Data Article

Microarray data and pathway analyses of peripheral blood mononuclear cells from healthy subjects after a three weeks grape-rich diet

Rosa Anna Milella [a],[**], Marica Gasparro [a],
Fiammetta Alagna [a],[1], Maria Francesca Cardone [a],
Silvia Rotunno [a],[2], Concetta Tiziana Ammollo [b],
Fabrizio Semeraro [b], Apollonia Tullo [c], Flaviana Marzano [c],
Domenico Catalano [d], Donato Antonacci [a], Mario Colucci [b],
Domenica D'Elia [d],[*]

[a] Research Centre for Viticulture and Enology, Council for Agricultural Research and Economics, Turi, 70010, Bari, Italy
[b] Department of Biomedical Sciences and Human Oncology, University of Bari "Aldo Moro", 70124, Bari, Italy
[c] Institute of Biomembranes, Bioenergetics and Molecular Biotechnologies, National Research Council, 70126, Bari, Italy
[d] Institute for Biomedical Technology, National Research Council, 70126, Bari, Italy

ARTICLE INFO

ABSTRACT

Using Human Gene Expression Microarrays (Agilent) technologies, we investigated changes of the level of gene expression in peripheral blood mononuclear cells of healthy subjects after 21 days of fresh table grape-rich diet and after an additional 28-day washout. Several hundreds of genes were differentially expressed after grape intake or after washout. The functional analysis of these genes detected significant changes in key processes such as inflammation and immunity, thrombosis, DNA and protein repair, autophagy and mitochondrial biogenesis. Moreover, fresh grape intake was found to influence the expression of many long non-

Diets
Human health
Health benefits

coding RNA genes. The data can be valuable for researchers interested in nutrigenetics and nutrigenomics studies and are related to the research article "Gene expression signature induced by grape intake in healthy subjects reveals wide-spread beneficial effects on PBMCs" [1].

Specifications Table

Subject	Nutrition
Specific subject area	Nutrigenomic science
Type of data	Table
	Chart
How data were acquired	SurePrint G3 Human Gene Expression Microarray (Agilent Technologies), GeneSpring software (Agilent) - version 13.1.1,
	Gene Set Enrichment analysis (g:GOSt)
	Incrementally profiling of ordered gene lists (g:Cocoa)
Data format	Raw
	Analysed
	Filtered
Parameters for data collection	Peripheral blood mononuclear cells of healthy subjects were collected before starting the grape diet (T0), at the end of the 21-day grape diet (T1), and after 4-week washout period (T2)
Description of data collection	Blood samples, RNA isolation from peripheral blood mononuclear cells, microarray analysis, statistical analysis, functional analysis
Data source location	Institution: Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria, Unità di ricerca per l'uva da tavola e la vitivinicoltura in ambiente mediterraneo (UTV)
	City/Town/Region: Turi, Bari, Apulia
	Country: Italy
	Latitude and longitude (and GPS coordinates) for collected samples/data: 40.921336, 17.010942
Data accessibility	Data is within this article
Related research article	Rosa Anna Milella, Marica Gasparro, Fiammetta Alagna, Maria Francesca Cardone, Silvia Rotunno, Concetta Tiziana Ammollo, Fabrizio Semeraro, Apollonia Tullo, Flaviana Marzano, Domenico Catalano, Donato Antonacci, Mario Colucci, Domenica D'Elia
	Gene expression signature induced by grape intake in healthy subjects reveals wide-spread beneficial effects on PBMCs
	Journal of Functional Foods
	Vol. 64
	Year: 2020
	DOI: https://doi.org/10.1016/j.jff.2019.103705

Value of the Data

- Data provided by this work are the first, of this type, obtained by a study in vivo on healthy subjects in response to a daily supplementation of fresh table grape.
- Other researchers can benefit from these data for comparative analysis or integrative omics studies with data from similar or complementary experiments with the same or similar type of nutrients on healthy subjects or patients.
- Data produced by this work can be used for validation experiments or system biology studies to elucidate the specific role, and molecular mechanisms, of grape on the expression of genes with a pivotal role on chronic inflammatory diseases, cardiovascular disorders or metabolic disorders such as diabetes and obesity.
- The additional value of results provided resides in two main findings not previously reported in the literature, the long-term effect of grape on gene expression and the unexpected number of long non-coding RNAs involved.
- LncRNAs are well known to have regulatory functions at genomic, transcriptomic and translational level. Many of the lncRNAs differentially expressed under the effects of grape intake are still uncharacterised. These data may contribute to their functional characterisation.
- Data obtained by these experiments can be useful to demonstrate the potential role of nutraceuticals derived from grape in preventing chronic diseases and in contributing to cell homeostasis and human health.

1. Data description

Microarray analyses of mRNAs isolated from peripheral blood mononuclear cells (PBMC) of six healthy subjects at baseline (T0), after 21 days of grape-rich diet (T1) and after a 4-week washout period (T2) showed altered expression (*p-value* ≤ 0.05 and Fold change ≥ 1.5) of several genes (Supplemental material, Table S1) [1]. Clustering analysis of these genes identified five main sub-clusters (Supplemental material, Table S2). The analysis of subcluster 3, the most influenced by grape intake, with g:GOSt highlighted the statistical enrichment of numerous down-regulated genes in biological processes or pathways related to the innate and cell-mediated immune response and to the regulation of signal transduction (Supplemental material, Table S3). The functional enrichment analysis of all differentially expressed genes was performed grouping genes in three classes (i.e., Class A, B and C) using the gene grouping obtained by the Venn diagram (Supplemental material, Table S4) and g:GOSt (Supplemental material, Table S5) for the statistical enrichment. Moreover, a comparative analysis of the three gene Class lists was performed using g:Cocoa and results are provided in two different formats (Excel: Supplemental material, Table S6; and pdf: Supplemental material, File S1). Microarray analysis revealed significantly altered expression of many long non-coding RNAs (Supplemental material, Table S7).

