Data in Brief 38 (2021) 107397

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Transcriptomic dataset from *Arabidopsis thaliana* seedlings in response to *Pseudomonas aeruginosa* mono-rhamnolipids

Noadya Monnier[a,b,c], Catherine Sarazin[b], Sonia Rippa[a,*]

[a] *Unité de Génie Enzymatique et Cellulaire, CNRS UMR 7025, Sorbonne Universités, Université de Technologie de Compiègne (UTC), Compiègne, France*
[b] *Unité de Génie Enzymatique et Cellulaire, CNRS UMR 7025, Université de Picardie Jules Verne (UPJV), Amiens, France*
[c] *Éléphant Vert France SAS, 5 cours du Danube, Serris 77700, France*

ARTICLE INFO

Article history:
Received 12 April 2021
Revised 20 August 2021
Accepted 16 September 2021

Keywords:
Rhamnolipids
Arabidopsis
Microarray
Elicitor
Biocontrol
Plant defense
Plant protection

ABSTRACT

The present data profile the large scale transcriptome changes in *Arabidopsis thaliana* Col-0 seedlings exposed to mono-rhamnolipids (Mono-RLs) from *Pseudomonas aeruginosa* secretome. Bacterial rhamnolipids (RLs) are biosurfactant known to trigger plant defense mechanisms and have a great potential for crop culture protection as environmental-friendly biocontrol solution. They are thought to interact directly with membrane lipids to induce plant defense gene expression and protection towards phytopathogens. However, to date, data on the global transcriptomic modifications induced by these natural amphiphilic glycolipids in plants are missing. Ten-day-old seedlings were treated for 1 or 3 h with 100 µM Mono-RLs in liquid growth medium for root absorption. Total RNA samples were extracted, purified, labelled and hybridized to Agilent V4 Gene Expression Microarrays 4 × 44 K (design ID 021169) carrying 43803 ssDNA probes of 60-mer covering the entire genome of *A. thaliana*. The dataset was validated by quality assessments including RNA sample quality, microarray quality and global gene expression profiling. The raw and normalized formats of these transcriptomic data are available via GEO repository with the

* Corresponding author.
E-mail address: sonia.rippa@utc.fr (S. Rippa).
Social media: (S. Rippa)

https://doi.org/10.1016/j.dib.2021.107397

accession number GSE168830. The dataset can be used to provide insights into the plant's early and later mechanisms induced or repressed by RLs. It can be compared to data obtained with other plant defense elicitors, including the well described compounds perceived by membrane protein receptors.

Specifications Table

Subject	Biology
Specific subject area	Transcriptomics of plants elicited by rhamnolipids
Type of data	Raw data in .txt files Data table in Excel files (QC reports, expression value matrix, LIMMA output, Fold change data) Figures
How data were acquired	Agilent Arabidopsis V4 Gene Expression Microarrays
Data format	Raw and normalized data
Parameters for data collection	5 samples of ten 10-day-old Col-0 Arabidopsis seedlings were transferred in liquid growth medium containing 0.5% EtOH (as control) or 100 µM Mono-RLs diluted in 0.5% EtOH (as Mono-RLs treated). They were harvested at 0, 1 and 3 h for total RNA extraction. cRNA were labelled and randomly hybridized on microarrays.
Description of data collection	Background correction, quantile normalization and filtering were realized as described in the materials and methods section.
Data source location	Université de Technologie de Compiègne, Compiègne, France
Data accessibility	Normalized and raw data are available in Gene Expression Omnibus (GEO) database of NCBI with accession number GSE168830 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE168830 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171040 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171041 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171042 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171043 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171044 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171045 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171046 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171047 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171048 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171049 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171050 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171051 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171052 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171053 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171054 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171055 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171056 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171057 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171058 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171059 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171060 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171061 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171062 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171063 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5171064

Value of the Data

- The data provide Arabidopsis seedling transcriptomic changes that occur after short or longer exposure to mono-RLs.
- The data can be used by the scientific community studying RL effects on plants and more largely plant defense triggering compounds. They can also be used by the phytosanitary industry developing biocontrol products based on RLs.
- The data can be used as a resource to identify early or later genes differentially expressed after a rhamnolipid treatment and to identify marker genes to follow the RL defense triggering activity in plants.
- The data can be used to decipher the way of perception of amphiphilic compounds with plant defense triggering activity as well as affinity for plasma membrane lipids and to compare them with other plant defense elicitors.

