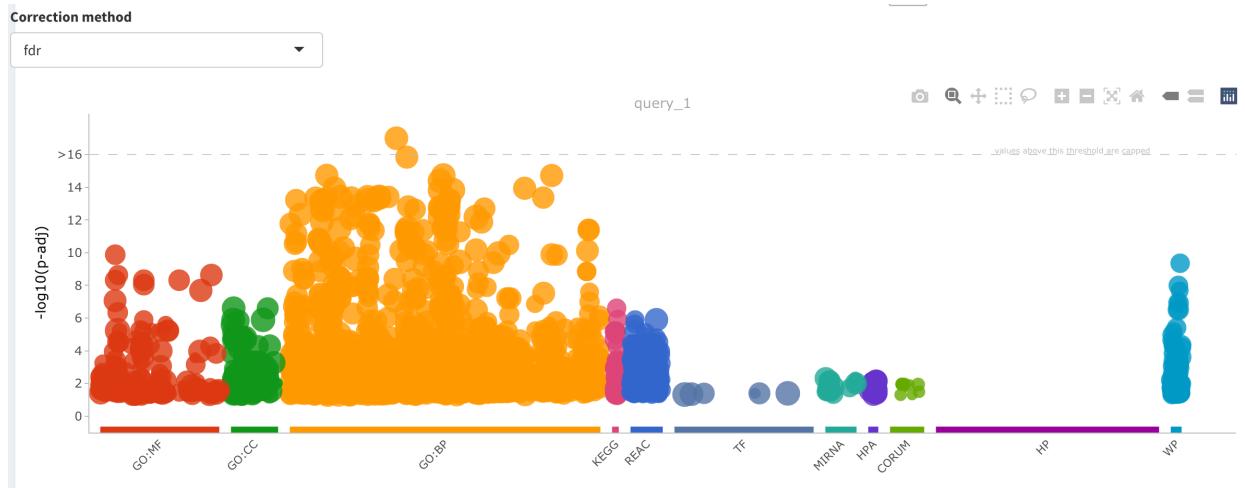


Figure 40: Gene pairs correlation

- **(KEs) Correction method:** Choose a method to adjust p-values for KE enrichment results, such as Bonferroni correction or False Discovery Rate (FDR).
- **(AOPs) Correction method:** Choose a method to adjust p-values for AOP enrichment results, such as Bonferroni correction or False Discovery Rate (FDR).
- **Percentage of enriched KE in AOP:** Specify the minimum proportion of Key Events (KEs) that are enriched within the entire AOP.

User has the option to add a background file or use existing background from BMDx app by selecting “use annotate gene as background”.

There are options to select that allow you to focus on KEs and AOPs that are significant.

### KE enrichment results

In this section, KE enrichment results can be visualized as a table. The results are exportable. In addition, POD distribution values can be visualized by KE type, KE level, and across KE annotation biological systems as box plots. The results are exportable.

### KE-KE network

Within the KEr window, results can be visualized in a KE-KE network. Select the experiment you are interested in and create a KE-KE network. If you want to focus on KEs associated with specific SSbD categories related to safety, you can choose among the SSbD categories. The default visualization “enlarges KE”, meaning it considers both enriched KEs and non-enriched KEs connected to enriched KEs, providing a comprehensive view of the network. This MOA reconstruction is based on a recent approach that prioritizes plausible molecular initiating events (MIEs) and adverse outcomes (AOs) and maps them back to the AOP network. For more details, refer to Del Giudice et al. (2024). This metric considers the proximity to an enriched KE on the KE-KE network and the presence of significantly altered genes in the event’s annotated gene sets. Non-enriched KEs are colored in gray, while enriched KEs are colored in red. Different shapes represent the types of KEs. The user can also filter based on the max path length, number of AOs and MIEs to focus on a specific section of the network. Clicking on events in the network allows you to see more information about the specific event.

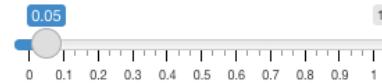
KE-KE networks of different conditions can also be compared by selecting common elements between conditions/experiments or unique elements. For example, “Network A minus Network B” shows what is unique to Network A.

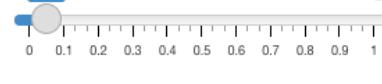
Enrich KEs and AOPs ×

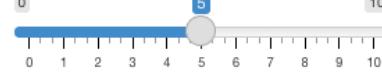
---

**Select organism:**  
Human

**Select gene identifier:**  
Ensembl

**(KEs) Enrichment Pval:**  0.05

**(AOPs) Enrichment Pval:**  0.05

**Minimum length AOP:**  5

**(KEs) Correction method:** fdr  (KEs) Significant Results only

**(AOPs) Correction method:** fdr  (AOPs) Significant Results only

**Select aggregation function:** Median

**Percentage of enriched KE in AOP:**  0.33

Use annotate genes as background

**Choose background file**  
Browse... No file selected

Enrich KE Close

Figure 41: AOP and KE enrichment parameters.

| Metadata Data                                                                                                                                                                                       | Experimental Data           | Pre-filtering      | BMD Analysis    | FunMapONE                                             | Adverse Outcome Pathway | Pairwise gene analysis | Transcription wide POD |      |                        |             |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|--------------------|-----------------|-------------------------------------------------------|-------------------------|------------------------|------------------------|------|------------------------|-------------|-------|
| Key events table                                                                                                                                                                                    | Kr                          | AOP table          | AOP fingerprint |                                                       |                         |                        |                        |      |                        |             |       |
| Enrichment table                                                                                                                                                                                    | POD distribution across KEs |                    |                 |                                                       |                         |                        |                        |      |                        |             |       |
| <b>Experiment</b>                                                                                                                                                                                   |                             | <b>Time Points</b> |                 |                                                       |                         |                        |                        |      |                        |             |       |
| bleomycin                                                                                                                                                                                           |                             | All                |                 | <input checked="" type="checkbox"/> Convert to symbol |                         |                        |                        |      |                        |             |       |
| <a href="#">Download KEs enrichment results</a>                                                                                                                                                     |                             |                    |                 |                                                       |                         |                        |                        |      |                        |             |       |
| Show 5 entries                                                                                                                                                                                      |                             |                    |                 | Search: <input type="text"/>                          |                         |                        |                        |      |                        |             |       |
| Experiment                                                                                                                                                                                          | timepoint                   | TermID             | Ke_description  | BMD                                                   | BMDL                    | BMDU                   | pval                   | padj | relevantGenesInGeneSet | GeneSetSize | Genes |
| All                                                                                                                                                                                                 | All                         | All                | All             | All                                                   | All                     | All                    | All                    | All  | All                    | All         | All   |
| 1 bleomycin 24 Event:1028 Activation of specific nuclear receptors, PPAR-gamma activation 0.054857 0.014127 0.518135 0.002102726 0.01916716 11 410 AREG;HBEGF;H2BC4;H2AC6;H2BC5;H2BC8;H4C8;H4       |                             |                    |                 |                                                       |                         |                        |                        |      |                        |             |       |
| 2 bleomycin 24 Event:1028 Activation of specific nuclear receptors, PPAR-gamma activation 0.054857 0.014127 0.518135 0.002102726 0.01916716 11 410 AREG;HBEGF;H2BC4;H2AC6;H2BC5;H2BC8;H4C8;H4       |                             |                    |                 |                                                       |                         |                        |                        |      |                        |             |       |
| 3 bleomycin 48 Event:1028 Activation of specific nuclear receptors, PPAR-gamma activation 0.0612205 0.015028 0.765075 0.000009575544 0.0002377233 25 410 GNG12;H2BC21;DHR59;AREG;HBEGF;H2BC4;H2AC6; |                             |                    |                 |                                                       |                         |                        |                        |      |                        |             |       |

Figure 42: KE enrichment table results.

