

# DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways

Ioannis S. Vlachos<sup>1,2,3</sup>, Nikos Kostoulas<sup>4</sup>, Thanasis Vergoulis<sup>4,5</sup>,  
Georgios Georgakilas<sup>1,2</sup>, Martin Reczko<sup>1,6</sup>, Manolis Maragkakis<sup>1,7</sup>,  
Maria D. Paraskevopoulou<sup>1,2</sup>, Kostantinos Prionidis<sup>1,8</sup>, Theodore Dalamagas<sup>4,5,\*</sup> and  
Artemis G. Hatzigeorgiou<sup>1,\*</sup>

<sup>1</sup>DIANA-Lab, Institute of Molecular Oncology, Biomedical Sciences Research Center 'Alexander Fleming', 16672 Vari, <sup>2</sup>Department of Informatics and Telecommunications, Postgraduate Program in 'Information Technologies in Medicine and Biology', University of Athens, 15784 Athens, <sup>3</sup>Laboratory for Experimental Surgery and Surgical Research 'N.S. Christeas', Medical School of Athens, University of Athens, 11527 Athens, <sup>4</sup>IMIS Institute, 'Athena' Research Center, 11524 Athens, <sup>5</sup>Knowledge and Database Systems Lab, Department of Computer Science, School of Electrical and Computer Engineering, National Technical University of Athens, 15773 Zografou, <sup>6</sup>Synaptic Ltd., 71110 Heraklion, Greece, <sup>7</sup>Department of Pathology and Laboratory Medicine, Division of Neuropathology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA and <sup>8</sup>Department of Computer and Communication Engineering, University of Thessaly, 38221, Volos, Greece

Received March 3, 2012; Revised May 2, 2012; Accepted May 4, 2012

## ABSTRACT

MicroRNAs (miRNAs) are key regulators of diverse biological processes and their functional analysis has been deemed central in many research pipelines. The new version of DIANA-miRPath web server was redesigned from the ground-up. The user of DNA Intelligent Analysis (DIANA) DIANA-miRPath v2.0 can now utilize miRNA targets predicted with high accuracy based on DIANA-microT-CDS and/or experimentally verified targets from TarBase v6; combine results with merging and meta-analysis algorithms; perform hierarchical clustering of miRNAs and pathways based on their interaction levels; as well as elaborate sophisticated visualizations, such as dendrograms or miRNA versus pathway heat maps, from an intuitive and easy to use web interface. New modules enable DIANA-miRPath server to provide information regarding pathogenic single nucleotide polymorphisms (SNPs) in miRNA target sites (SNPs module) or to annotate all the predicted and experimentally validated miRNA targets in a selected molecular pathway (Reverse Search

module). DIANA-miRPath v2.0 is an efficient and yet easy to use tool that can be incorporated successfully into miRNA-related analysis pipelines. It provides for the first time a series of highly specific tools for miRNA-targeted pathway analysis via a web interface and can be accessed at <http://www.microrna.gr/miRPathv2>.

## INTRODUCTION

The discovery of microRNAs (miRNAs) in *Caenorhabditis elegans* in 1993 opened the door for the study of phenomena, such as RNA interference and paved the way for the RNA revolution. miRNAs are ~23-nt long single-stranded RNA species that regulate post-transcriptionally protein coding genes by mRNA cleavage, direct translational repression and/or mRNA destabilization (1).

miRNAs have been deemed as key regulators of diverse biological processes, including development, stem cell proliferation, division and differentiation, regulation of innate and adaptive immunity, apoptosis, cell signaling and metabolism (2–4). miRNAs play an important role in a vast array of human pathologies, such as cancer, viral infections, cardiovascular diseases, metabolic

\*To whom correspondence should be addressed. Tel: +30 210 9656310 (ext.190); Fax: +30 210 9653934; Email: [artemis@fleming.gr](mailto:artemis@fleming.gr)  
Correspondence may also be addressed to Theodore Dalamagas. Tel: +30 210 6875415; Fax: +30 210 6856804;  
Email: [dalamag@imis.athena-innovation.gr](mailto:dalamag@imis.athena-innovation.gr)

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

© The Author(s) 2012. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

disorders, autoimmune pathologies, as well as neuropsychiatric pathological conditions (2,3,5–9). They are intensely researched, not only for their role in physiology and pathology but also as potential therapeutic targets.

