

Review

Computational Methods and Software Tools for Functional Analysis of miRNA Data

Adrian Garcia-Moreno ¹  and Pedro Carmona-Saez ^{1,2,*} 

¹ Bioinformatics Unit, Centre for Genomics and Oncological Research (GENyO)—Pfizer/University of Granada/Andalusian Regional Government, PTS Granada, 18016 Granada, Spain; adrian.garcia@genyo.es

² Department of Statistics, University of Granada, 18071 Granada, Spain

* Correspondence: pcarmona@ugr.es

Received: 31 July 2020; Accepted: 26 August 2020; Published: 28 August 2020



Abstract: miRNAs are important regulators of gene expression that play a key role in many biological processes. High-throughput techniques allow researchers to discover and characterize large sets of miRNAs, and enrichment analysis tools are becoming increasingly important in decoding which miRNAs are implicated in biological processes. Enrichment analysis of miRNA targets is the standard technique for functional analysis, but this approach carries limitations and bias; alternatives are currently being proposed, based on direct and curated annotations. In this review, we describe the two workflows of miRNAs enrichment analysis, based on target gene or miRNA annotations, highlighting statistical tests, software tools, up-to-date databases, and functional annotations resources in the study of metazoan miRNAs.

Keywords: functional analysis; miRNA; ncRNA; databases; enrichment; tools

1. Introduction

Since the discovery of interfering RNAs in 1993 in *Caenorhabditis elegans* [1] miRNAs have been continuously characterized by high throughput experimental techniques. miRNAs are non-coding RNA molecules of ~22 nucleotides that mediate gene silencing by guiding Argonaute (AGO) proteins to target sites in the 3' untranslated region (UTR) of mRNAs. Over the past few decades, more than 2000 miRNAs were discovered in humans [2], and their key roles in many development and biological processes were characterized. miRNAs were also studied in different human diseases and are currently being pursued in clinical diagnostics and as therapeutic targets [3].

There are several resources that store information about miRNAs. miRBase [4] is one of the main databases that contain a complete miRNA catalogue with sequence and functional information covering more than 271 organisms, including 38,589 hairpin precursors and 48,860 mature miRNAs. Other important miRNAs databases are miRCarta [5] and mirGeneDB [6]. MiRCarta contains miRNA and precursor data from miRBase, predicted miRNAs from sequencing data through miRMaster [7] and different publications. mirGeneDB encloses information regarding curated miRNAs across the metazoan phylum. These are invaluable resources, but important considerations must be kept in mind related to false positives. For example, many entries in miRBase were reported to be fragments of other classes of small RNAs including tRNAs, snoRNAs and rRNA. In an effort to discard human false-positive miRNAs collected in these databases, Alles et al. [2] used small RNA sequencing data from almost 30,000 samples from different sources, such as miRMaster, The Cancer Genome Atlas (TCGA) and National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA), to report 2300 true human mature miRNAs, of which 1115 are currently annotated in miRBase. Increasing knowledge about the role of miRNAs in human diseases also led to the development of dedicated resources, such as Human miRNADisease Database (HMDD) [8], which contains experimentally validated miRNAs,

or miRCancer database [9], which incorporates miRNAs mainly associated with different types of cancers. Similar to HMDD, the Mammal NcRNA-Disease Repository (MNDR) [10] gathers further mammalian organisms, offering a broader catalogue of associations between diverse ncRNAs and diseases. The majority of miRNAs databases are dedicated to animals; nevertheless, several are also dedicated to plants, for example, Plant miRNA Encyclopedia (PmiREN) [11] is a complete, up-to-date catalogue of plant miRNAs that encompass other existing plant databases.

Next-generation sequencing (NGS) and microarray technologies were widely used in the past decade to analyze gene mRNAs and miRNAs at the genome level. In the field of transcriptomics, one common experimental set up is to compare different experimental conditions, i.e., disease versus healthy samples, to define a list of dysregulated miRNAs, which could range from dozens to hundreds. A main challenge in this context, and in the characterization of miRNA networks and pathways, is the drawing of conclusions from these miRNAs lists. This is approached by functional analysis, also referred to as enrichment or over-representation analysis, which generally, involves deciding whether miRNAs are significantly enriched in a specific pathway or biological process which may indicate that the process is associated with the observed phenotype. To perform this analysis, functional annotations can be retrieved from databases, such as Gene Ontology (GO) [12] or Kyoto Encyclopedia of Genes and Genomes (KEGG) [13], and statistical tests can be applied to evaluate which terms are over-represented in the list.

