

What do all the (human) micro-RNAs do?

Ultsch and Löttsch

RESEARCH ARTICLE

Open Access

What do all the (human) micro-RNAs do?

Alfred Ultsch¹ and Jörn Lötsch^{2,3*}

Abstract

Background: Micro-RNAs (miRNA) are attributed to the systems biological role of a regulatory mechanism of the expression of protein coding genes. Research has identified miRNAs dysregulations in several but distinct pathophysiological processes, which hints at distinct systems-biology functions of miRNAs. The present analysis approached the role of miRNAs from a genomics perspective and assessed the biological roles of 2954 genes and 788 human miRNAs, which can be considered to interact, based on empirical evidence and computational predictions of miRNA versus gene interactions.

Results: From a genomics perspective, the biological processes in which the genes that are influenced by miRNAs are involved comprise of six major topics comprising biological regulation, cellular metabolism, information processing, development, gene expression and tissue homeostasis. The usage of this knowledge as a guidance for further research is sketched for two genetically defined functional areas: cell death and gene expression. Results suggest that the latter points to a fundamental role of miRNAs consisting of hyper-regulation of gene expression, i.e., the control of the expression of such genes which control specifically the expression of genes.

Conclusions: Laboratory research identified contributions of miRNA regulation to several distinct biological processes. The present analysis transferred this knowledge to a systems-biology level. A comprehensible and precise description of the biological processes in which the genes that are influenced by miRNAs are notably involved could be made. This knowledge can be employed to guide future research concerning the biological role of miRNA (dys-) regulations. The analysis also suggests that miRNAs especially control the expression of genes that control the expression of genes.

Keywords: Micro-RNA, Gene expression, Regulation, Computational biology, Machine-learning, Knowledge-discovery, Genetics

Background

Micro ribonucleic acids (miRNAs), first described in 1993 [1], have been recognized as a major player in cellular regulation by conferring RNA interference [2]. MiRNAs are initially transcribed from host genes as longer primary transcripts or pri-miRNAs, from which shorter approximately 70 nucleotide-long pre-miRNAs are excised by the RNase III enzyme “Drosha” [3], pri-miRNA transcripts may code for more than one miRNA [4]. Pre-miRNAs are exported from the nucleus to the cytoplasm by the RNA-binding protein exportin 5 [5]. There, they are cleaved to the ~22 nucleotides-long mature

miRNAs by the endoribonuclease “Dicer” [6]. Mature miRNAs impede gene translation by binding at complementary messenger RNA sequences, thereby initiating mRNA cleavage or obstructing mRNA incorporation in ribosomes.

More than 2000 human miRNAs have been identified [7,8], potentially regulating the transcription of the 21,000 human protein-encoding genes [9]. Research during the last decade [10,11] identified miRNAs dysregulations in several pathophysiological processes [12] such as cancer [13], cardiovascular diseases [14], viral infections [15] and pain [16]. In these and further context, miRNAs have been repeatedly found to modulate a wide range of physiological functions such as cellular differentiation, proliferation and apoptosis [17]. This suggests that miRNA-mediated control targets a range of typical biological processes hinting at a distinct systems-biology function of miRNAs.

* Correspondence: j.loetsch@em.uni-frankfurt.de

²Institute of Clinical Pharmacology, Goethe - University, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany

³Fraunhofer Institute of Molecular Biology and Applied Ecology - Project Group Translational Medicine and Pharmacology (IME-TMP), Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany

Full list of author information is available at the end of the article

The present analysis approached the role of human miRNAs from a genomics perspective and assessed the biological roles of those genes that can be considered to interact with miRNAs, based on empirically evidence [18,19] or computational prediction [20]. Computational methods, publicly available databases and data mining tools (Table 1) were used to combine the knowledge about miRNA versus gene interactions with the acquired knowledge about higher-level organization of gene products into biological pathways [21], of which the gold-standard is the Gene Ontology (GO) knowledge base [22].

Methods

Empirical validated miRNA/gene interactions

The genes likely to be regulated by miRNAs were identified by connecting several lines of evidence using publicly available computational methods, databases and data mining tools (Table 1). A first source of miRNA regulated genes consisted of empirically shown interactions of miRNA with genes. The majority of genes with empirical evidence for interaction with a miRNA was identified from miRTarBase database [18] that hosts the currently largest amount of experimentally validated miRNA versus target interactions. From this database the miRNA versus gene interactions were used for which strong experimental evidence was indicated, which in this database was defined as being provided in the form of reporter assays or western blots (file: miRTarBase_SE_WR.xls, Release 4.5 from <http://mirtarbase.mbc.nctu.edu.tw/php/download.php>). This gave a set of $n = 360$ different miRNAs acting on $n = 1472$ different genes. Additional miRNA regulated genes were queried from the TarBase database [19] that hosts further experimentally validated miRNA-gene interactions. In that database, experimentally validated, or supported, interactions are derived from specific, as well as high throughput experiments, such as microarrays and proteomics (for full details, see <http://diana.cslab.ece.ntua.gr/?sec=home>). From this database the reported direct

interactions were used. This gave a set of $n = 136$ different miRNAs acting on $n = 798$ different genes. The size of unions and intersections of these gene sets are given in Figure 1.

Computational prediction of miRNA/gene interactions

To reduce the impact of a possible research bias on the results, a second source of miRNA regulated genes was added from a computational prediction of miRNA regulated genes. A sufficiently credible prediction of miRNA regulated genes was obtained by querying the TargetScan Human software (version 6.2 [20]) for all human miRNAs known to this database. To obtain valid predictions an intensive correction against false positive predictions was performed. Considering the complexity of computational identification of miRNA targets [4,27], a subsequent analysis of the distribution of the output of TargetScan, the so called "Total Context⁺ scores" (TCP scores) [28,29] was performed. To minimize the risk of false positive predictions, this distribution was compared with the scores for empirical validated miRNA targets and only those interactions were kept for which a probability of more than 98% for a valid interaction could be derived (Additional file 1). This filtering reduced the $n = 14610$ unique genes and $n = 1539$ human miRNA for which TargetScan predicted a miRNAa interaction to only $n = 1355$ genes and $n = 548$ human miRNA for which the computer prediction is sufficiently reliable. The union of the miRNA from empirical evidence and filtered computational predictions resulted in $n = 788$ different human miRNAs with interactions on $n = 2954$ different genes (Additional file 2: Table S1).