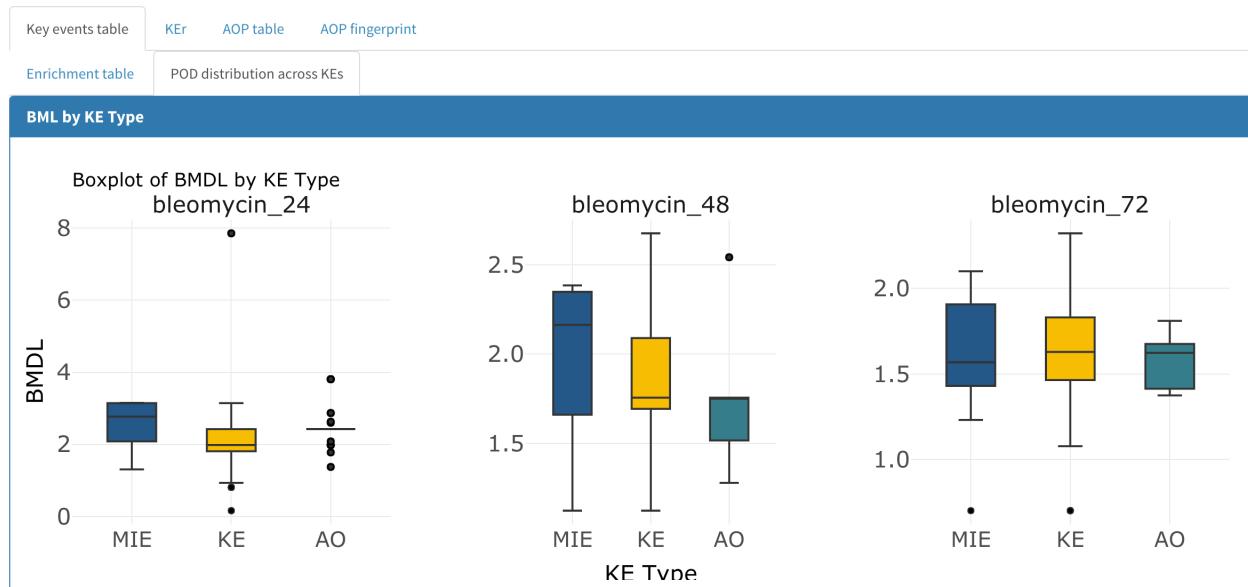


Figure 43: BMDL by KE type.

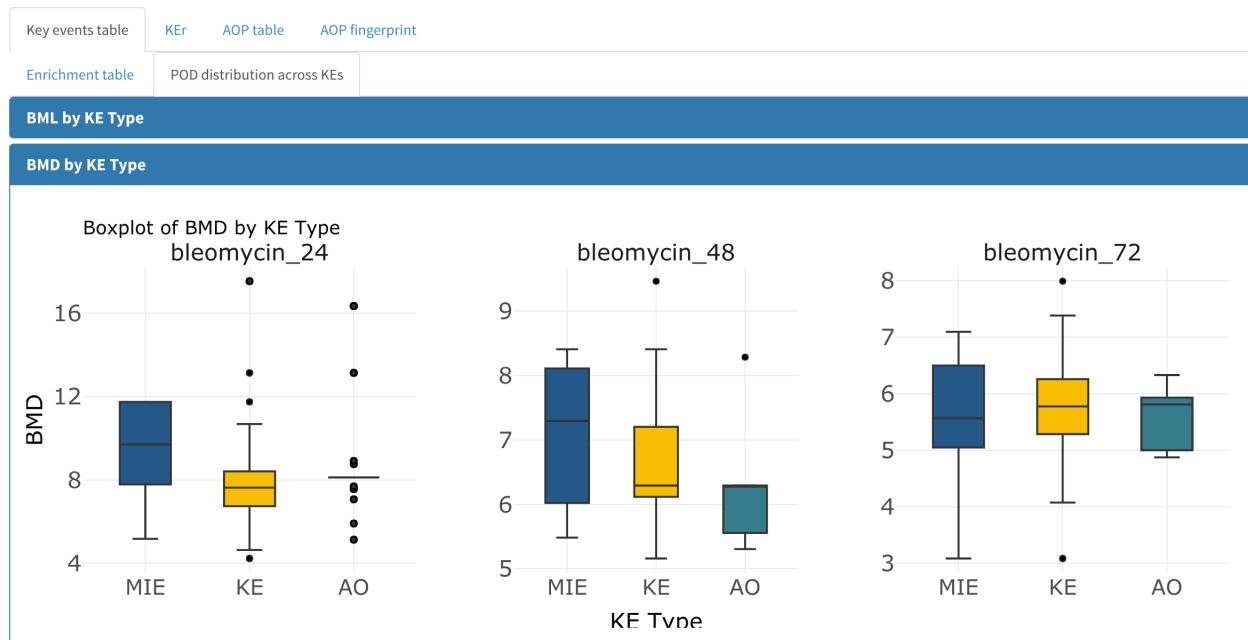


Figure 44: BMD by KE type.