2. Experimental design, materials, and methods

2.1. Study design

Twenty healthy volunteers (10 females and 10 males; aged 24−48) consumed 5g of fresh grape per kg of body weight per day of black grape Autumn royal for three weeks. Blood samples were collected from overnight fasting subjects before starting the grape diet (T0), at the end of the 21-day grape diet (T1), and after the 4-week washout period (T2). PBMCs were isolated by using the Fycoll−Hypaque method.

2.2. RNA isolation and microarray analysis

Total RNA was isolated from PBMCs using the PureLink® RNA Mini Kit (Ambion by Life Technologies, Carlsbad, CA, USA). The RNA concentration and purity were determined spectrophotometrically at 230, 260 and 280nm using a NanoDrop 2000 UV−Vis Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). The RNA integrity was measured with the 2100 Bioanalyser using the RNA 6000 Nano LabChip kit (Agilent Technologies, Santa Clara, CA, USA).

Samples from six subjects (three males and three females), with a ratio A260/A280 between 1.8 and 2.0, an A260/A230 ratio ≥ 2.0 and an RNA Integrity Number ≥ 7, were generated using the Low Input Quick Amp Labeling kit, according to the Two-Color Microarray-Based Gene Expression Analysis (Version 6.7) manufacturer's protocol (Agilent Technologies, Santa Clara, CA, USA).

Each RNA sample was labelled with the dyes swapped to eliminate artefacts resulting from variation in the different dyes' incorporation. Hybridisation was carried out on SurePrint G3 Human Gene Expression Microarrays, 8 × 60K v2 (Agilent Technologies, Santa Clara, CA, USA).

In the paired study design, each subject acted as its own control: on the same chip we have hybridised one array with Cy3-cRNA at T0 vs Cy5 cRNA at T1 of one subject and another array with Cy3-cRNA at T1 vs Cy5 cRNA at T2 of the same subject. We have also hybridised the respective dyes swapped on the same chip to rule out any possible dye effect.

2.3. Statistical analysis

The row-data, produced by the array experiments, were imported in the GeneSpring software grouped using the time point (T0, T1, T2). To evaluate the reproducibility and reliability of data, a Principal Component Analysis (PCA) was carried out to calculate the PCA scores. This allowed us to identify if some of the samples were an outlier with respect to the others. Subsequent analyses were

performed following the GeneSpring 13.5 wizard pipeline, including statistical steps and associated results, typically entity lists. To identify differentially expressed genes, the fold change analysis was performed using cut-off = 1.5 and choosing as condition the time-points. Time intervals considered were T1vsT0, T2vsT0 e T2vsT1, and the analysis repeated for each sample treated as biological replicates. Genes up- and down-regulated as resulted by the Fold change analysis, were selected for the calculation of significant p-values using a one-way ANOVA test (choosing the asymptotic p-value computation algorithm) followed by post hoc Tukey's Honestly Significant Difference test (p-value $\leq$ 0.05) corrected by the application of Benjamini-Hochberg multiple testing for the false discovery rate. Filtered gene lists were generated for expression changes greater than 1.5-fold and a p-value $\leq$ 0.05 (Supplementary material S1).

2.4. Bioinformatic functional analysis

For the functional interpretation of results, we used the Gene Group Functional Profiling (g:GOSt) and the Compact Compare of Annotations (g:Cocoa) tools available in the g:Profiler web server (update 2018-10-02) [2]. g:GOSt performs statistical enrichment analysis of single-gene lists and returns statistics using multiple sources of functional evidence, including Gene Ontology terms, biological pathways, regulatory motifs of transcription factors and microRNAs, human disease annotations and protein-protein interactions. g:Cocoa performs functional analysis of multiple gene lists. In our case, g:Cocoa was used for the comparative analysis of differentially expressed gene lists at different timeframes. The hierarchical clustering analysis was performed with the GeneSpring software (Agilent) using the Euclidean distance metric on conditions and entities and Average linkage rule [1]. K-means associates entities into groups to produce groups of entities with a high degree of similarity within each group and a low degree of similarity between groups. This analysis clustered gene differentially expressed in five subgroups whose significant enrichment in biological processes and pathways was performed using the g:GOSt tool. Lists of up- and down-regulated genes at different timeframes (T1vsT0, T2vsT0 e T2vsT1), as returned by the statistical analysis in GeneSpring, where analysed separately also considering the specific timeframe during which they showed an altered expression as returned by the analysis carried out with the Venn diagram tool. The gene lists obtained, six in total, were used separately for the functional analysis in g:GOSt and all together for the comparative analysis in g:Cocoa. Parameters, options and sources used for the analysis are included in Supplementary Tables as annotated by g:Profiler.

Acknowledgements

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2020.105278.

References

[1] R.A. Milella, M. Gasparro, F. Alagna, M.F. Cardone, S. Rotunno, C.T. Ammollo, F. Semeraro, A. Tullo, F. Marzano, D. Catalano, D. Antonacci, M. Colucci, D. D'Elia, Gene expression signature induced by grape intake in healthy subjects reveals wide-spread beneficial effects on PBMCs, J. Funct. Foods 64 (2020) 103705, https://doi.org/10.1016/j.jff.2019.103705.

[2] J. Reimand, T. Arak, P. Adler, L. Kolberg, S. Reisberg, H. Peterson, et al., g:Profiler-a web server for functional interpretation of gene lists (2016 update), Nucleic Acids Res. 44 (W1) (2016) W83—W89, https://doi.org/10.1093/nar/gkw199.