A long series of algorithmic tools and databases have been implemented and utilized by miRNA researching groups as tools of the trade (10–12). Gene target prediction applications were some of the first sectors to receive intense research focus (10). The tools that followed were designed to address the need for functional characterization of miRNAs and their predicted targets. Few of the available servers are specifically focused on the identification of miRNA targeted pathways (e.g. miTalos, DIANA-miRPath v1) (13,14), while others provide a relevant functionality as an extension of a pre-existent pipeline (e.g. mirTar, Gene Trail) (15,16).

miTalos (13) can be used for the analysis of a subset of human signaling pathways. It identifies targets using five different external miRNA target prediction algorithms while also considering expression data. miRTar (16) can be utilized for investigating alternatively spliced miRNA targets, which are identified by integrating external prediction algorithms. miRTar can also perform gene set enrichment analysis for the identification of miRNA targeted pathways. GeneTrail (15) is a web server hosting gene set enrichment and over-representation capabilities against various databases, such as GO and KEGG. GeneTrail has been extended with a tool, which queries miRNA identifiers against the MicroCosm Targets database for putative gene targets (17). It provides many options during the gene enrichment analysis process, such as *P*-value thresholds and multiple testing significance correction but is lacking miRNA-specific functionalities, such as information regarding binding positions and binding type.

DIANA-miRPath v1 (14) was one of the first available applications focused on the enrichment analysis of predicted target genes, capable of detecting pathways targeted by single or multiple miRNAs. Here, we present the next version of DIANA-miRPath webserver (v2.0), which was entirely redesigned. The new web server aims to significantly increase the accuracy of utilized algorithms and statistics, as well as to enhance computational speed, compared to the previous miRPath version. Importantly, DIANA-miRPath v2.0 offers for the first time a series of tools specifically focused on miRNA-targeted pathway analysis.

The user of DIANA-miRPath v2.0 can utilize predicted or experimentally validated targets; combine results with merging and meta-analysis algorithms; perform hierarchical clustering of microRNAs and pathways based on their interaction levels; as well as elaborate sophisticated visualizations, such as dendrograms or miRNA/pathway interaction heat maps, from an intuitive and easy to use web interface. The new server provides additional information regarding pathogenic SNPs in predicted miRNA target sites. Furthermore, the reverse analysis module annotates all predicted or experimentally validated miRNAs targeting a selected molecular pathway.

Specific details regarding design, implementation and use of DIANA-miRPath v2.0 are presented in the following section.

## METHODS AND RESULTS

### miRPath v2.0 database

DIANA-miRPath v2.0 is based on a new relational schema, specifically designed to accommodate this as well as future miRPath updates. miRNA and pathway related information was obtained from miRBase 18 (11) and Kyoto Encyclopedia of Genes and Genomes (KEGG) v58.1 (18), respectively, for *Mus musculus* and 'homo sapiens'.

In DIANA-miRPath v2.0, the user can select between *in silico* predicted miRNA gene targets and a large set of experimentally validated targets, or both. The *in silico* target prediction is performed using DIANA-microT-CDS (19), the latest version of the DIANA-microT algorithm, which computes a miRNA–gene interaction score for predicted targets in 3'-UTR and coding sequence (CDS) transcript regions. DIANA-microT-CDS has been tested against a variety of high-throughput experimental data, exhibiting the highest sensitivity and specificity in comparison with other broadly used miRNA target prediction algorithms (19). DIANA-microT-CDS is fully integrated in the new database schema, enabling the filtering of miRNA–gene interactions based on a user defined threshold. The server is using 0.8 score as a default threshold, which provides an average of 350 targets per miRNA. Information regarding the number of targets and the derived predictive accuracy for different thresholds, as calculated against experimentally verified miRNA targets (20) is provided in Supplementary Table S1.

TarBase v6.0 database (12) is also incorporated into the new schema. Tarbase v6.0 is currently the largest manually curated, experimentally validated miRNA–gene interactions database, hosting more than 65 000 miRNA–gene interactions; 16.5- to 175-fold more than any other available manually curated database.

### Pathogenic single nucleotide polymorphisms identification in predicted miRNA target sites

The coincidence of clinically significant single nucleotide polymorphisms (SNPs) and predicted miRNA target sites was calculated, in order to relate miRNA gene targets in molecular pathways to pathogenic polymorphisms. A total of 9899 experimentally verified human SNPs, specifically indicated as 'pathogenic' or 'probable pathogenic' and with known genomic coordinates were derived from dbSNP BUILD 135 (21). The SNPs' coordinates are overlapped with all miRNA target binding sites predicted with DIANA-microT-CDS, resulting in 16 578 miRNA target sites containing a pathogenic or probable pathogenic SNP. The user has access to the results of the pathogenic SNP analysis directly from the DIANA-miRPath v2.0 web server (Supplementary Figure S2). The web server provides the dbSNP id, as well as a direct link to the relevant dbSNP entry, enabling the user to further

assess the available evidence supporting the pathogenicity of the detected SNP.

### Algorithms, implementation and results computation

The DIANA-miRPath v2.0 website is implemented by combining PHP and AJAX. It is specifically designed to provide results in real-time, offering an application-like user experience. All the statistical and analysis algorithms are implemented in R.