1. Data Description

This dataset represents a large scale transcriptome comparison of Arabidopsis seedlings not exposed or exposed to Mono-RLs for 1 h or 3 h. Total RNA samples were extracted from the seedling tissues and were subjected to the microarray experiments using Agilent Arabidopsis V4 Gene Expression Microarrays (Agilent Technologies, Santa Clara, CA, USA). The flowchart of the experimental design of the study is presented (Fig. 1). The raw and normalized formats of these data are available in National Center for Biotechnology information (NCBI) Gene Expression Omnibus (GEO) gene expression data repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE168830). The data obtained after each processing step are provided as supplementary data files. Quality Control (QC) reports generated after array scanning and data extraction are provided in Supplemental Appendix S1. Final linear fit output values generated on the R software are gathered in Table 1. Expression value matrix after background correction, normalization and without probe replicates are supplied in Table 2. Unfiltered fold change data are gathered in Table 3. Reproducibility of the transcriptomic data from 5 independent replicates of each sample was confirmed by a principal component analysis (PCA) (Fig. 2). The magnitude of the transcriptomic modulations was globally assessed with the dispersion of Fold Changes (FC) (Fig. 3) and Venn diagrams of significantly regulated transcripts (Fig. 4).

2. Experimental Design, Materials and Methods

2.1. Mono-RLs solution preparation

Mono-RLs (mainly composed of α-L-rhamnopyranosyl-β-hydroxydecanoyl-β-hydroxydecanoate) were isolated from a commercial mixture (Jeneil Biotech, Inc, Saukville, USA). They were kindly provided by Dr Stéphan Dorey (RIBP, URCA, Reims, France). They were purified by the ICMR (URCA, Reims, France) by fast centrifugal partition chromatography (FCPC, Kromaton Technologies apparatus (Angers, France)) as previously described in [1]. Mono-RLs initial stock was prepared in absolute ethanol at 20 mM concentration. To ensure complete dilution, this stock was diluted in water at 2 mM concentration (10% ethanol) [2].

2.2. Plant growth and treatment

A. thaliana genotype Col-0 corresponding to Nottingham Arabidopsis Stock Center (NASC) reference N1092 was used. The whole protocol was realized in sterile conditions and independently repeated 5 times, to collect 5 independent biological replicates. Seeds were sterilized for 5 min