Biological roles of miRNA regulated genes

To assess the role of miRNA regulation, the biological roles of the genes were identified based on the Gene Ontology (GO) knowledgebase [22] where the knowledge about genes is formulated using a controlled vocabulary of

Table 1 Publicly available data sources and freeware computational tools used to identify miRNA- targeted genes and to classify and visualize their biological functions (accessed November 22, 2013)

Site name	URL	Reference
AmiGO (search utility for GO)	http://amigo.geneontology.org/	[23]
Gene Ontology (GO)	http://www.geneontology.org/	[22]
Gene Trail	http://genetrail.bioinf.uni-sb.de/	[24]
HUGO Gene Nomenclature Committee	http://www.genenames.org/	[25]
miRTarBase	http://mirtarbase.mbc.nctu.edu.tw/	[18]
NCBI gene index database	http://www.ncbi.nlm.nih.gov/gene/	
PubMed	http://www.ncbi.nlm.nih.gov/pubmed	
R software (version 3.0.2)	http://CRAN.R-project.org/	
TarBase database	http://diana.cslab.ece.ntua.gr/tarbase/	[19]
TargetScan Human	http://www.targetscan.org/	[7]

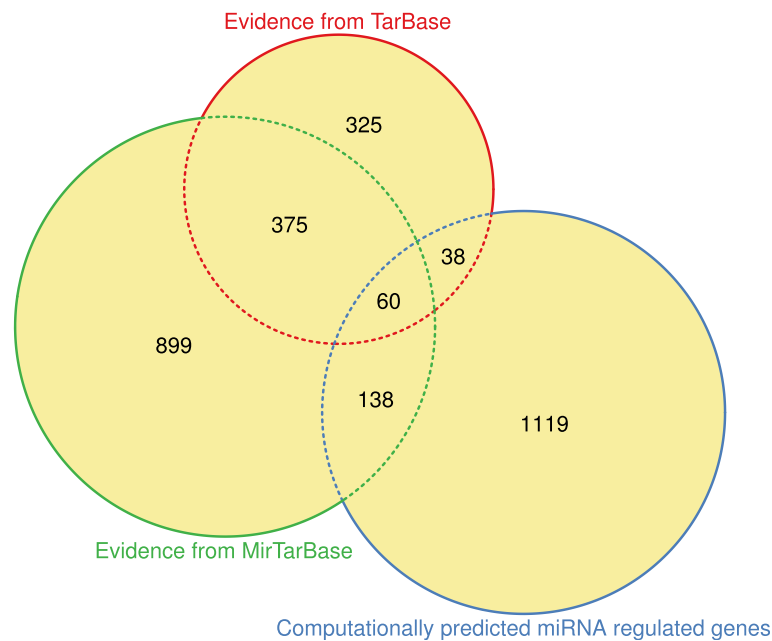


Figure 1 Venn diagram [26] visualizing the sets of genes and the sizes of their intersections. The present analysis was based on the miRNAs that resulted as the union of the three sources, i.e., evidence-based miRNA interacting genes from the miRTarBase database [18] evidence-based miRNA interacting genes from the TarBase database [19] and computationally predicted miRNA regulated genes based on an analysis using the TargetScan Human [20] software (for details of the prediction method, see appendix).

GO terms (categories), to which the genes [30] are annotated [31]. GO terms are related to each other by “is-a”, “part-of” and “regulates” relationships forming a polyhierarchy (i.e., a directed acyclic graph (DAG [32], knowledge representation graph). Particular biological processes, cellular localizations or molecular functions annotated to the miRNA-regulated genes were found by means of an overrepresentation analysis (ORA [33]) using the web-based GeneTrail [24] tool. This tool calculated the significance of the occurrence of the genes of the set of miRNA regulated genes at each term of the GO with respect to the expected occurrence of the genes given by all GO annotations. Statistical significance (p-values) was calculated by the GeneTrail program by applying Fisher’s exact test with Bonferroni α correction [34]. The result was a representation of the complete knowledge about the biological roles of miRNA-regulated genes (complete DAG). To perform this information more intelligible, functional abstraction [35] was applied identifying a special set of GO terms, i.e., “functional areas”, that represent the knowledge contained in the complete DAG at a maximum of coverage, certainty, information value and conciseness [35]. Finally, for GO terms describing biological processes the functional areas could be subsumed to topics to further enhance the conciseness of the description.

To assess the validity of the GO overrepresentation analysis (ORA) in a ten-fold repeated experiment $n = 3000$ genes were randomly chosen from the set of all $n = 17794$

genes for which GeneTrail contained annotations. For a p-value threshold of $t_p = 0.05$ and Bonferroni α correction none of these gene sets produced any significant go term. It could be observed that a small subset of miRNA interacts with many, i.e. up to 229, genes and on the other hand a large subset of miRNA ($n = 304$ of the $n = 788$ miRNA) interacts only with one gene. To address a potential bias of this unequal distribution the set of $n = 788$ miRNA was split into two separate subsets A and B. Set A contains 23% ($n = 181$) miRNA which interact with 75% of the $n = 2954$ genes. Set B contained the other miRNA that interacted with only a few ($n < 6$) genes. Set A produced the same set of functional areas as the set of all $n = 2954$ genes with a median p-value of $1.0 \cdot 10^{-38}$. Set B reproduced the functional areas of the set of all genes (median p-values of $1.0 \cdot 10^{-13}$) with the exception of “biological adhesion” and “response to stimulus” (details given in the supplement).

Results

The analysis of the biological roles by the human miRNA regulated genes could be based on a total of 2954 genes obtained by unifying (Figure 1) the evidence-based sets of miRNA-interacting genes of $n = 1472$ queried from the miRTarBase database [18] and $n = 898$ queried from TarBase database [19]. This set of empirical evidences was augmented by $n = 1355$ genes obtained by computational prediction on the basis of the output of TargetScan [20]

(Additional file 1). With overlaps between the gene sets (Figure 1), the analytical basis comprised of 62% evidence-based and 28% computationally predicted miRNA-regulated genes that based on the same sources of evidence or predictions interact with 344 different miRNAs.

This set of $n = 2984$ human genes empirically shown to be regulated by miRNAs or sufficiently credible computationally predicted to interact with miRNAs was used for an over-representation analysis (ORA [33]) with a p-value threshold of $t_p = 1.0 \cdot 10^{-5}$ and Bonferroni α correction. This resulted in a polyhierarchy of 187 significantly over-represented GO terms in GO categories “biological process”, “cellular component” and “molecular function” (Additional file 3: Figure S1, ORA_Empirical_plus_Predicted.png). By contrast, no under-represented terms were seen.