### Filtering based on expression data

DIANA-miRPath v2.0 offers the option to filter miRNA targets based on expression data. The user can upload a predefined list of genes that are expressed in investigated tissues. The program will filter all miRNA targets based on this list and will use only the expressed subset of genes for the pathway enrichment analysis.

### Enrichment analysis

The user can select the desired microRNAs by the web interface or by uploading a file with the desired miRNA identifiers.

Following the selection of miRNAs and (optional) genes sets, the server subsequently performs an enrichment analysis of microRNA gene targets in KEGG pathways (Table 1). The default algorithm utilized is the one-tailed Fisher's exact test, which is often performed in such context (22,23). The one-tailed test is selected because the initial hypothesis was to detect enriched (and not depleted) pathways with targets of specific miRNAs. This test is often encountered in relevant literature also as hypergeometric test. It provides exact probabilities and is appropriate even for small number statistics (23). DIANA-miRPath v2.0 can also calculate a more conservative adjustment of Fisher's exact test (24). The derived score represents the upper bound of the distribution of the jack-knife Fisher exact probabilities and it favors pathways supported with more gene targets.

The server offers to the user the option to perform the false discovery rate (FDR) method, as a correction for multiple hypothesis testing. The FDR algorithm is implemented as described by Benjamini and Hochberg (25). In both cases (corrected or uncorrected significance levels), the user can specifically set the desired level of significance and to filter out all pathways with an associated  $P$ -value > threshold from visualizations or further analyses.

**Table 1.** Over-representation analysis performed by the DIANA-miRPath web server

|              | $\in$ Pathway A | $\notin$ Pathway A | Total |
|--------------|-----------------|--------------------|-------|
| Targeted     | $n_{1+}$        | $n_{2+}$           | $N_+$ |
| Non-targeted | $n_{1-}$        | $n_{2-}$           | $N_-$ |
| Total        | $n_1$           | $n_2$              | $N$   |

$n_1$ : number of nodes in pathway A,  $n_{1+}$ : number of targeted nodes in pathway A,  $n_{1-}$ : number of non-targeted nodes in pathway A,  $n_2$ : number of nodes  $\notin$  Pathway A,  $n_{2+}$ : number of targeted nodes  $\notin$  Pathway A,  $n_{2-}$ : number of non-targeted nodes  $\notin$  Pathway A,  $N_+$ : all targeted nodes,  $N_-$ : all non-targeted nodes,  $N$ : all nodes.

### Multiple miRNA effect analysis

The new server can accommodate many different analysis scenarios. Specific attention has been paid to the methodologies used for combining and merging predicted and/or experimentally validated gene targets of multiple miRNAs. The user can now select how gene targets or targeted pathways will be combined and even if the results combination will take place *a priori* (Union and Intersection of Genes) or *a posteriori* of the statistical analysis (Union and Intersection of Pathways). Currently, the supported analysis scenarios are as follows.

**Union of Genes:** this option utilizes the union of targeted genes by the selected microRNAs (genes targeted by at least one selected miRNAs) prior to the statistical calculation. The resulting targeted genes superset is used for the over-representation statistical analysis.

**Intersection of Genes:** in this analysis, only the intersection of targeted genes (genes targeted by all selected miRNAs) is utilized for the over-representation statistical analysis.

**Union of Pathways:** in this mode, the server identifies all the significantly targeted pathways by the selected microRNAs.

**Intersection of Pathways:** the final option provides the intersection of targeted pathways by the selected microRNAs. The resulting subset contains only the pathways with statistically significant results for all the selected microRNAs.

In the first two options (Intersection and Union of Genes), the server combines the gene targets of the selected microRNAs into a common subset (intersection) or a superset (union). The resulting set is then incorporated in the enrichment analysis. In the last two options (Union and Intersection of Pathways), the server initially calculates the significance levels between all possible microRNA–pathway pairs, by performing repeatedly the enrichment analysis against the selected microRNAs and all available pathways. In the second step of the process, the server combines the previously calculated significance levels and provides a merged  $P$ -value for each pathway, by applying Fisher's combined probability method (Fisher's method).

Fisher's method is a meta-analysis algorithm, which can be used to combine the results of more than one independent tests bearing upon the same hypothesis. It exhibits asymptotical optimality and has been successfully used in studies of medical as well as biological context (26–28). The meta-analysis' null hypothesis is that all of the separate null hypotheses are true; while the alternative hypothesis is that at least one of the separate alternative hypotheses is true. The resulting  $P$ -values depict the probability that the examined pathway is significantly enriched with gene targets of the selected miRNAs. By adopting this technique, DIANA-miRPath v2.0 provides accurate statistical results and  $P$ -values in all available analysis options.

All provided results can be also corrected in the combined enrichment analysis for multiple hypotheses testing, by applying Benjamini and Hochberg's FDR.