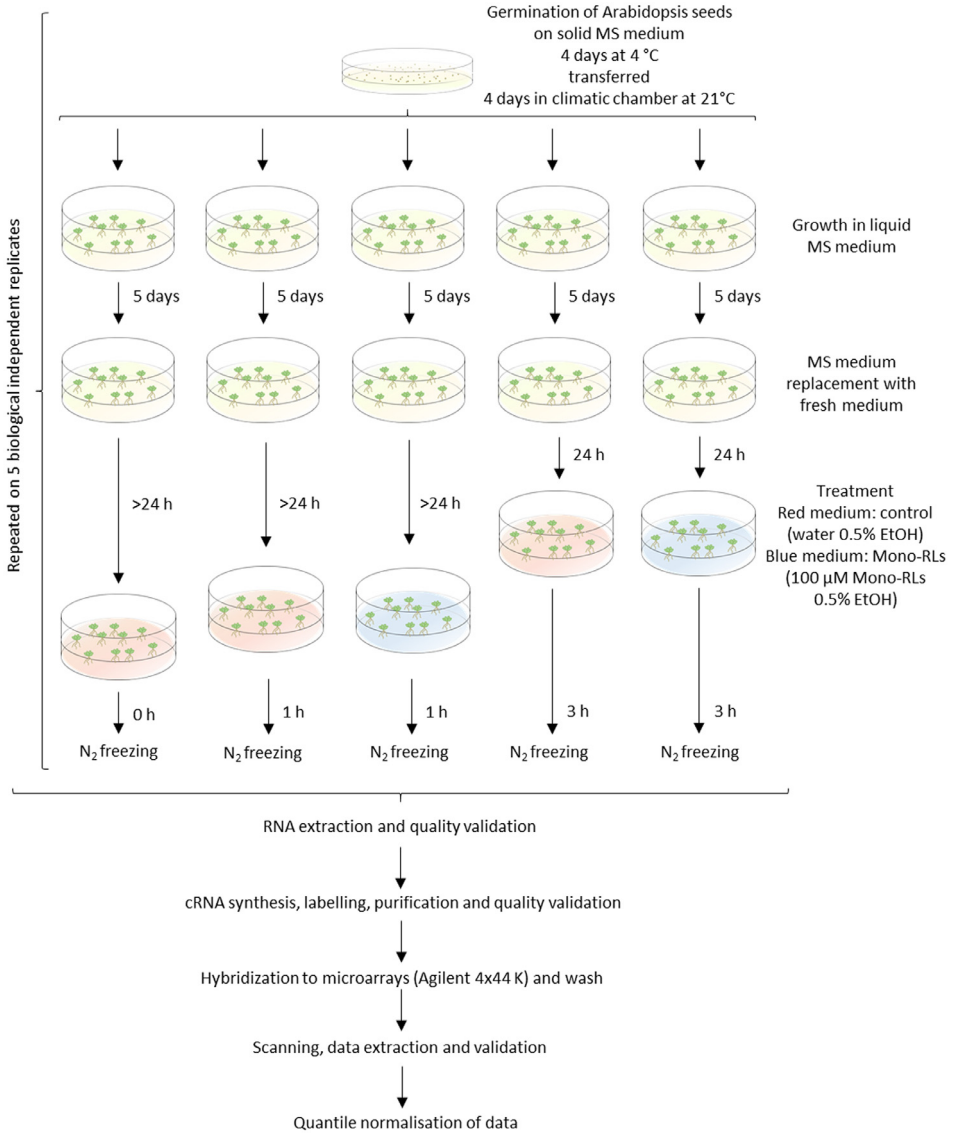

Fig. 1. Flowchart of experimental design of the study.

in 70% ethanol and 15 min in bleach (2.5% chloride) before being extensively rinsed with sterile water. They were then placed in petri dishes containing Murashige and Skoog [3] (MS) solid medium (4.3 g L^{-1} MS (ref M5524), 0.5 g L^{-1} 4-Morpholineethanesulfonic acid (MES), 5 g L^{-1} sucrose, 7 g L^{-1} agarose, all from Sigma Aldrich, Saint-Louis, USA), pH 5.7 (adjusted with KOH). Petri dishes were then placed for 4 days at 4 °C in the dark. For germination, seeds were placed in a climatic chamber at 21 °C, 60% relative humidity, 150 µmol m^{-2} s^{-1} light intensity with a 12-h photoperiod. After 4 days, 50 selected plants with similar development were divided in 5 groups of 10 plants, each one being placed in 1 mL of liquid MS medium (4.3 g L^{-1} MS, 0.5 g L^{-1} MES, 5 g L^{-1} sucrose). After 5 supplementary days into the climatic chamber, the growth medium of each group was replaced by 950 µL of fresh liquid MS medium. Treatments were