For the largest GO category, i.e., “biological process” containing 156 significant terms, functional abstraction [35] provided 17 functional areas (Table 2). This described the biological processes in which the genes that are influenced by miRNAs are involved by six major topics comprising biological regulation, cellular metabolism, information processing, development, gene expression and tissue homeostasis. The identified functional areas can be exploited to split the specific knowledge representation graph (DAG; Additional file 3: Figure S1,

ORA_Empirical_plus_Predicted.png) of the 156 terms in the GO category biological process into smaller hierarchies (aspects). This is demonstrated below for the functional areas “cell death” (GO:0008219, Figure 2) and “gene expression” (GO:0010467, Figure 3).

Following functional abstractions of the further GO categories (Table 3) the GO category “cellular component” (Figure 4) indicated 2.5 times more miRNA-interacting genes annotated to the nucleus ($n = 688$ genes) than to the cytoplasm ($n = 274$). This significantly ($p < 10^{-20}$) exceeded the $n = 504$ genes that were expected to be annotated to the nucleus. Finally, the analysis of “molecular function” (Figure 4) indicated a particular role of miRNAs in selective, non-covalent interaction of a molecule with one or more specific sites on another molecule, i.e., “binding” (GO:0005488, $p < 10^{-33}$), including DNA binding (GO:0003677, $p < 10^{-15}$), and the regulation of “transcription factor activity” (GO:0003700, $p < 10^{-10}$) or “transcription factor binding” (GO:0008134, $p < 10^{-23}$).

Discussion

Published literature attributes miRNAs to a systems biological role by a direct regulatory mechanism on classic protein coding genes, mainly by RNA interference impeding gene translation via destabilizing messenger RNA transcripts [36]. Considering that a miRNA may target

Table 2 Functional areas (GO terms of the category “biological process”), topically sorted (left column), of the genes interacting with miRNAs, i.e., for which a gene versus miRNA interaction has been experimentally shown, sorted for the number of genes included

	Functional area	GO term ID	Number of genes			
			Fraction [%]	Observed	Expected	[− log ₁₀ p-val]
BR	Biological regulation	GO:0065007	38.9	1149	817	54.3
Metabolism	Primary metabolic process	GO:0044238	34.2	1009	805	20
	Cellular macromolecule metabolic process	GO:0044260	29.2	864	615	34.4
	Nitrogen compound metabolic process	GO:0006807	23.2	685	478	27.6
	Cellular biosynthetic process	GO:0044249	21	620	428	25.5
	Information transmission	Signaling	GO:0023052	22.7	671	489
	Response to stimulus	GO:0050896	17	501	381	10.4
	Cell communication	GO:0007154	10.8	319	207	15.2
Development	Developmental process	GO:0032502	17.3	512	352	20.5
	Cellular component organization	GO:0016043	15.5	459	341	10.9
	Multicellular organismal development	GO:0007275	15.2	448	305	18
	Cell proliferation	GO:0008283	8.9	262	152	19.2
	Cellular component movement	GO:0006928	4.9	142	77	11.8
	Biological adhesion	GO:0022610	4.2	123	77	5.4
	GE	Gene expression	GO:0010467	18.8	554	380
TH	Cell death	GO:0008219	8.2	243	138	18.8
	Cell cycle	GO:0007049	6.1	180	115	8.1

BR: biological regulation, GE: gene expression, TH: tissue homeostasis.

Significant and remarkable gene ontology (GO) terms (for definition see the AmiGO search tool for GO [23]) resulted from over-representation analysis (ORA) of the 2945 genes with experimentally shown or computationally predicted miRNA interaction that were annotated to the GO category “biological process”.