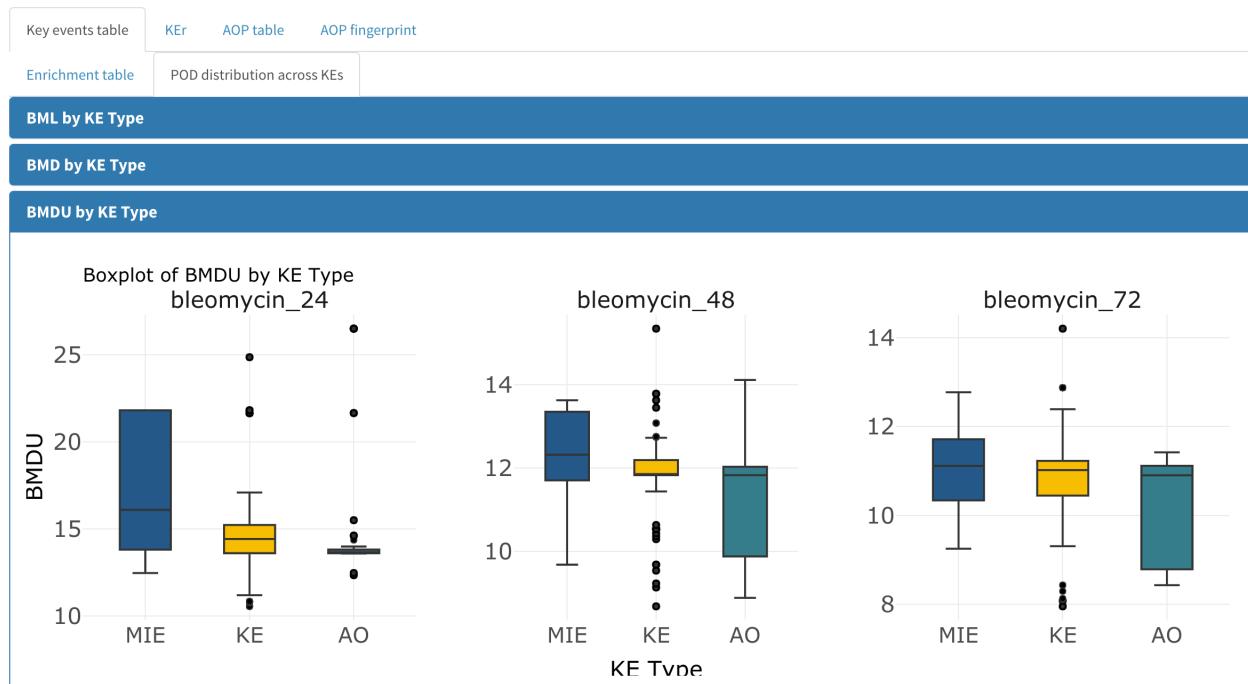


Figure 45: BMDU by KE type.

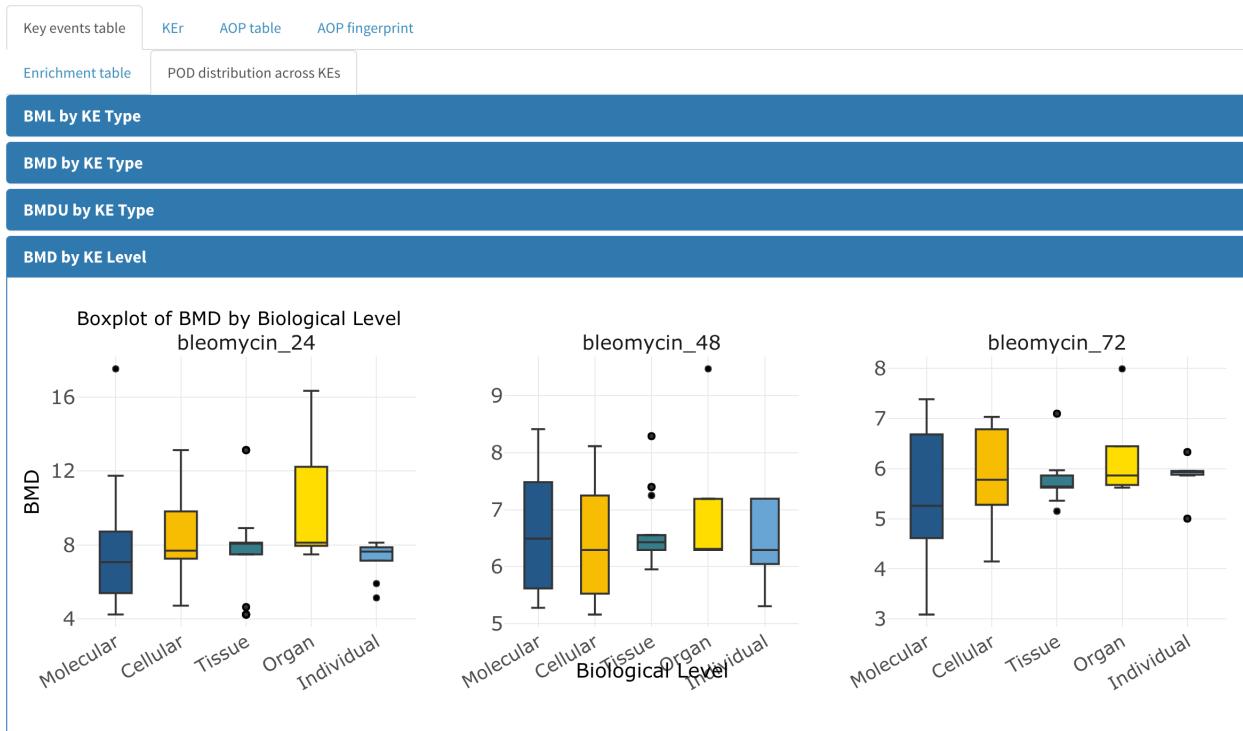


Figure 46: BMD by KE level.

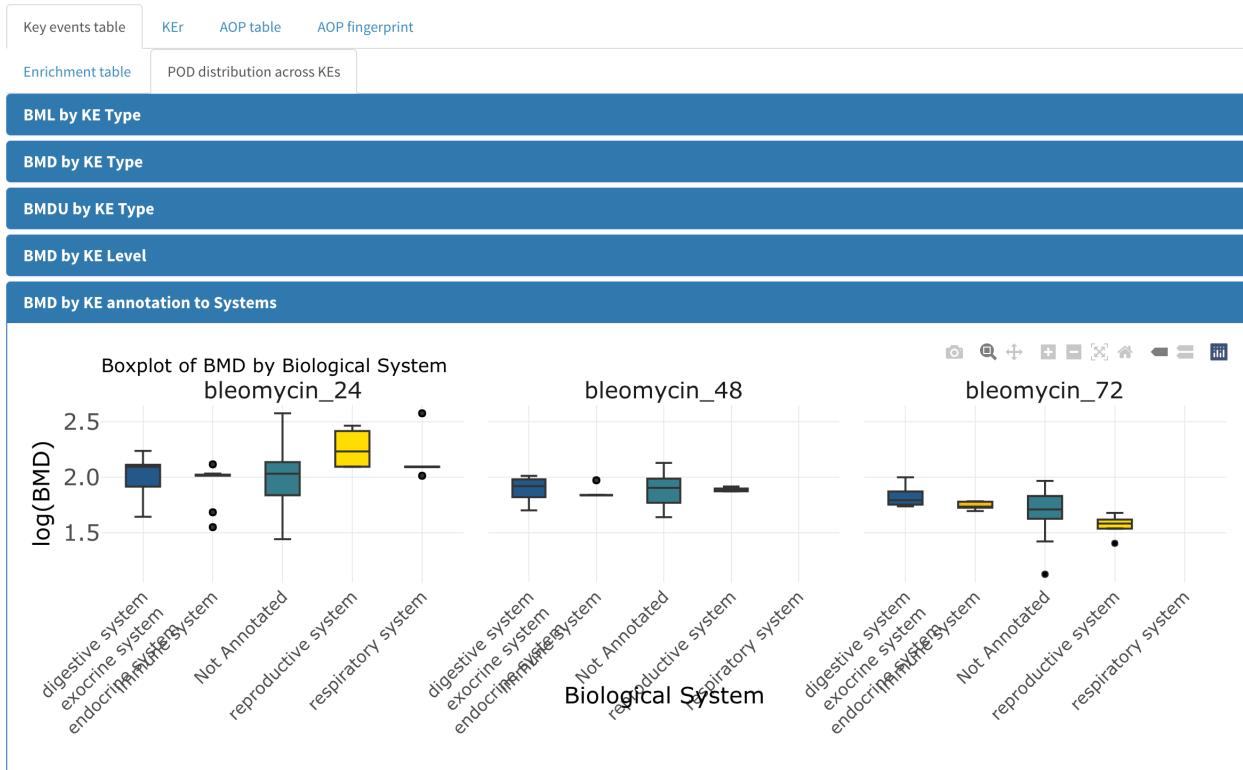


Figure 47: BMD by KE annotation to biological systems.