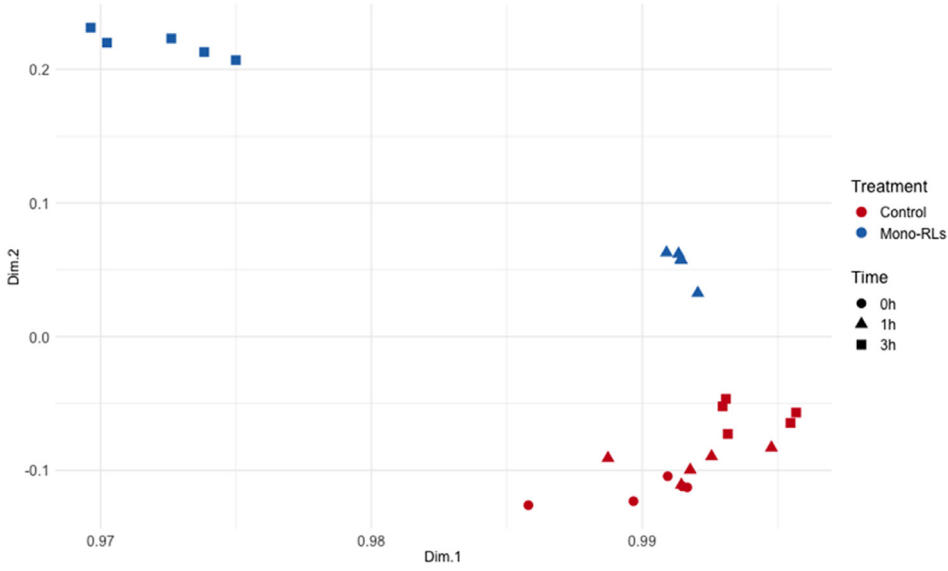

Fig. 2. PCA (principal component analysis) of the data from the different conditions.

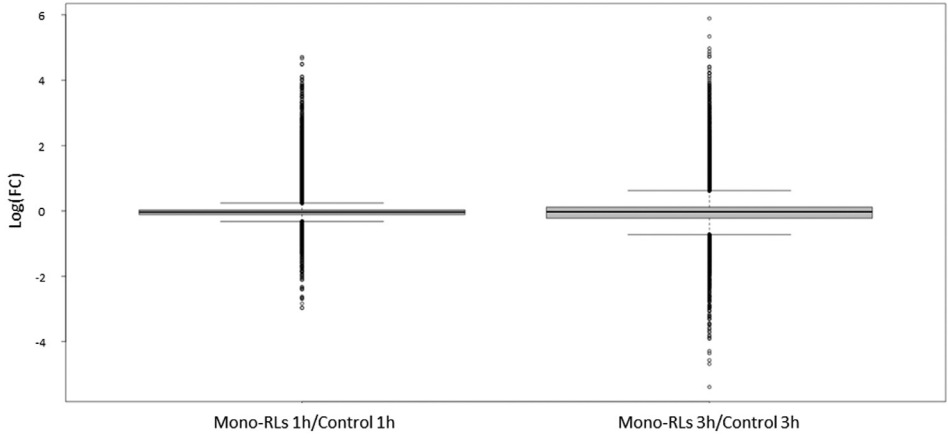

Fig. 3. Dispersion of all log(FC) (Fold Changes) obtained from seedlings treated by Mono-RLs 1h or 3h compared to the control conditions.

realized 24 h after, by adding 50 µL of a solution containing 10% of ethanol (control) or 2 mM Mono-RLs (in 10 % ethanol) to culture medium in order to reach a final concentration of 0.5% ethanol (control) or 100 µM Mono-RLs (in 0.5% ethanol). Treatments were performed starting with the 3 h treatment in order to collect all samples at the same time (the 1 h treatment started 2 h later than the 3 h treatment). Seedlings were delicately harvested with a clamp, dried on filter paper, immediately frozen in liquid nitrogen and stored at -80 °C until RNA extraction. Five independent replicates were carried out with 10 seedlings.