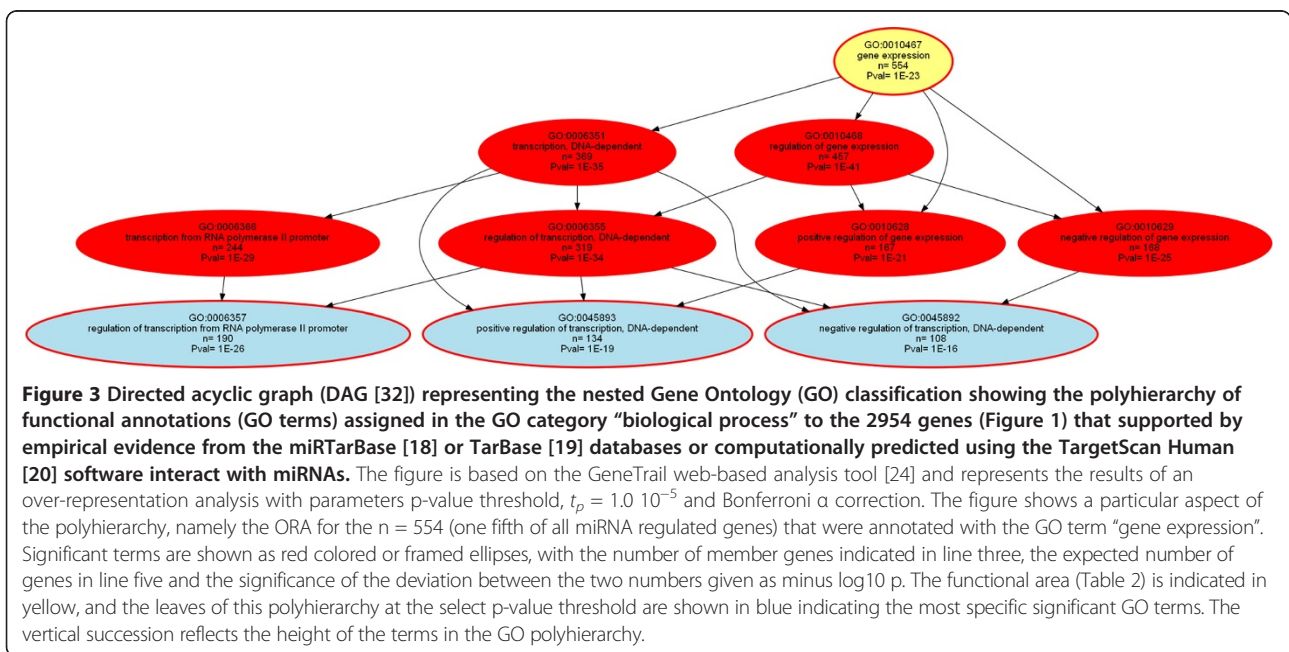
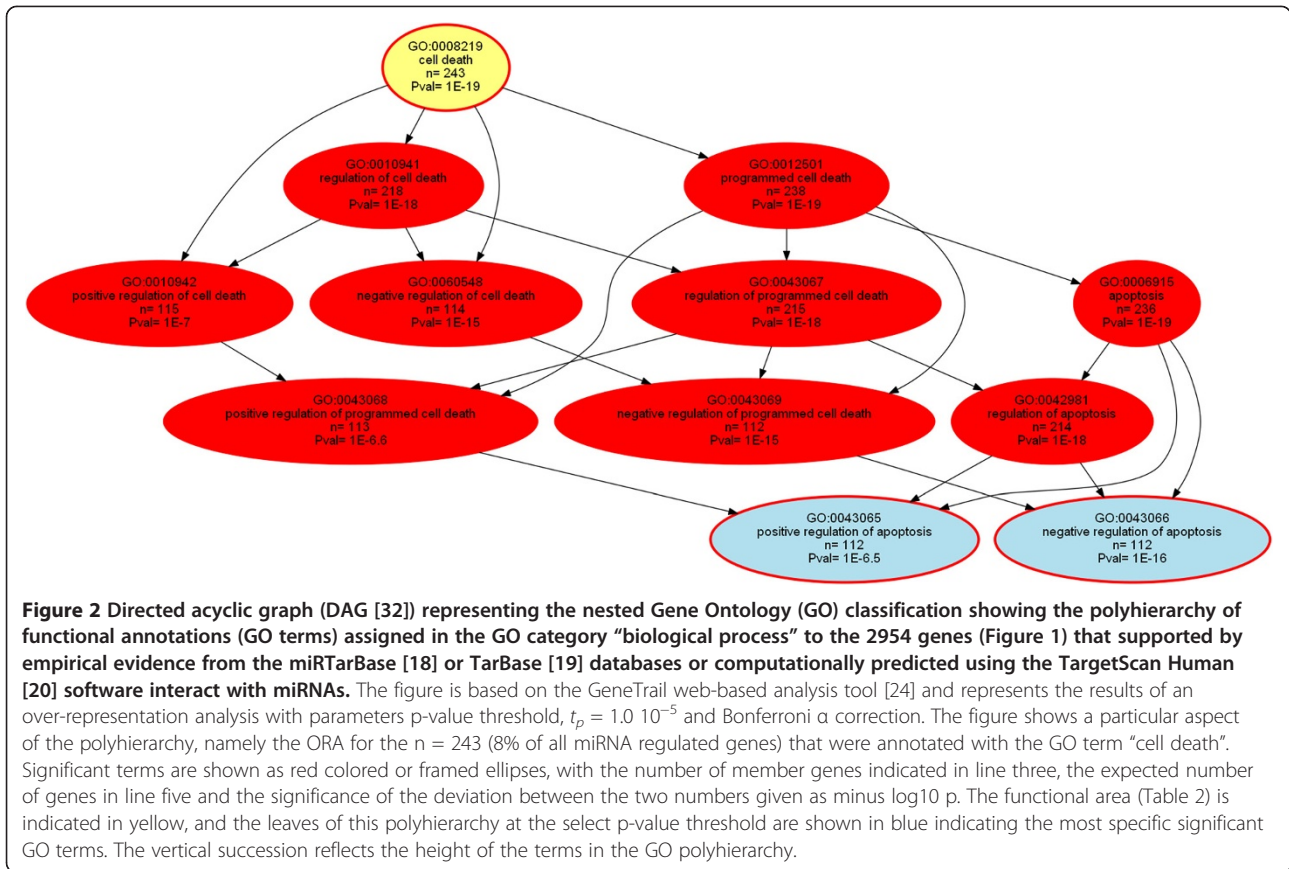


Table 3 Functional areas (GO terms of the categories “cellular component” and “molecular function”) of the genes interacting with miRNAs, i.e., for which a gene versus miRNA interaction has been experimentally shown, sorted for the number of genes included

Functional area	GO term ID	Number of genes			
		Fraction [%]	Observed	Expected	[- log ₁₀ p-val]
Cellular component					
Nucleus	GO:0005634	23.4	688	505	20.8
Cytosol	GO:0005829	9.3	274	205	5.1
Nucleoplasm	GO:0005654	7.4	219	139	10.6
Molecular function					
Binding	GO:0005488	54.7	1612	1351	34.3
Transferase activity, transferring phosphorus-containing groups	GO:0016772	5.9	174	103	10.7
Transcription factor activity	GO:0003700	5.1	151	84	11.6

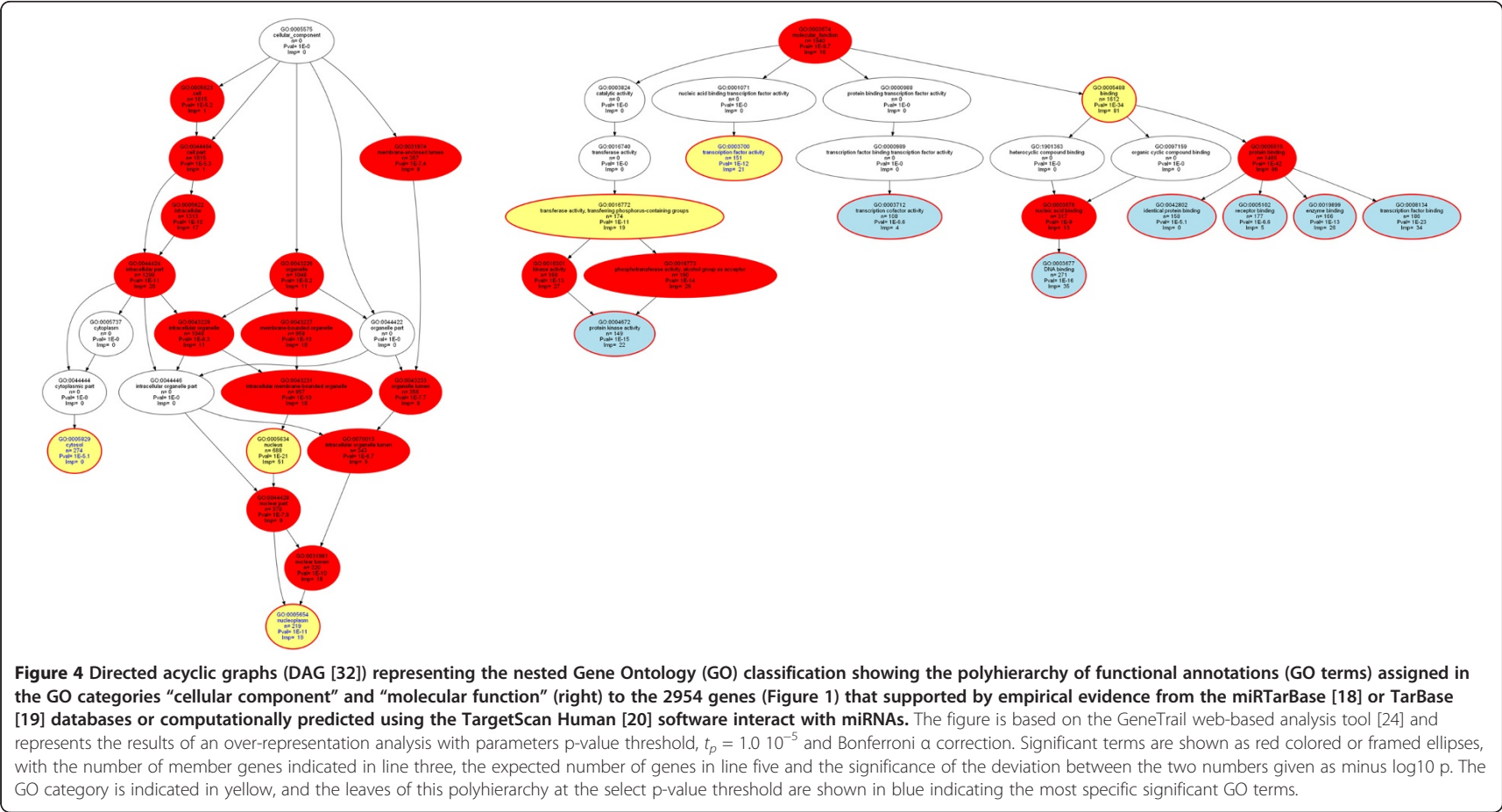
Significant and remarkable gene ontology (GO) terms (for definition see the AmiGO search tool for GO [23]) resulted from over-representation analysis (ORA) of the 2945 genes with experimentally shown or computationally predicted miRNA interaction.