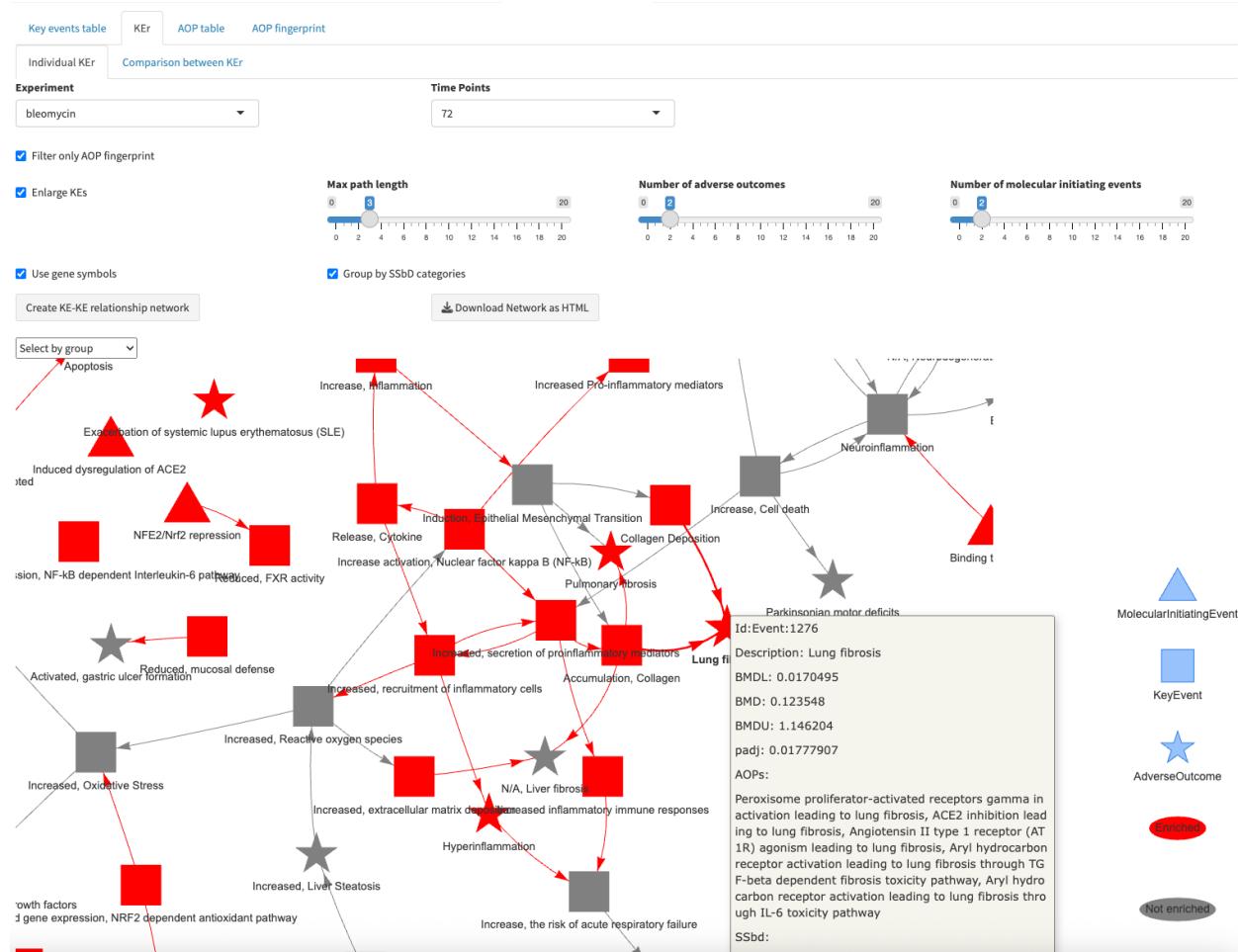


Figure 48: MOA reconstruction of lung fibrosis.