### Results visualization

The new DIANA-miRPath v2.0 interface (Figure 1) has been designed to be highly adaptable to different use-case

The screenshot shows the DIANA miPath v2.0 web interface. At the top, there's a red header with 'DIANA LAB' and 'DNA Intelligent Analysis'. Below this is a navigation bar with links: HOME, SOFTWARE, DATABASES, MEMBERS, PUBLICATIONS. The main content area is titled 'Mirpath' and includes a brief description of the tool. The interface is divided into several sections:
 

- Species:** A dropdown menu set to 'Human'.
- Gene filter:** A link to 'determine a list of genes (optional)'.
- Add miRNAs:** A text input field, a 'Tarbase' dropdown, and a '+ or upload a file' button.
- Selected miRNAs (lists):** A list of miRNAs: hsa-let-7a-5p, hsa-miR-21-5p, and hsa-miR-140-5p. Each has a 'microT-CDS' dropdown (set to 'Tarbase') and a 'disable' link.
- Add extra miRNAs (lists):** A section for adding more miRNAs.
- Show genes for this list:** A link to view genes for the selected miRNAs.
- Select the way to merge results:** Radio buttons for 'genes union', 'genes intersection' (selected), and 'pathways union'.
- P-value threshold:** A text input set to '0.05' with an 'Apply' button.
- MicroT threshold:** A text input set to '0.8' with an 'Apply' button.
- Heatmap & clustering options:** Radio buttons for 'Significance Clusters/Heatmap' (selected) and 'Targeted Pathways Clusters/Heatmap'.
- Download results in file:** A link to download the results.
- KEGG pathway table:** A table with columns: #, KEGG pathway, p-value, #genes, #miRNAs, and details. It lists two pathways: 'ECM-receptor interaction (hsa04512)' and 'Glycosaminoglycan biosynthesis - heparan sulfate (hsa00534)'.
- List of related miRNAs:** A link to view related miRNAs for a specific pathway.
- Show/Hide miRNAs related to the pathway:** A link to toggle the visibility of related miRNAs.

**Figure 1.** From the DIANA-miPath v2.0 interface, the user can select the microRNAs that will be included in the analysis or upload miRNA lists in the form of text files. The user can select if miRNA gene targets will be experimentally validated (derived from TarBase 6) or predicted (derived from DIANA-microT-CDS). Optionally, the user can upload a predefined list of genes expressed in investigated tissues, which will be used to focus the enrichment analysis only on the specified subset. Subsequently, the user can determine the result merging method and statistics/enrichment calculation methodologies. The number of provided results, as well as the sensitivity and specificity of the DIANA-microT-CDS prediction algorithm can be set by user-defined thresholds. By selecting pathways union/intersection merging methods, the user obtains access to the advanced visualizations, which include miRNA/pathway clusters and miRNAs versus pathways heat maps. All significantly targeted pathways, with *P*-values lower than the user-defined threshold are presented in the interactive table. Pathway names, KEGG ids, significance levels, number of miRNAs targeting each pathway and targeted genes are some of the provided information. The table provides also access to enriched KEGG representations, DIANA-microT-CDS prediction details and experimental validation information, in the case of TarBase derived targets. The reverse search module can be used to detect miRNAs targeting (experimentally validated or predicted) a specified pathway. All results can be downloaded in a portable .csv format.

scenarios and to provide results in real time. In order to perform the analysis, the user can select one or more microRNAs and the source of gene targets for each miRNA (*in silico* predicted or experimentally validated). Optionally, a list of expressed genes can be also loaded. Subsequently, the server presents the significantly enriched pathways, the targeted genes in each pathway and the number of miRNAs with positively identified targets for each pathway in the form of an interactive table.