many different genes and vice versa, a gene may be targeted by several different miRNAs [37,38], the ~2000 miRNAs identified for *Homo sapiens* [7,8] may potentially regulate the transcription of all 21,000 human protein-encoding genes [9] and thus be involved in any biological process known to the GO database. However, the present analysis suggested that only a seventh of the human genes seem to be miRNA regulated. Moreover, while the analysis also suggested that this regulation might be involved in any biological process, which is supported by the absence of under-represented GO terms in the ORA, the observed significant over-representation of GO terms clearly indicates that miRNAs play distinct biological roles, which exceed a general evenly-distributed function in gene regulation.

In the present work, a precise and comprehensive view of the systems biological role of miRNAs was obtained via analyzing the functions of a set of genes supported by published evidence for direct miRNA interaction [18,19] combined with a trustworthy computational prediction of miRNA interactions. Using the knowledge about the biological processes, cellular localizations and molecular functions related to genes in the Gene Ontology (GO) knowledge base, the analysis provided a complete and precise description of the involvement of miRNAs in particular physiological and pathophysiological processes. The identification of these distinct roles, represented by functional areas (Tables 2 and 3), was a major finding of this analysis. These functional areas can be considered as a primary answer to the question “What do all those miRNAs do?” from a genomics point of view. Moreover, a further finding of this analysis was, that miRNAs, while exported from the nucleus as pre-miRNAs and in the cytosol processes to mature miRNAs where they exert their RNA interfering function, importantly regulate genes with products acting in the nucleus.

Several of the identified functional areas agree with the current knowledge about the involvement of miRNA in physiological and pathophysiological processes. Specifically, the topic “development” covers the roles of miRNAs reported in neuronal, muscle, and germline development, embryonic stem cell development and differentiation and immune development [17] in developmental regulatory pathways [39], neuronal specification and differentiation [40] or B cell development [41]. The topic information transmission containing the functional areas “signaling”, “response to stimulus” and “cell communication”, reflects the roles of miRNA regulation in immune response modulation and responses to immune-cell stimulation [17], response to stimuli [39,42], autoimmune and inflammatory responses including the toll-like receptor pathway [40] and T cell receptor signaling [41].

The application of the comprehensive overview on the role of miRNAs in organisms may be demonstrated at two particular functional areas, namely cell death (GO:0008219) and gene expression (GO:0010467). The appearance of the first may be attributed to bias of present miRNA research, whereas the second leads to a possible new insight to the general purpose of miRNAs in organisms. MiRNAs are known to play a role in “cell death” including programmed cell death and apoptosis [17,39,40,42,43]. However, while the present analysis verifies this function, the suggestion that the regulation of cell death is a particular role of miRNAs, outstanding from their general role as a ubiquitous regulatory mechanism of gene expression, cannot be maintained when analyzing the evidence-based and computationally predicted sets separately. That is, cell death and related GO terms were only over-represented when analyzing the evidence-based gene set (Additional file 4: Figure S2, ORA_Empirical.png), suggesting a possible research bias since in the set of computationally predicted miRNA-interacting genes no GO term related to