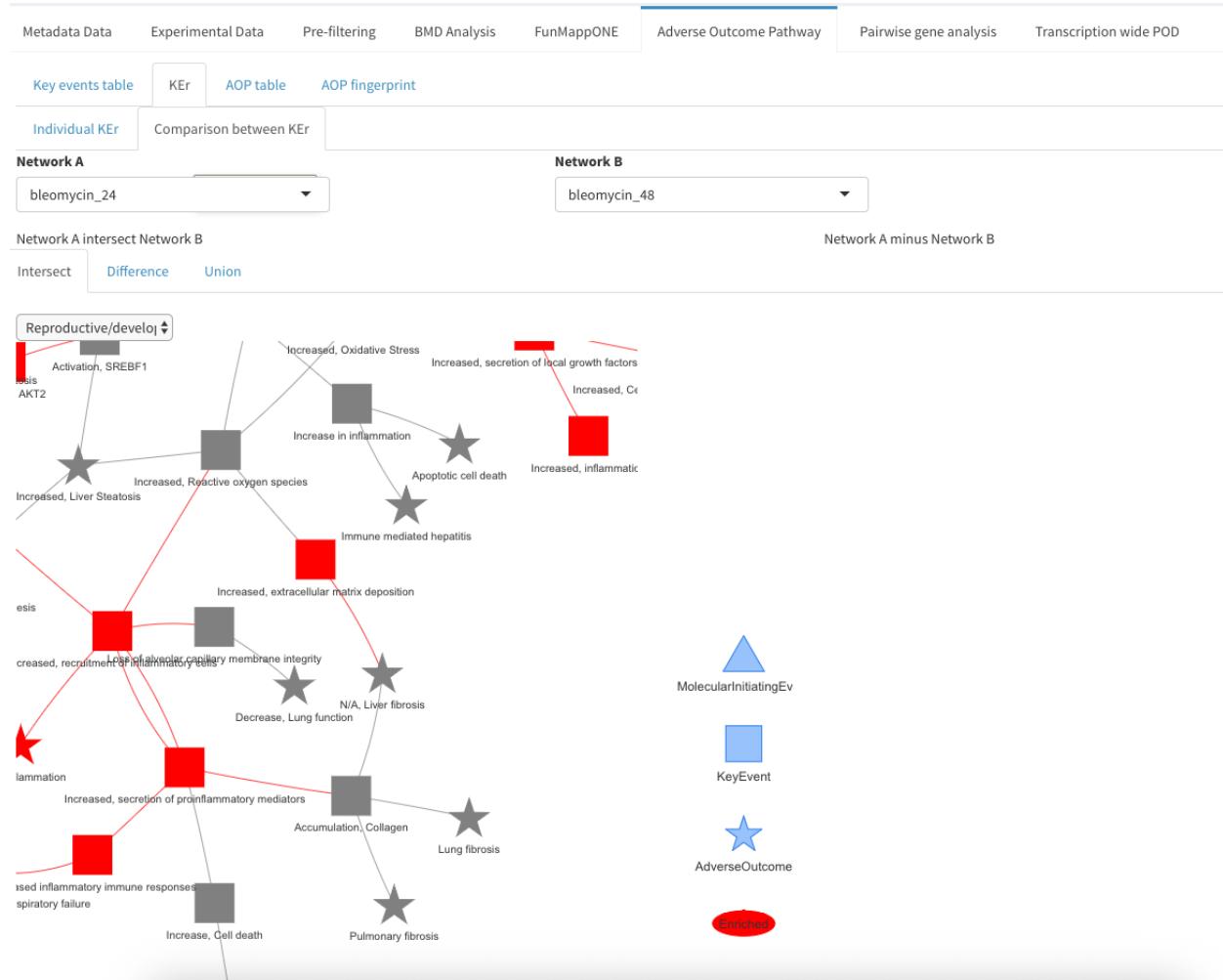


Figure 49: Comparison of KEKE network across conditions.

## AOP Enrichment results

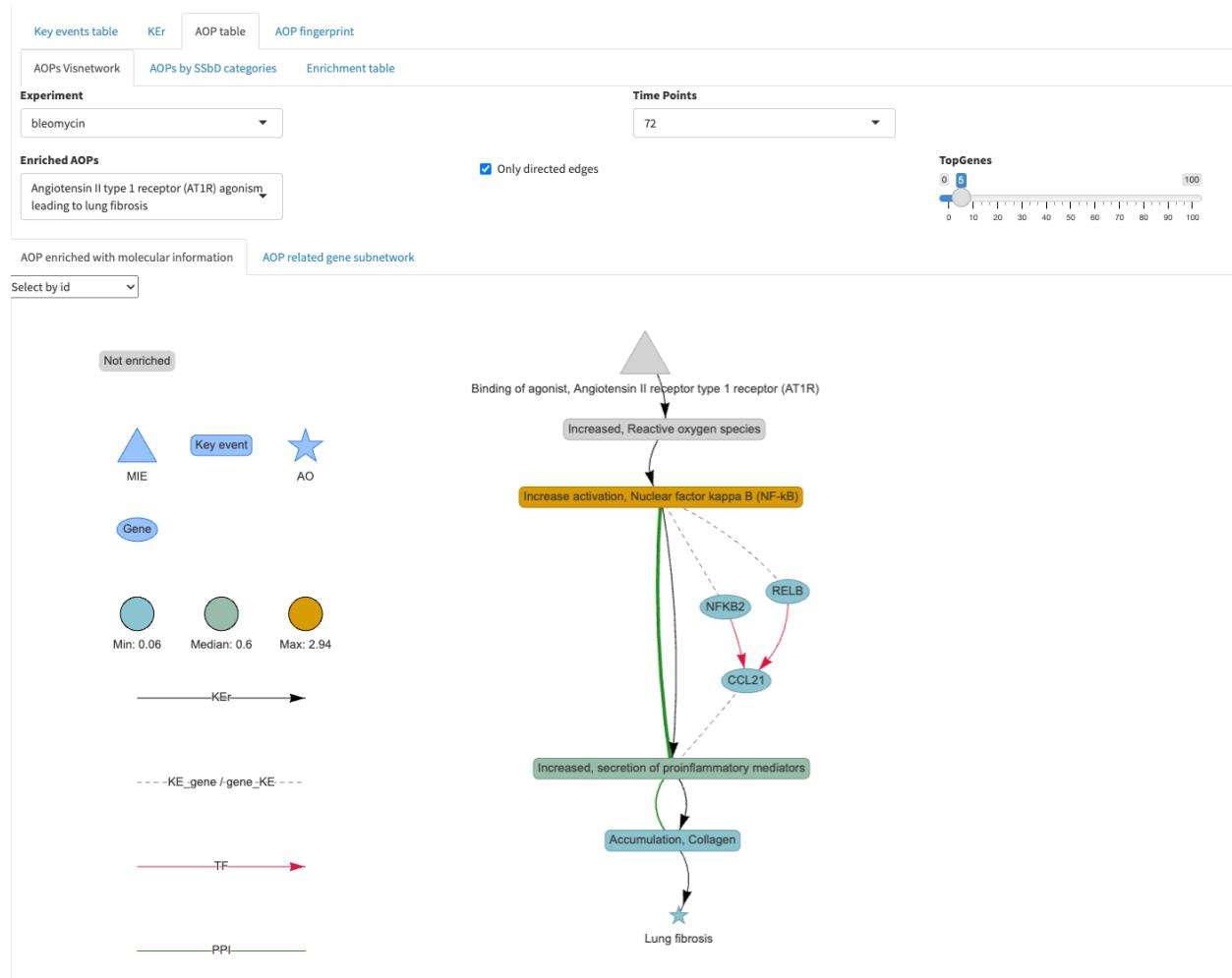
The AOP enrichment table presents all enrichment results. The results are exportable.

| TermID | Experiment | timepoint | relevantGenesInGeneSet | GeneSetSize | pval | Genes                                                                                             |
|--------|------------|-----------|------------------------|-------------|------|---------------------------------------------------------------------------------------------------|
| 1      | Aop:1      | bleomycin | 24                     | 35          | 1303 | 2.120174e-8 ENSG00000117399;ENSG00000123080;ENSG00000184371;ENSG00000058085;ENSG00000162772;E1    |
| 2      | Aop:1      | bleomycin | 24                     | 35          | 1303 | 2.120174e-8 ENSG00000117399;ENSG00000123080;ENSG00000184371;ENSG00000058085;ENSG00000162772;E1    |
| 3      | Aop:1      | bleomycin | 24                     | 35          | 1303 | 2.120174e-8 ENSG00000117399;ENSG00000123080;ENSG00000184371;ENSG00000058085;ENSG00000162772;E1    |
| 4      | Aop:1      | bleomycin | 24                     | 35          | 1303 | 2.120174e-8 ENSG00000117399;ENSG00000123080;ENSG00000184371;ENSG00000058085;ENSG00000162772;E1    |
| 5      | Aop:1      | bleomycin | 48                     | 63          | 1303 | 1.160465e-8 ENSG00000049249;ENSG00000117399;ENSG00000123080;ENSG00000081985;ENSG00000184371;E1    |
| 6      | Aop:1      | bleomycin | 72                     | 53          | 1303 | 0.000005947637 ENSG00000049249;ENSG00000162493;ENSG00000060656;ENSG00000117399;ENSG00000123080;E1 |

Figure 50: AOP enrichment table results.