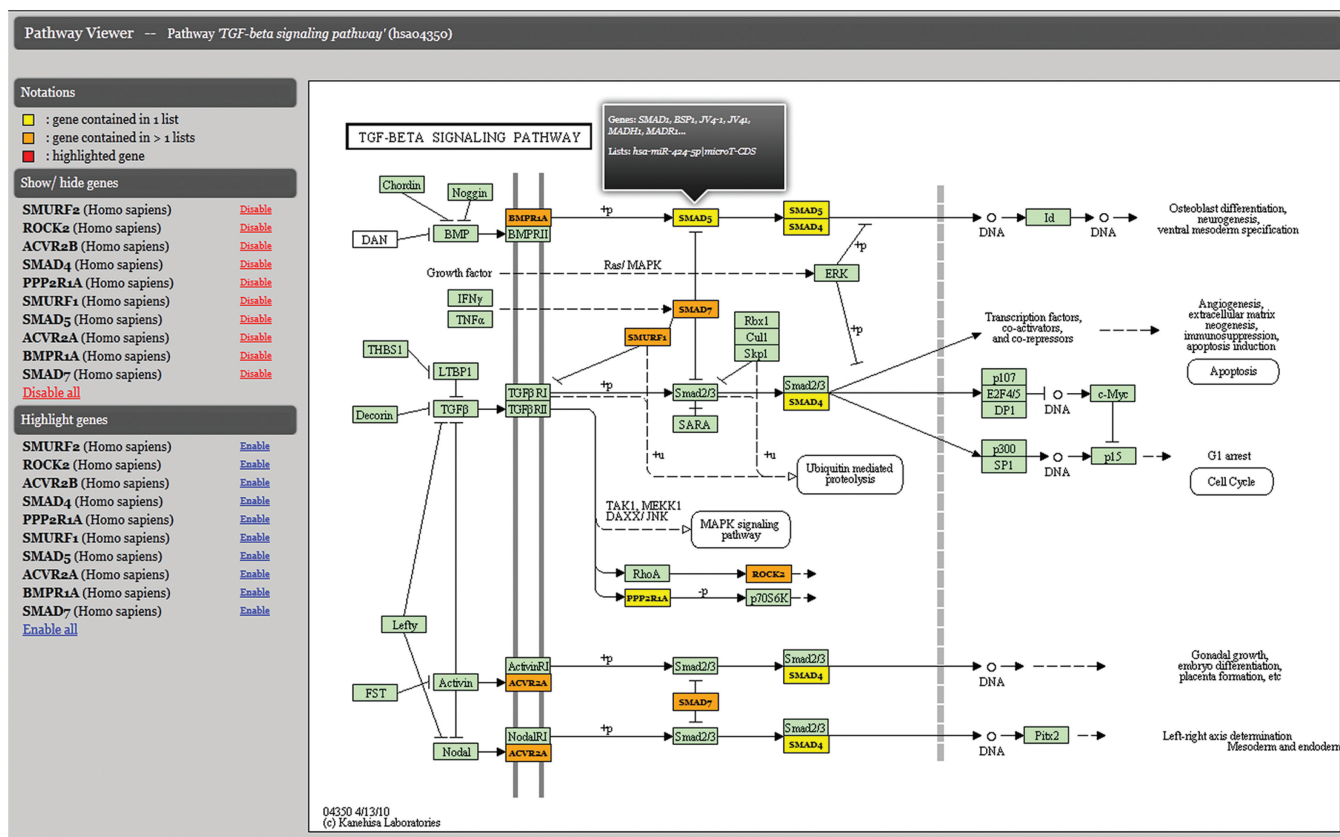
In the case of predicted miRNA–gene interactions, the server provides a link to the relevant DIANA-microT-CDS server entries. There, the user can further inspect the predicted miRNA–gene interaction. Such interactions include the binding region, position and type. If a miRNA–gene interaction is experimentally validated, the server provides a link to the specific section of the TarBase 6.0 website. The relevant entry provides information

regarding the implemented experimental method used for validation and the supporting literature.

DIANA-miPath v2.0 offers enriched KEGG pathway visualizations, where the targeted genes are specifically marked for easier inspection (Figure 2).

#### Targeted pathways (reverse search module)

The new reverse search module can be used to identify all miRNAs which are predicted or experimentally validated to target a specific KEGG pathway. The module takes as input a KEGG pathway name or identifier and the source of miRNA targets. It subsequently identifies all the miRNAs targeting the selected pathway. The new module can become a powerful asset to scientists studying specific pathways. It can help examining validated relationships between pathways and miRNAs expressed in the available literature (TarBase targets) or to study



**Figure 2.** DIANA-miRPath v2.0 offers enriched KEGG pathway visualizations, where the targeted genes are specifically marked for easier inspection. The server provides three levels of gene labeling: yellow (gene targeted by 1 selected miRNA), orange (gene targeted by >1 selected miRNAs) and red (gene specifically marked by the user). The user can also enable/disable gene marking and hide/show targeted genes. By simply hovering over a target gene (tooltip), the web server provides information regarding the source of the interaction (TarBase or DIANA-microT-CDS) and the implicated miRNAs. Selection of any of the pathway's constituents will lead the user directly to the relevant entry on the KEGG website.

novel miRNA–pathway interactions (DIANA-microT-CDS targets). If the analysis is performed *in silico*, the user can determine the desired levels for sensitivity and precision by applying a DIANA-microT-CDS score threshold.

### Advanced analysis pipelines and visualizations

DIANA-miRPath v2.0 provides also advanced features, statistics and visualization aids, which significantly increase the depth of the analysis and maximize the user's influence on results' calculation and presentation.