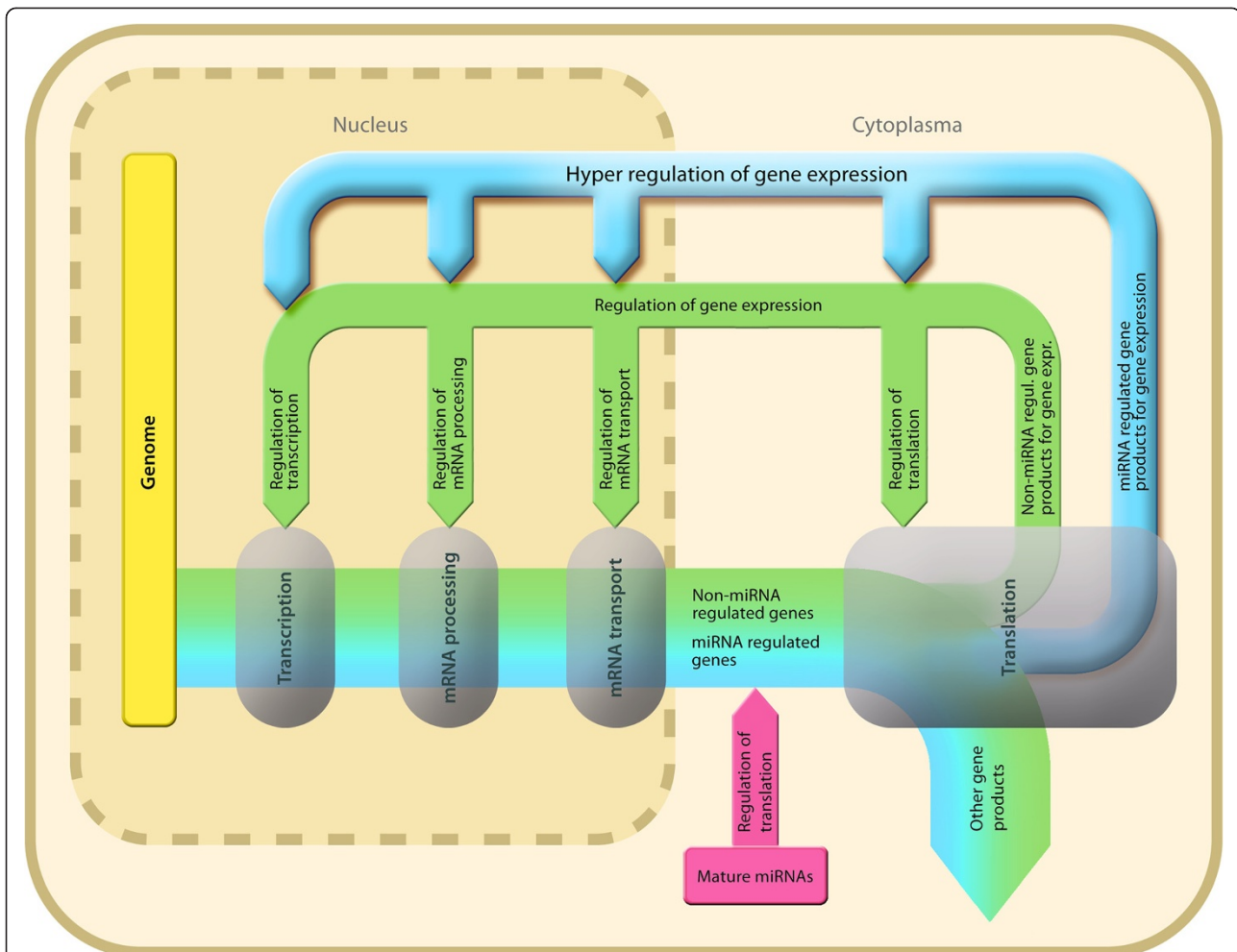


Figure 5 Proposed "hyper-regulation" of gene expression by miRNAs. The figure shows the role of miRNAs in the complex transcriptional network (blue arrow). By regulating (blue arrow) the expression of genes that are involved in the regulation of the expression of genes, a miRNA-dependent regulatory mechanism of gene regulation is formed on top of the miRNA-independent regulation of gene expression (green arrow). By this regulatory mechanism, proposed as "hyper-regulation" of gene expression (blue arrow), miRNAs interfere with the whole transcriptome mainly including intranuclear mechanisms besides the well-known extranuclear (red arrow) mechanisms. Hyper-regulation accommodates observations of global gene down-regulation in the absence of miRNAs [47] which downregulate gene product that reduce gene transcription such as DNA methyltransferases (Additional file 2: Table S1. RegulatedGenes_vs_miRNAs_Matrix.xlsx).

apoptosis emerged as significant (Additional file 5: Figure S3, ORA_Predicted.png).

By contrast, the functional area "gene expression" seems to be a particularly important biological role of miRNA regulation. This GO term was significant in both the evidence-based and computationally predicted gene sets (Additional file 4: Figures S2, ORA_Empirical.png, and 3, ORA_Predicted.png). About one fifth of the miRNA influenced genes ($n = 554$) are involved in the regulation of gene expression as reflected by the analysis of the GO category "biological process". Moreover, the ORAs for the GO categories "cellular localization" and "molecular function" also seemed to converge to gene expression. That is, the molecular functions included an over-representation of transcriptional functions such as transcription factor

regulation and DNA binding. The cellular components where the products of the miRNA regulated genes are located, were more often than expected found in the nucleus. When considering the definition of the GO term "gene expression" (GO:0010467) as the biological processes in which a gene's genomic sequence is converted into a mature gene product or products (proteins or RNA) from the production of an RNA transcript, the processing toward a mature RNA and the translation into proteins [23], miRNA-regulation covers it completely.

Thus, miRNA control applies in particular to the expression of genes that control the expression of genes, which we propose as "hyper-regulation" (Figure 5). The accepted role of miRNAs is the steering (inhibition) of the abundance of gene products, which is mechanistically

exerting its functional infraction mainly in the cytoplasm. Hyper-regulation adds to mechanisms of gene expression control. It points at so far unappreciated increased complexity of gene expression control exceeding current paradigms. It can be hypothesized that miRNA mediated control represents an ancient major mechanism of cellular control providing small versatile molecules at comparably less metabolic effort for respective synthesis compared to protein translation. These systems are being found at all levels of gene expression from transcriptional fine-tuning. This was shown for the transcription activator Ets-1 where variable phosphorylation serves to fine-tune transcription at the level of DNA binding [44], the increasingly populated system of non-protein-coding regulatory RNAs increasing the diversity of control of genome dynamics and developmental programming [45], and the tight control of p53 as “guardian of the genome” shown to be closely regulated by miR-34a [46]. When considering that regulatory mechanisms may also repress genes that repress gene expression, such as all three DNA methyltransferases (DNMT 1, 3a and 3b; Additional file 2: Table S1. RegulatedGenes_vs_miRNAs_Matrix.xlsx), present findings also accommodate observations of genes being down-regulated following the deletion of dicer and thus abolishing the presence of miRNAs [47].

Based on the broad basis of current knowledge, the present data mining and computer science-based approach extends laboratory approaches to the role of miRNAs human biology. However, the analyses relied on external information and therefore, crucially depended on the accuracy and completeness of the empirical evidence entered into the queried databases. Limitation of possible research or publication bias was attempted by adding computational predicted miRNA/gene interactions (TargetScan), which were conservatively filtered to reduce false positives. The consequences of this have been discussed above, revealing that parts of the results cannot exclude a research bias whereas other parts such as the hyperregulation of gene expression prevail regardless of the source of miRNA versus gene interactions. While the intention to exclude false positives nevertheless required conservative statistics throughout all analyses, the procedure might have triggered underestimations of the number of miRNAs versus gene interactions which could affect the results.