Moreover, the AOP section includes a mechanistic visualization that contextualizes KEs within an AOP and it integrates with data from PPI to illustrate how they are interconnected through molecular interactions.

## Individual AOP visualization



The top relevant genes have been ranked according to multiple network properties that reflect their connectivity within the PPI network. To visualize the top most relevant genes based on a user defined number, the user should double click on the enriched KEs in the figure at the bottom of the page. These genes will then appear. KEs that share those genes will be considered together and separated by semicolon. Each time a different id of interest is selected, it is important to reinitialize the clustering to ensure accurate and updated results.

In addition, the tool allows to visualize the distribution of BMD values across AOPs grouped by SSbD categories and reports the number of AOPs per category.

### AOP fingerprint

Once the enrichment analysis is completed, users can visualize the enriched AOPs using a bubble plot or range plot. They have the option to display either AOP names or AOP IDs on the y-axis. The font size, as well as the plot's height and width, can be adjusted for optimal export as a PDF. Additionally, AOP enrichment results can be exported. Users can also choose to focus on a specific range of AOPs that fall within particular SSbD categories related to safety, for example, "Carcinogenicity".

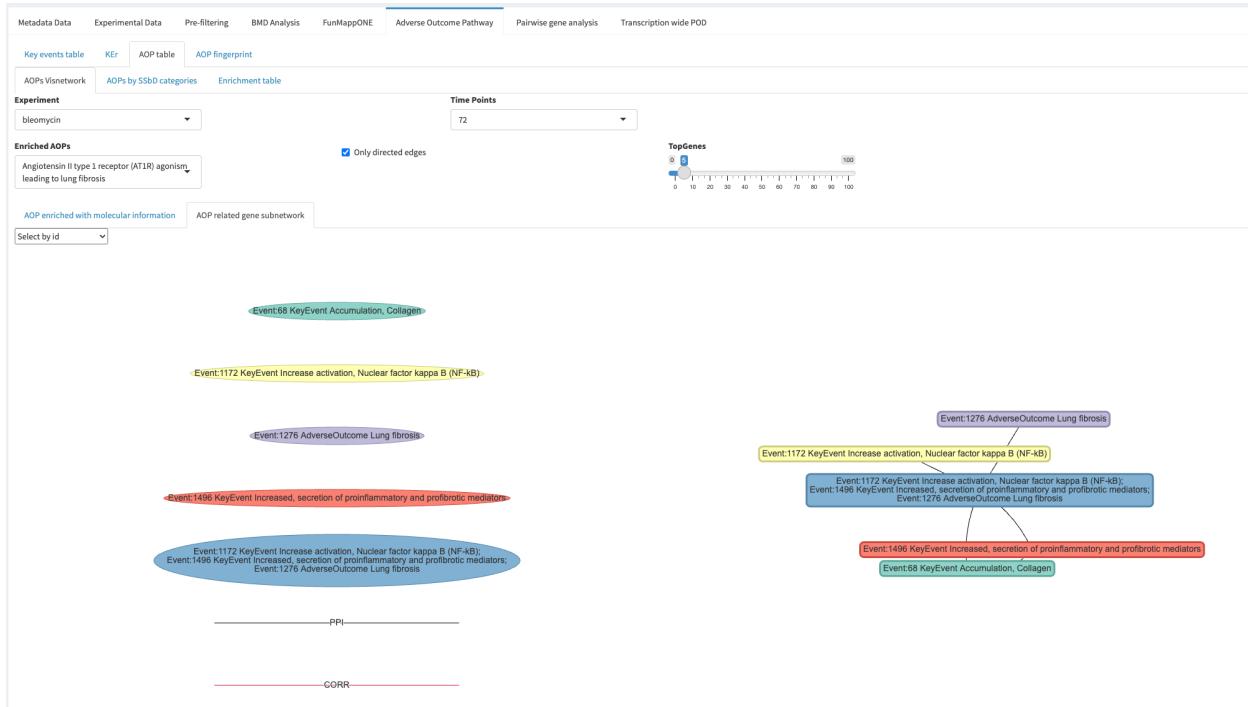


Figure 51: Enriched KEs of the AOP Angiotensis II leading to fibrosis.

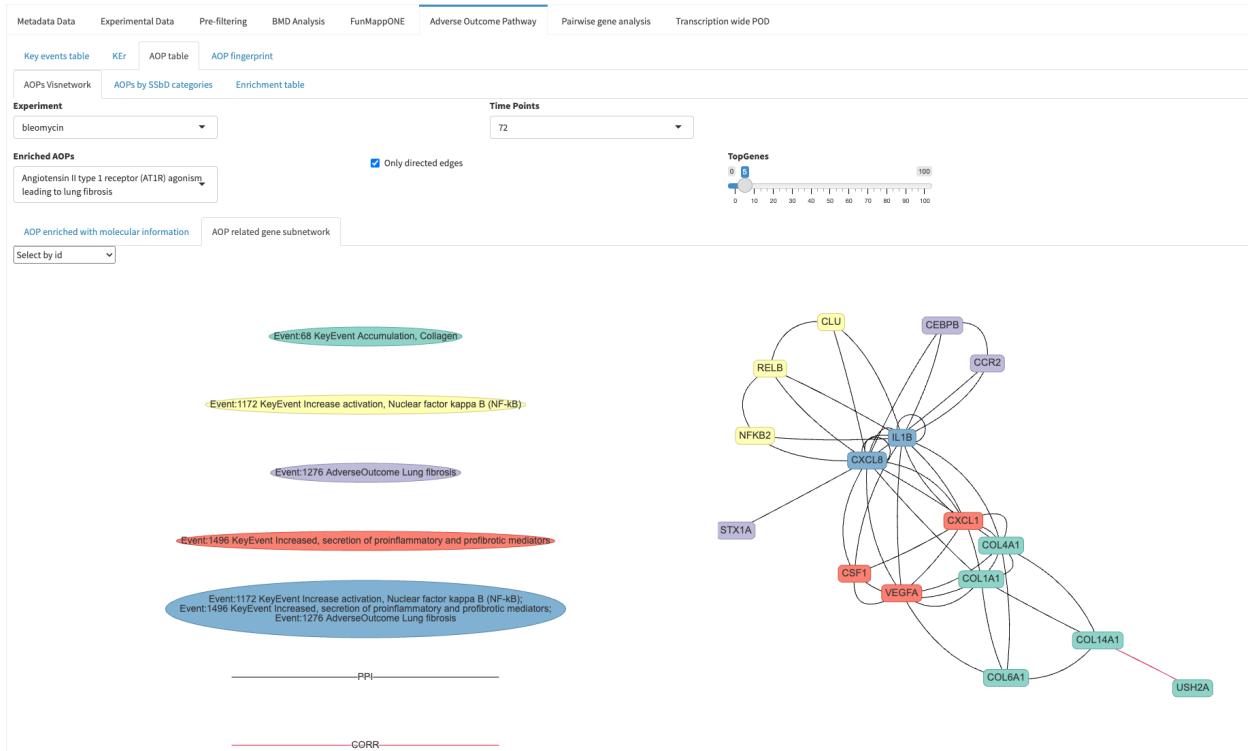


Figure 52: Top genes within the enriched KEs of the AOP Angiotensis II leading to fibrosis.