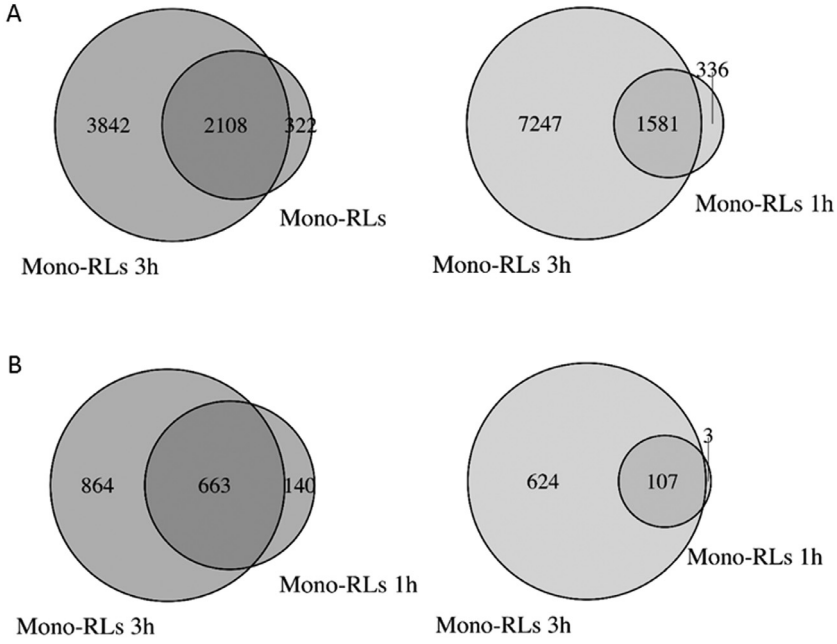

Fig. 4. (A) Venn diagrams of significantly regulated transcripts (p-value < 0.01). (B) Venn diagrams of regulated transcripts with a $|\log(FC)| > 1$ and p-value < 0.01. Orange color: up-regulated, green color: down-regulated transcripts. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.3. Total RNA extraction, cRNA preparation and purity analysis

For each condition and each replicate, the 10 seedlings harvested were crushed together with mortar and pestle. RNA extractions were then conducted with the RNeasy Plant mini kit (Qiagen, Hilden, Germany, Ref 74904) according to provider specifications, including the DNA digestion step with the RNase free DNase set (Qiagen, Ref 79254,). RNA concentration and purity were determined using Thermo Fisher (Waltham, USA) Scientific Nanodrop 2000. RNA integrity was checked using Bio-Rad (Hercules, USA) Experion™ and the corresponding kit Experion RNA StdSens Analysis (Bio-Rad, Ref 7007103) according to provider specifications. The gene expression study was realised using 100 ng RNA for each sample. The whole process was realized according to the Agilent protocol One-color Microarray-Based Gene Expression Analysis Low Input Quick Amp Labelling and the Agilent equipment (Agilent Technologies, Santa Clara, CA, USA). First, spike-in was added to RNA (Spike-In-One-Color, Agilent, Ref 5188–5282). cDNA were then synthetized and used as template for labelled and amplified cRNA using the Low Input Quick Amp Labelling Kit One-Color (Agilent, Ref 5190–2305). cRNA were then purified using the RNeasy Plant Mini Kit (Qiagen, Ref 74904). Purified cRNA were analyzed with a Nanometer 2000 (Thermo Fisher Scientific, Whaltam, USA). RNA integrity was validated by Experion™ chips. Purity (DO 260/DO 280 $\geq$2; 260/230 $\geq$2), yields ($\geq$1.65 ng µL x 10^{-3} elution volume in µL) and specific activities ($\geq$6 pmol Cy3 / µg RNA) were checked.

2.4. Microarray hybridization

The microarrays used were V4 Gene Expression Microarray, 4 $\times$ 44 K, (4 repeats of 43803 ssDNA probes of 60-mer covering all genome of *A. thaliana*, Agilent design ID 021169). cRNA were hybridized for 17 h on microarrays using the Gene Expression Hybridization Kit (Agilent, Ref

5188–5327) in the Shel Lab Agilent Microarray Hybridization Oven at 65 °C (ref G2545A, rotating system Ref G2530–6029). Hybridization chambers (ref G2534A) and Hybridization chambers gasket slides (Ref G2534–60011) were used. Deposits were organized by crossing biological independent experiences. The 5 repetitions on 7 arrays were hybridized in two times in order to minimize hybridizations. Slides were then washed 3 times with the Gene Expression Wash Buffer Kit (Agilent, ref 5188-5327).