Finally, the present computational approach to the role of miRNAs emphasizes the increasing use of bioinformatics in the interpretation of miRNA functions. This accommodates the vast complexity of the acquired information about the role of miRNAs in biology and pathophysiology that probably exceeds human comprehension. Therefore, advances in research increasingly require computer science. This has been shown, for example, in two recent reports where current knowledge from databases was included in generating the research results via

computational means. Specifically, the biological role of miRNAs found by array analyses in regenerating lungs was approached using integrative systems biology assessments including a GO analysis [48]. Interestingly, although this research was aimed towards the role of miRNAs in lung injury and tissue regeneration, one of the results was, that the GO term “gene expression” appeared as an important functional area of those genes that are influenced by the miRNAs particular identified in that experiments (see Figure six in [48]). Thus, the result that miRNAs seem to preferentially regulate genes that regulate the expression of genes obtained presently seems to appear in other analysis on a completely independent data basis as well, supporting its generality and improbability to merely present a bias in the presently queried evidence based miRNA versus gene interactions, which is further supported by the above-mentioned persistence of this results in the computationally predicted miRNAs regulated genes. A further recent example of the utility of computational biology is successful prediction of survival of glioblastoma patients by analyzing the inter-relation between miRNA and gene expression [49].

Conclusions

Laboratory research identified contributions of miRNA regulation to several distinct biological processes. The present analysis transferred this knowledge to a systems-biology level. A comprehensible and precise description of the biological processes in which the genes influenced by miRNAs are notably involved was obtained. This identified seven different topics subsuming 17 functional areas for the genetic role of miRNA regulations: biological regulation, cellular metabolism, information processing, development, gene expression and tissue homeostasis. The present analysis explicitly intended to exploit all the current knowledge about miRNAs versus gene interactions and about the function of genes. This includes the knowledge gathered in databases and the computational means to make predictions. Indeed, the use of knowledge from different sources, when analyzed separately such as for the regulation of genes that regulate the expression of genes, agreed between empirical and predicted interactions, however, bears the potential of disagreements which need to be addressed in the laboratory. Therefore, the knowledge that has emerged from the present analysis can be employed to guide future research concerning the biological role of miRNA (dys-) regulations.

Additional files

Additional file 1: An appendix with the detailed description of computational prediction of miRNA versus gene interactions.

Additional file 2: Table S1. A cross-table displaying the miRNA versus gene interactions (RegulatedGenes_vs_miRNAs_Matrix.xlsx).

Additional file 3: Figure S1. Displaying the results of overrepresentation analysis of the set of genes with experimentally shown or computationally predicted miRNA interaction: ORA_Empirical_plus_Predicted.png.

Additional file 4: Figure S2. Displaying the results of overrepresentation analysis of the set of genes with experimentally shown miRNA interaction: ORA_Empirical.png.

Additional file 5: Figure S3. Displaying the results of overrepresentation analysis of the set of genes with computationally predicted miRNA interaction: ORA_Predicted.png.

Competing interests

The authors declared that they have no competing interests.

Authors' contributions

Conceived and designed the analysis: AU, JL. Analyzed the data: AU. Wrote the paper: JL, AU. Both authors read and approved the final manuscript.

Acknowledgement

This work was supported by the "Landesoffensive zur Entwicklung wissenschaftlich-ökonomischer Exzellenz": "LOEWE-Schwerpunkt: Anwendungsorientierte Arzneimittelforschung" (JL) and the Else Kröner-Fresenius Foundation (EKFS), Research Training Group Translational Research In-novation – Pharma (TRIP, JL). The research has also received funding from the European Union Seventh Framework Programme (FP7/2007 - 2013) under grant agreement no 602919 (JL). The funders had no role in method design, data selection and analysis, decision to publish, or preparation of the manuscript. We thank Ms. Sabine Herda for her help with database queries, Ms. Catharina Lippmann for her work on the graphical representation of ORA results, and Wolf von Waldow for his artistic help with Figure 5. We also thank Prof. Michael Parnham for manuscript language editing. A preliminary version of this analysis, with a limited database, has been published as an abstract of the 46th Workshop of Statistical Computing, July 2104, Ulm, Germany (ISSN 0939–5091, p.16).

Author details

¹DataBionics Research Group, University of Marburg, Hans-Meerwein-Straße, 35032 Marburg, Germany. ²Institute of Clinical Pharmacology, Goethe - University, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany. ³Fraunhofer Institute of Molecular Biology and Applied Ecology - Project Group Translational Medicine and Pharmacology (IME-TMP), Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany.