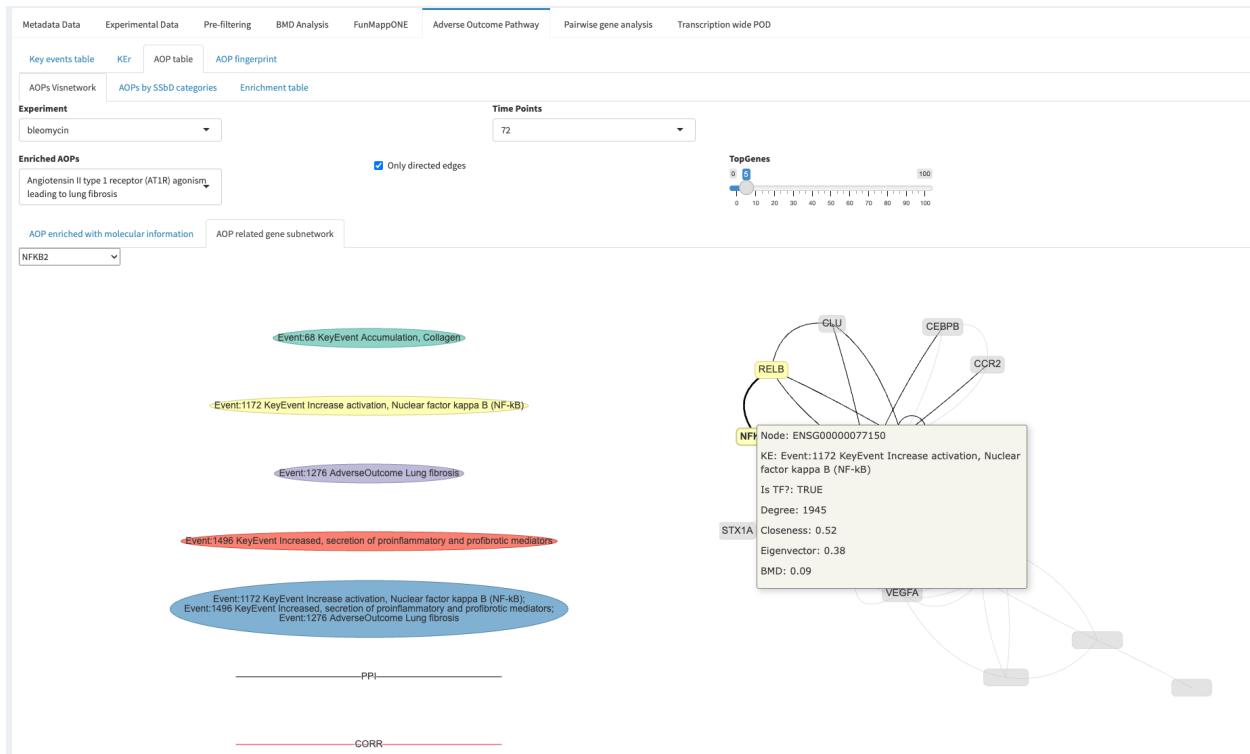


Figure 53: Leading genes found in the enriched KEs of the Angiotensin II AOP linked to fibrosis, with info boxes detailing each gene characteristics.