### Hierarchical clustering and dendrograms

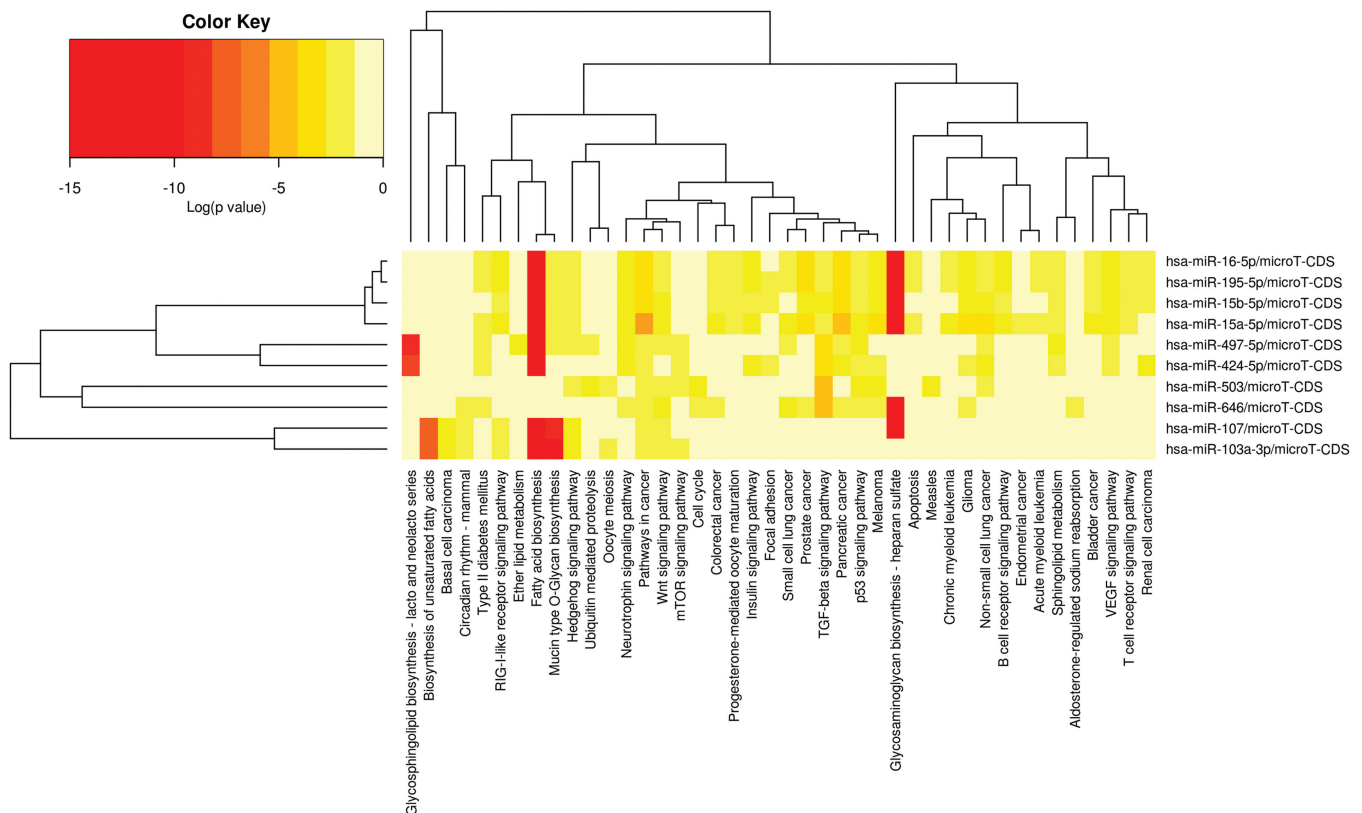
The new web server can perform follow-up analyses, such as hierarchical clustering of targeted pathways and miRNAs. DIANA-miRPath v2.0 realizes clustering of the selected miRNAs based on their influence on molecular pathways. It provides clustering of pathways based on the subset of miRNAs that target each pathway and the significance level of the interaction. The server performs the hierarchical cluster analysis based on a complete linkage clustering method, where squared Euclidean distances are calculated as distance measures. The web server can utilize absolute *P*-values in all calculations (option: 'Significance Clusters') or binary values (0: not targeted,

1: targeted), if the option 'Targeted Pathways Clusters' is selected. By utilizing these options, the algorithm can cluster together microRNAs targeting similar lists of pathways, as well as pathways, which are targeted by similar lists of microRNAs (Targeted Pathways Clusters); or take also into account the significance levels of the interactions (Significance Clusters) during the clustering process.

These advanced features can help the user identify relations between miRNAs or pathways depending on the effect size of the miRNA–pathway interactions. The web server provides visualizations of the hierarchical clustering in the form of miRNA and pathway dendrograms.