Received: 21 May 2014 Accepted: 13 October 2014

Published: 18 November 2014

References

- Lee RC, Feinbaum RL, Ambros V: The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 1993, **75**(5):843–854.
- He L, Hannon GJ: MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 2004, **5**(7):522–531.
- Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN: The nuclear RNase III Drosha initiates microRNA processing. *Nature* 2003, **425**(6956):415–419.
- Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG: Lost in translation: an assessment and perspective for computational microRNA target identification. *Bioinform* 2009, **25**(23):3049–3055.
- Bohnsack MT, Czaplinski K, Gorlich D: Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* 2004, **10**(2):185–191.
- Bernstein E, Caudy AA, Hammond SM, Hannon GJ: Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 2001, **409**(6818):363–366.
- Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ: miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 2006, **34**(Database issue):D140–D144.
- Chuang JC, Jones PA: Epigenetics and microRNAs. *Pediatr Res* 2007, **61**(5 Pt 2):24R–29R.
- Pennisi E: Genomics. ENCODE project writes eulogy for junk DNA. *Science* 2012, **337**(6099):1159–1161.
- Ruvkun G: Molecular biology. Glimpses of a tiny RNA world. *Science* 2001, **294**(5543):797–799.
- Lee R, Feinbaum R, Ambros V: A short history of a short RN. *Cell* 2004, **116**(2 Suppl):S89–S92. 81 p following S96.
- Esteller M: Non-coding RNAs in human disease. *Nat Rev Genet* 2011, **12**(12):861–874.
- Croce CM, Calin GA: miRNAs, cancer, and stem cell division. *Cell* 2005, **122**(1):6–7.
- Small EM, Olson EN: Pervasive roles of microRNAs in cardiovascular biology. *Nature* 2011, **469**(7330):336–342.
- Kulkarni S, Savan R, Qi Y, Gao X, Yuki Y, Bass SE, Martin MP, Hunt P, Deeks SG, Telenti A, Pereyra F, Goldstein D, Wolinsky S, Walker B, Young HA, Carrington M: Differential microRNA regulation of HLA-C expression and its association with HIV control. *Nature* 2011, **472**(7344):495–498.
- Niederberger E, Kynast K, Lötsch J, Geisslinger G: MicroRNAs as new players in the pain game. *Pain* 2011, **152**(7):1455–1458.
- Dong H, Lei J, Ding L, Wen Y, Ju H, Zhang X: MicroRNA: function, detection, and bioanalysis. *Chem Rev* 2013, **113**(8):6207–6233.
- Hsu SD, Tseng YT, Shrestha S, Lin YL, Khaleel A, Chou CH, Chu CF, Huang HY, Lin CM, Ho SY, Jian TY, Lin FM, Chang TH, Weng SL, Liao KW, Liao IE, Liu CC, Huang HD: miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res* 2014, **42**(Database issue):D78–D85.
- Papadopoulos GL, Reczko M, Simossis VA, Sethupathy P, Hatzigeorgiou AG: The database of experimentally supported targets: a functional update of TarBase. *Nucleic Acids Res* 2009, **37**(Database issue):D155–D158.
- Lewis BP, Burge CB, Bartel DP: Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 2005, **120**(1):15–20.
- Hu P, Bader G, Wigle DA, Emili A: Computational prediction of cancer-gene function. *Nat Rev Cancer* 2007, **7**(1):23–34.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000, **25**(1):25–29.
- Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, Ami GOH, Web Presence Working G: AmiGO: online access to ontology and annotation data. *Bioinformatics* 2009, **25**(2):288–289.
- Keller A, Backes C, Al-Awadhi M, Gerasch A, Kuntzer J, Kohlbacher O, Kaufmann M, Lenhof HP: GeneTrailExpress: a web-based pipeline for the statistical evaluation of microarray experiments. *BMC Bioinform* 2008, **9**:552.
- Seal RL, Gordon SM, Lush MJ, Wright MW, Bruford EA: genenames.org: the HGNC resources in 2011. *Nucleic Acids Res* 2011, **39**(Database issue):D514–D519.
- Venn J: On the diagrammatic and mechanical representation of propositions and reasonings. *Dublin Philos Mag J Sci* 1880, **9**:1–18.
- Zheng H, Fu R, Wang J-T, Liu Q, Chen H, Jiang S-W: Advances in the Techniques for the Prediction of microRNA Targets. *Int J Mol Sci* 2013, **14**(4):8179–8187.
- Grimson A, Farh KK, Johnston WK, Garrett-Engle P, Lim LP, Bartel DP: MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* 2007, **27**(1):91–105.
- Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP: Weak seed-pairing stability and high target-site abundance decrease the proficiency of Isy-6 and other microRNAs. *Nat Struct Mol Biol* 2011, **18**(10):1139–1146.
- Camon E, Magrane M, Barrell D, Lee V, Dimmer E, Maslen J, Binns D, Harte N, Lopez R, Apweiler R: The Gene Ontology Annotation (GOA) Database: sharing knowledge in uniprot with gene ontology. *Nucleic Acids Res* 2004, **32**(Database issue):D262–D266.
- Camon E, Magrane M, Barrell D, Binns D, Fleischmann W, Kersey P, Mulder N, Oinn T, Maslen J, Cox A, Apweiler R: The Gene Ontology Annotation (GOA) project: implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome Res* 2003, **13**(4):662–672.
- Thulasiraman K, Swamy MNS: *Graphs: Theory and Algorithms*. New York; NY: Wiley; 1992 [u.a.].
- Backes C, Keller A, Kuntzer J, Kneissl B, Comtesse N, Elnakady YA, Muller R, Meese E, Lenhof HP: GeneTrail-advanced gene set enrichment analysis. *Nucleic Acids Res* 2007, **35**(Web Server issue):W186–W192.

34. Hochberg Y: **A Sharper bonferroni procedure for multiple tests of significance.** *Biometrika* 1988, **75**(4):800–802.
35. Ultsch A, Lötsch J: **Functional abstraction as a method to discover knowledge in gene ontologies.** *PLoS One* 2014, **9**(2):e90191.
36. Morozova N, Zinovyev A, Nonne N, Pritchard LL, Gorban AN, Harel-Bellan A: **Kinetic signatures of microRNA modes of action.** *RNA* 2012, **18**(9):1635–1655.
37. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N: **Combinatorial microRNA target predictions.** *Nat Genet* 2005, **37**(5):495–500.
38. Friedman RC, Farh KK, Burge CB, Bartel DP: **Most mammalian mRNAs are conserved targets of microRNAs.** *Genome Res* 2009, **19**(1):92–105.
39. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**(2):281–297.
40. Ullah S, John P, Bhatti A: **MicroRNAs with a role in gene regulation and in human diseases.** *Mol Biol Rep* 2013.
41. Pauley KM, Cha S, Chan EK: **MicroRNA in autoimmunity and autoimmune diseases.** *J Autoimmun* 2009, **32**(3–4):189–194.
42. Ambros V: **The functions of animal microRNAs.** *Nature* 2004, **431**(7006):350–355.
43. Alvarez-Garcia I, Miska EA: **MicroRNA functions in animal development and human disease.** *Development* 2005, **132**(21):4653–4662.
44. Pufall MA, Lee GM, Nelson ML, Kang H-S, Velyvis A, Kay LE, McIntosh LP, Graves BJ: **Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region.** *Science* 2005, **309**(5731):142–145.
45. Amaral PP, Dinger ME, Mercer TR, Mattick JS: **The eukaryotic genome as an RNA machine.** *Science* 2008, **319**(5871):1787–1789.
46. He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y, Xue W, Zender L, Magnus J, Ridzon D, Jackson AL, Linsley PS, Chen C, Lowe SW, Cleary MA, Hannon GJ: **A microRNA component of the p53 tumour suppressor network.** *Nature* 2007, **447**(7148):1130–1134.
47. Zhao J, Lee MC, Momin A, Cendan CM, Shepherd ST, Baker MD, Asante C, Bee L, Bethry A, Perkins JR, Nassar MA, Abrahamsen B, Dickenson A, Cobb BS, Merckenschlager M, Wood JN: **Small RNAs control sodium channel expression, nociceptor excitability, and pain thresholds.** *J Neurosci* 2010, **30**(32):10860–10871.
48. Tan KS, Choi H, Jiang X, Yin L, Seet JE, Patzel V, Engelward BP, Chow VT: **Micro-RNAs in regenerating lungs: an integrative systems biology analysis of murine influenza pneumonia.** *BMC Genomics* 2014, **15**:587.
49. Kim D, Shin H, Joung JG, Lee SY, Kim JH: **Intra-relation reconstruction from inter-relation: miRNA to gene expression.** *BMC Syst Biol* 2013, **7**(Suppl 3):S8.

doi:10.1186/1471-2164-15-976

Cite this article as: Ultsch and Lötsch: What do all the (human) micro-RNAs do? *BMC Genomics* 2014 **15**:976.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