2.5. Data acquisition and analysis

The arrays were immediately scanned with the SureScan Microarray Scanner (Agilent). The data extraction was performed with the Agilent software Feature Extraction version 11.5.1.1 using the default settings [4]. Quality Control (QC) reports generated showed that all evaluated metrics matched defined criteria. Processing of the dataset was realized using LIMMA (Linear Models for Microarray and RNA-Seq Data) package version 3.40.6 [5] of the bioconductor project [6] on R software version 3.6.1 [7]. The different steps were made following LIMMA user's guide [8] and in particular with the single-channel case study on Agilent arrays. Briefly, a background correction was realized before a quantile normalization between arrays using respectively the « backgroundCorrect » and «normalizeBetweenArrays» functions. Control probes and probes above background on less than five arrays were filtered out. No more transformation was done on data transmitted to GEO NCBI (except removing replicates as requested). To validate the experimental process, a principal component analysis was realized using the PCA function of FactorMineR package version 2.4 [9]. It confirmed three distinct groups of values corresponding to the control condition, the 1 h Mono-RL treatment and the 3 h Mono-RL treatment. Boxplots were generated with log(FC) values and were well centered on 0. Venn diagrams were realized, first considering probes with a corresponding p-value < 0.01 and with probes for which the condition $|\log(FC)| > 1$ is also respected.

Ethics Statement

No animal or human experiments were performed to obtain the data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

CRediT Author Statement

Noadya Monnier: Visualization, Formal analysis, Data curation, Writing – original draft, Writing – review & editing; **Catherine Sarazin:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing; **Sonia Rippa:** Conceptualization, Visualization, Formal analysis, Writing – original draft, Writing – review & editing.

Acknowledgments

We thank Dr Stéphan Dorey (Résistance Induite et Bioprotection des Plantes laboratory, Université Reims Champagne-Ardenne) for having provided us Mono-RLs purified by the Institut de Chimie Moléculaire de Reims. This research was supported by the Hauts-de-France (HdF) Council and European Regional Development Fund (ERDF), and was realized in the context of the

MAELIA project. N.M. Ph.D. thesis was co-funded by HdF Council and ERDF. HdF Council and ERDF also co-founded equipment utilized within CPER 2007-2020.

Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.dib.2021.107397.

References

[1] A.-L. Varnier, L. Sanchez, P. Vatsa, L. Boudesocque, A. Garcia-Brugger, F. Rabenoelina, A. Sorokin, J.-H. Renault, S. Kauffmann, A. Pugin, C. Clément, F. Baillieul, S. Dorey, Bacterial rhamnolipids are novel MAMPs conferring resistance to *Botrytis cinerea* in grapevine, Plant Cell Environ. 32 (2009) 178–193, doi:10.1111/j.1365-3040.2008.01911.x.

[2] N. Monnier, A. Furlan, C. Botcazon, A. Dahi, G. Mongelard, S. Cordelier, C. Clément, S. Dorey, C. Sarazin, S. Rippa, Rhamnolipids from *Pseudomonas aeruginosa* are elicitors triggering *Brassica napus* protection against *Botrytis cinerea* without physiological disorders, Front. Plant Sci. 9 (2018) 1170, doi:10.3389/fpls.2018.01170.

[3] T. Murashige, F. Skoog, A revised medium for rapid growth and bio assays with tobacco tissue cultures, Physiol. Plant 15 (1962) 473–497, doi:10.1111/j.1399-3054.1962.tb08052.x.

[4] Agilent, agilent feature extraction 12.0 reference guide, (2014) 328.

[5] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (2015) e47–e47, doi:10.1093/nar/gkv007.

[6] W. Huber, V.J Carey, R. Gentleman, S. Anders, M. Carlson, B.S. Carvalho, H. Corrada Bravo, S. Davis, L. Gatto, T. Girke, R. Gottardo, F. Hahne, K.D. Hansen, R.A. Irizarry, M. Lawrence, M.I. Love, J. MacDonald, V. Obenchain, A.K. Oleś, H. Pagès, A. Reyes, P. Shannon, G.K. Smyth, D. Tenenbaum, L. Waldron, M. Morgan, Orchestrating high-throughput genomic analysis with Bioconductor, Nat. Methods 12 (2015) 115–121, doi:10.1038/nmeth.3252.

[7] R. R core TeamR: A Language and Environment for Statistical Computing, R Found. Stat. Comput., Vienna, Austria, 2014, doi:10.1017/CBO9781107415324.004.

[8] G.K. Smyth, M. Ritchie, N. Thorne, Linear models for microarray data user' s guide, Bioinformatics 20 (2011) 3705–3706.

[9] F. Husson, S. Lê, J. Pagès, in: Exploratory Multivariate Analysis by Example Using R, 40, 1st Edition, 2010, pp. 1–225, doi:10.1201/b10345.