### miRNAs versus pathways heat maps

The new DIANA-miRPath v2.0 server enables also the user to create advanced visualizations such as miRNAs versus pathways heat maps (Figure 3 and Supplementary Figure S1). Heat maps are graphical representations of data where values in a matrix are represented as colors (29). These intuitive visualizations have been proven useful in numerous fields, since they enable the users identify patterns in the data, which were not easily discernible when examining the parameters individually. Furthermore, they enable the visualization of a very



**Figure 3.** miRNAs versus pathways heat map (clustering based on significance levels). Darker colors represent lower significance values. The attached dendrograms on both axes depict hierarchical clustering results for miRNAs and pathways, respectively. On the miRNA axis, we can identify miRNAs clustered together by exhibiting similar pathway targeting patterns. An analogous clustering can be observed also on the pathway axis. In this particular example, we can observe at least one pathway (fatty acid biosynthesis) that is clearly targeted by most investigated miRNAs with a very small *P*-value. More details regarding the methods and results of this example can be found in the Supplementary Material.

large number of variables, their in-between relationships and their levels of interaction. The web server utilizes the hierarchical clustering results on both axes (pathways and miRNAs), in order to construct the heat map visualization. As in the case of cluster analysis, the web server provides two options for heat map calculation: ‘Significance Heat Maps’ and ‘Targeted Pathways Heat Maps’. The former involves the use of absolute *P*-values in all calculations, while the latter substitutes all *P*-values lower than the user defined threshold with 0, and 1 otherwise. With the use of these advanced tools, the user can examine numerous miRNA–miRNA, miRNA–pathway and pathway–pathway relationships. Such representations can help researchers discover patterns and relationships hidden in the data. All plots are rendered in high resolution.

## DISCUSSION

The new web server described in this article is designed to accommodate various use case scenarios and provide results belonging to different research pipelines. For an applied demonstration of the web server’s functionality, we performed a case study based on the work of Finnerty *et al.* (30), where the researchers explore the function and role of the proposed miR-15/107 miRNA

group. It is a group exhibiting strong influences in human biology, highly involved in biological functions, such as cell division, metabolism, stress response and angiogenesis, in vertebrate species. In the performed analysis, the majority of the DIANA-miRPath v2.0 results are supported by the elaborate literature review of Finnerty *et al.* Detailed information regarding the methodology and the derived results from the example can be found in the provided Supplementary Material.

DIANA-miRPath v2.0 is a completely redesigned new version, implemented to provide utilities and functions specific to the analysis and identification of miRNA targeted pathways. All analyses are performed in real time, rendering possible for the user to experiment with different levels of sensitivity and specificity regarding the miRNA–gene interactions. A user researching if a specific pathway of interest is targeted by a group of miRNAs under definite conditions may utilize a more sensitive score (e.g. suggested score 0.7). Similarly, a user investigating in a more generalized experiment which pathways are mostly affected by the expressed miRNAs, may use a higher and more specific threshold (e.g. suggested score 0.9). To our knowledge, this is the first interactive web service for pathway analysis in this direction.

Another unique feature to DIANA-miRPath v2.0 is the integration of experimentally verified targets, offering

validated results derived from the extensive relevant literature.

We think that the new pathological SNP discovery module, which can be used to detect SNPs on miRNA target sites and pathways, may provide in several cases 'hidden jewels' of biological information. We hope that the increased number of additional advanced analysis and visualization tools, such as miRNA versus pathways heat maps, hierarchical miRNA/pathways clustering and meta-analysis merging of results, will help users explore and understand more clearly miRNA-related research data.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2 and Supplementary Figures 1 and 2.

## ACKNOWLEDGEMENTS

The authors would like to thank the two anonymous reviewers for their constructive comments and suggestions.

## FUNDING

The project [09 SYN-13-1055] 'MIKORNA' and the project [09SYN-13-901] 'EDGE' by the Greek General Secretariat for Research and Technology. Funding for open access charge: Project 'EDGE'.

*Conflict of interest statement.* None declared.