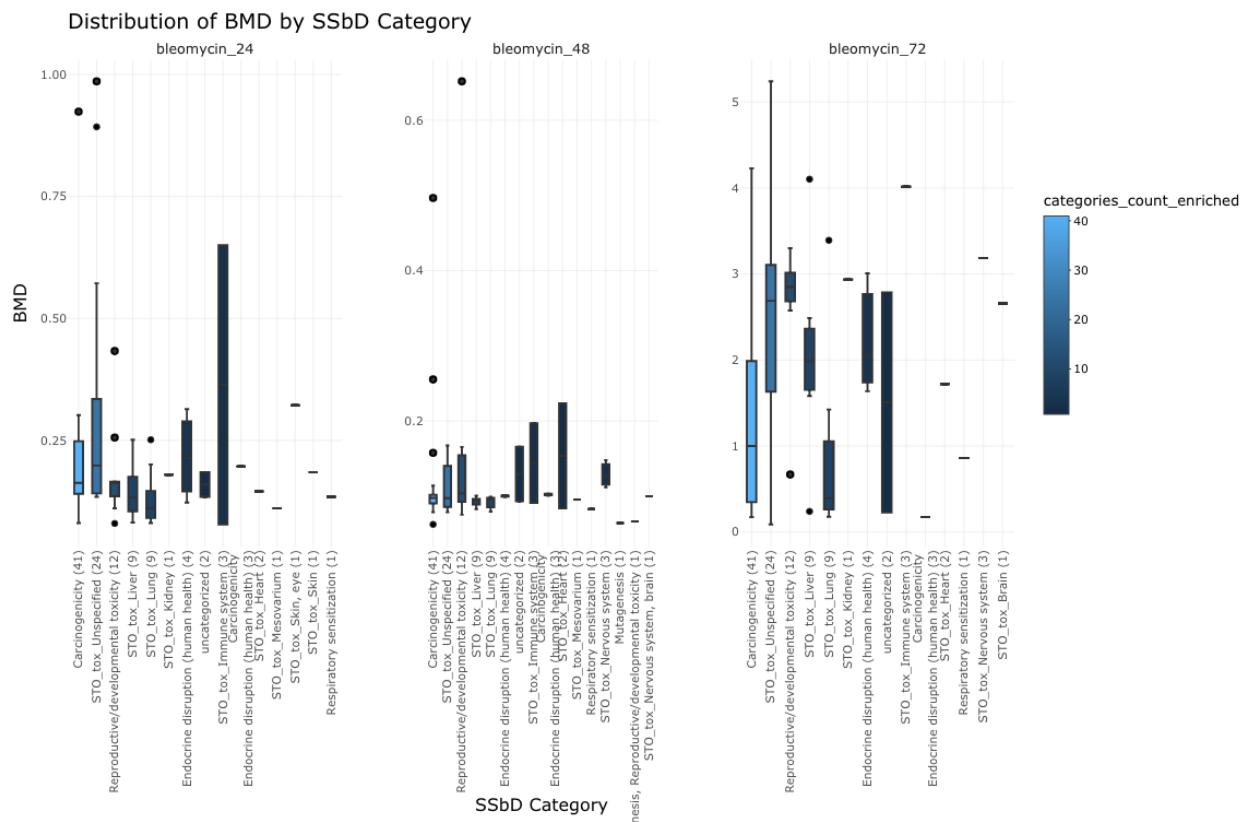


Figure 54: Distribution of BMD values grouped by SSbD categories

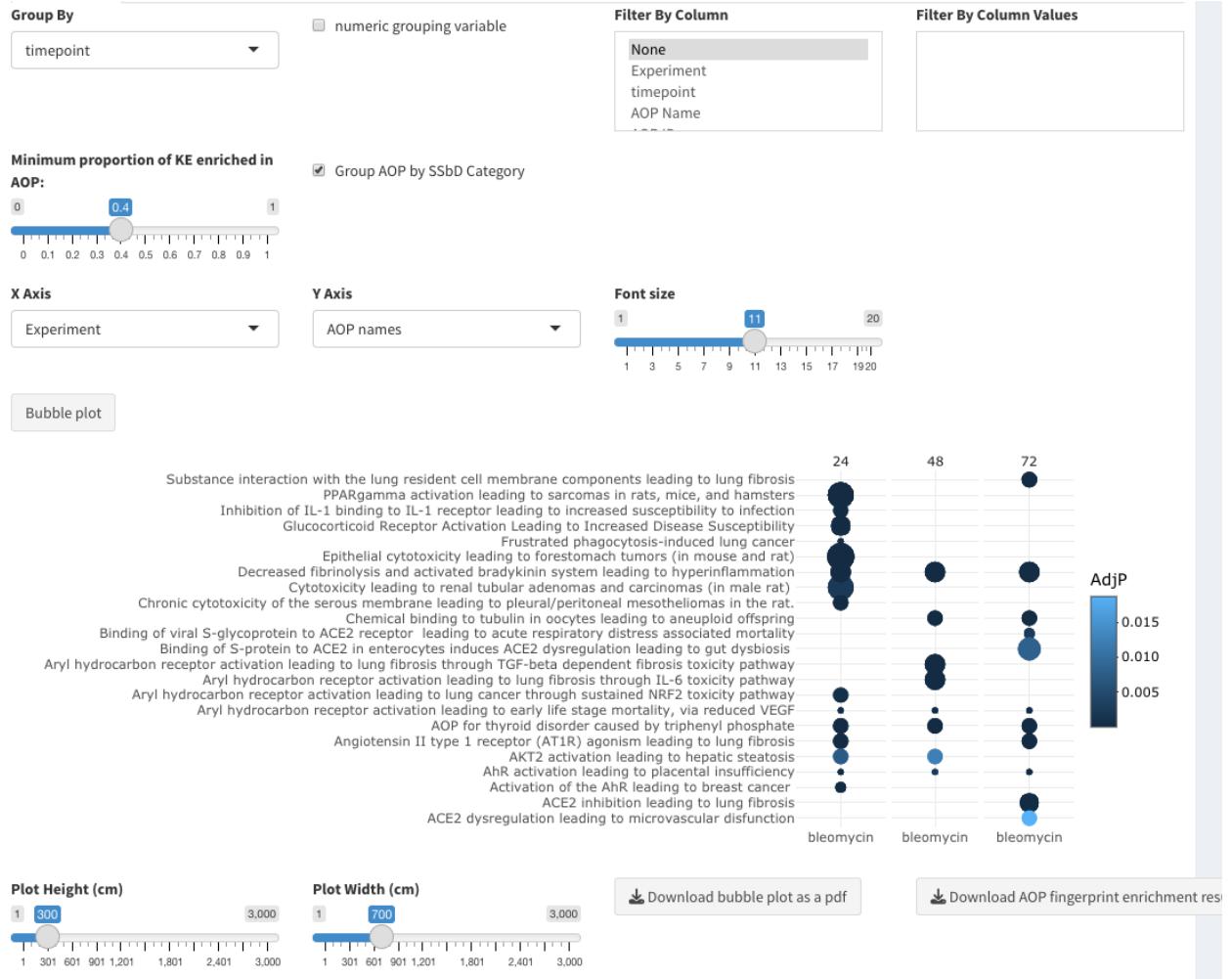


Figure 55: AOP fingerprint bubble plot.

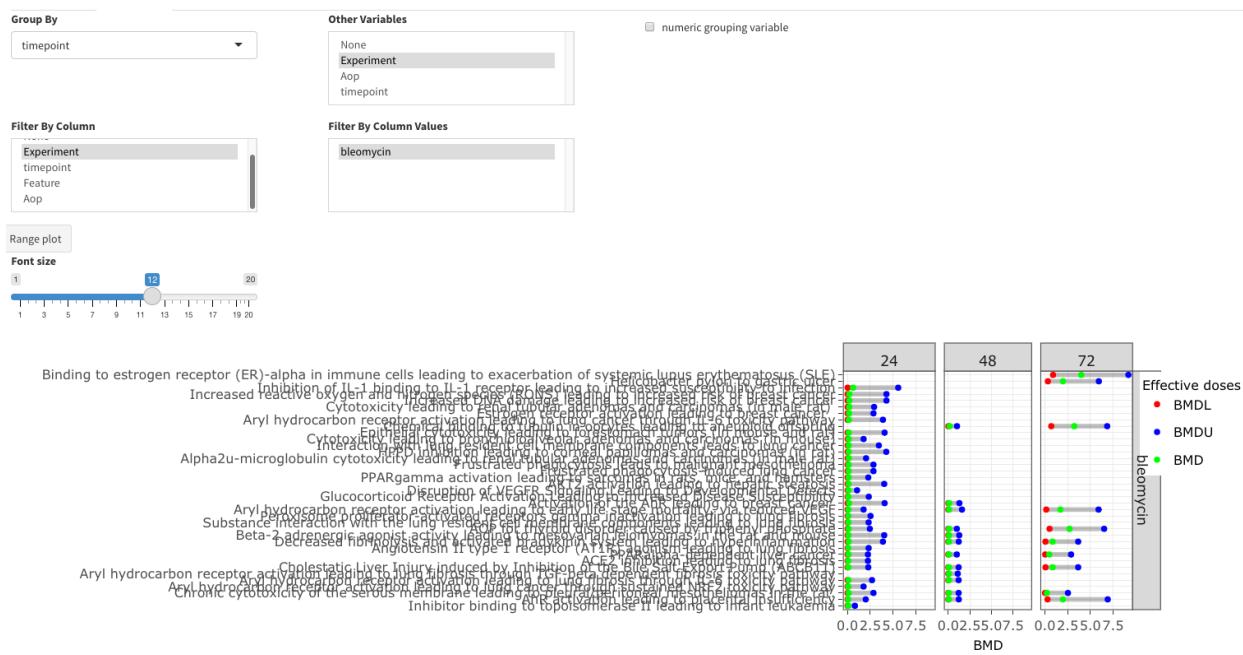


Figure 56: AOP fingerprint range plot.