## REFERENCES

- Bartel,D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215–233.
- Wang,H. and Peng,D.-Q. (2011) New insights into the mechanism of low high-density lipoprotein cholesterol in obesity. *Lipids Health Dis.*, **10**, 176.
- Dai,R. and Ahmed,S.A. (2011) MicroRNA, a new paradigm for understanding immunoregulation, inflammation, and autoimmune diseases. *Trans. Res.*, **157**, 163–179.
- Huntzinger,E. and Izaurralde,E. (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat. Rev. Genet.*, **12**, 99–110.
- Erson,A.E. and Petty,E.M. (2008) MicroRNAs in development and disease. *Clin. Genet.*, **74**, 296–306.
- Zampetaki,A. and Mayr,M. (2012) MicroRNAs in vascular and metabolic disease. *Circ. Res.*, **110**, 508–522.
- Haasnoot,J. and Berkhout,B. (2011) RNAi and cellular miRNAs in infections by mammalian viruses. *Methods Mol. Biol.*, **721**, 23–41.
- Guay,C., Roggli,E., Nesca,V., Jacovetti,C. and Regazzi,R. (2011) Diabetes mellitus, a microRNA-related disease? *Trans. Res.*, **157**, 253–264.
- Ono,K., Kuwabara,Y. and Han,J. (2011) MicroRNAs and cardiovascular diseases. *FEBS J.*, **278**, 1619–1633.
- Alexiou,P., Maragkakis,M., Papadopoulos,G.L., Reczko,M. and Hatzigeorgiou,A.G. (2009) Lost in translation: an assessment and perspective for computational microRNA target identification. *Bioinformatics*, **25**, 3049–3055.
- Kozomara,A. and Griffiths-Jones,S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.*, **39**, D152–D157.
- Vergoulis,T., Vlachos,I.S., Alexiou,P., Georgakakis,G., Maragkakis,M., Reczko,M., Gerangelos,S., Koziris,N., Dalamagas,T. and Hatzigeorgiou,A.G. (2012) TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. *Nucleic Acids Res.*, **40**, D222–D229.
- Kowarsch,A., Preusse,M., Marr,C. and Theis,F.J. (2011) miTALOS: analyzing the tissue-specific regulation of signaling pathways by human and mouse microRNAs. *RNA*, **17**, 809–819.
- Papadopoulos,G.L., Alexiou,P., Maragkakis,M., Reczko,M. and Hatzigeorgiou,A.G. (2009) DIANA-mirPath: integrating human and mouse microRNAs in pathways. *Bioinformatics*, **25**, 1991–1993.
- Backes,C., Keller,A., Kuentzer,J., Kneissl,B., Comtesse,N., Elnakady,Y.A., Muller,R., Meese,E. and Lenhof,H.P. (2007) GeneTrail: advanced gene set enrichment analysis. *Nucleic Acids Res.*, **35**, W186–W192.
- Hsu,J.B., Chiu,C.M., Hsu,S.D., Huang,W.Y., Chien,C.H., Lee,T.Y. and Huang,H.D. (2011) miRTar: an integrated system for identifying miRNA-target interactions in human. *BMC Bioinformatics*, **12**, 300.
- Backes,C., Meese,E., Lenhof,H.P. and Keller,A. (2010) A dictionary on microRNAs and their putative target pathways. *Nucleic Acids Res.*, **38**, 4476–4486.
- Kanehisa,M., Goto,S., Sato,Y., Furumichi,M. and Tanabe,M. (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.*, **40**, D109–D114.
- Reczko,M., Maragkakis,M., Alexiou,P., Grosse,I. and Hatzigeorgiou,A.G. (2012) Functional microRNA targets in protein coding sequences. *Bioinformatics*, **28**, 771–776.
- Selbach,M., Schwanhaussner,B., Thierfelder,N., Fang,Z., Khanin,R. and Rajewsky,N. (2008) Widespread changes in protein synthesis induced by microRNAs. *Nature*, **455**, 58–63.
- Sayers,E.W., Barrett,T., Benson,D.A., Bolton,E., Bryant,S.H., Canese,K., Chetvernin,V., Church,D.M., DiCuccio,M., Federhen,S. *et al.* (2011) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **39**, D38–D51.
- Beißbarth,T. and Speed,T.P. (2004) GOrat: find statistically over-represented Gene Ontologies within a group of genes. *Bioinformatics*, **20**, 1464–1465.
- Rivals,I., Personnaz,L., Taing,L. and Potier,M.-C. (2007) Enrichment or depletion of a GO category within a class of genes: which test? *Bioinformatics*, **23**, 401–407.
- Hosack,D.A., Dennis,G. Jr, Sherman,B.T., Lane,H.C. and Lempicki,R.A. (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol.*, **4**, R70.
- Benjamini,Y. and Hochberg,Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. Ser. B*, **57**, 289–300.
- Ramon,C.L. and Folks,J.L. (1971) Asymptotic optimality of Fisher's method of combining independent tests. *J. Am. Stat. Assoc.*, **66**, 802–806.
- Lees,J.G., Hriche,J.K., Morilla,I., Ranea,J.A. and Orengo,C.A. (2011) Systematic computational prediction of protein interaction networks. *Phys. Biol.*, **8**, 035008.
- Liu,W., Li,D., Liu,J., Chen,S., Qu,L., Zheng,J., Xu,G. and Yang,N. (2011) A genome-wide SNP scan reveals novel loci for egg production and quality traits in white Leghorn and brown-egg Dwarf layers. *PLoS One*, **6**, e28600.
- Perez-Llamas,C. and Lopez-Bigas,N. (2011) Gitools: analysis and visualisation of genomic data using interactive heat-maps. *PLoS One*, **6**, e19541.
- Finnerty,J.R., Wang,W.X., Hebert,S.S., Wilfred,B.R., Mao,G. and Nelson,P.T. (2010) The miR-15/107 group of microRNA genes: evolutionary biology, cellular functions, and roles in human diseases. *J. Mol. Biol.*, **402**, 491–509.