Enrichment analysis is widely used for analyzing gene lists, and functional information and annotation databases are usually centered on genes. To implement these types of methods for miRNAs, a common approach is to retrieve their target genes, which are used to infer processes or pathways in which miRNAs are involved. Alternatively, efforts to centralize the information of a direct association between miRNAs and biological processes are being carried out. These types of databases allow researchers to use curated annotations for miRNAs directly for functional analysis, thereby overcoming some known limitations of indirect miRNA-targets annotations. In this context, Huntley et al. [14,15] used Gene Ontology to annotate miRNAs, taking advantage of probably the most well-settled biological ontology in order to directly annotate miRNAs with their functions. They created the initial guidelines, resulting in 500 mature miRNAs from human, mouse, and rat being associated with nearly 4400 GO terms so far, of which over 3000 are linked with human miRNAs.

In this work, we provide a detailed overview of the functional analysis pipeline in miRNAs, including a revision of annotation databases, statistical tests and a comparative analysis of all available software tools. Further information about quantification methods and downstream miRNAs analyses should be regarded in other reviews. For example, smallRNA-seq aligners are compared by Ziemman et al. [16] and four commonly used miRNA-seq analysis tools are comparatively evaluated with a standard toxicogenomics study design by Bisgin et al. [17].

2. Functional Analysis Workflow in miRNAs

The functional analysis starts with a list of miRNAs (e.g., miRNAs differentially expressed between two phenotypes) and a set of annotations obtained from two different pipelines (see Figure 1). The most common approach is based on obtaining all target genes associated with the list of miRNAs, for which functional gene annotations can be associated, then applying a statistical test to determine their statistical significance; however, there are some reported drawbacks regarding this strategy. Blazard et al. [18] described that this approach results in certain functional categories being preferentially targeted by miRNAs, regardless of whether those miRNAs are differentially expressed in a biological state or not. This implies that random sets of miRNAs report significant p-values for certain specific terms. Equally, Godard and van Eyll [19] proved that the results are not specific and lead to the systematic identification of highly related biological processes, reporting that random miRNA lists of the same size as the analyzed signature, result in similar enrichment results and demonstrate bias toward cancer and cell cycle terms. Although these are important drawbacks, these analyses are still very useful to derive biological information and are used routinely in published studies. An alternative

approach is based on using direct and curated annotations of miRNAs obtained through expert-based annotation from the literature. As mentioned in the introduction, methods following this strategy gained interest in the last few years fueled by the work of Huntley et al. which integrated miRBase and Gene Ontology. Additionally, dedicated databases and enrichment tools using direct annotations are also beginning to be considered, namely, HMDD, MNDR and miRCancer; these databases are described in the following sections.

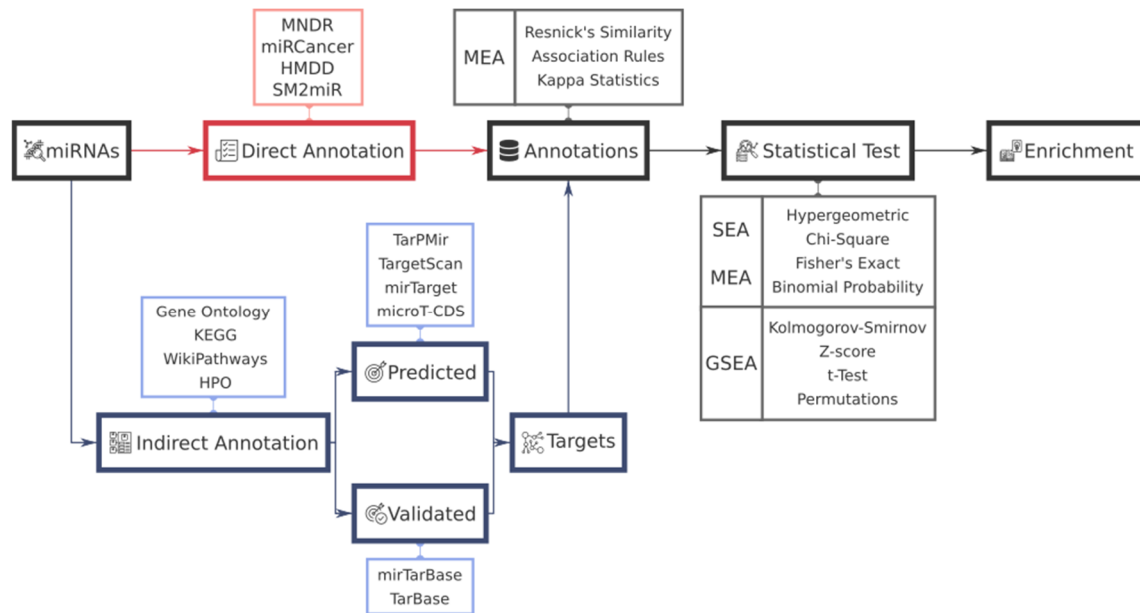


Figure 1. Overview of workflow for functional analysis of miRNAs. Given a list of miRNAs, functional annotations can be retrieved via direct (in red) or indirect (in blue) schemas. Direct annotations are obtained from dedicated databases (i.e., MNDR, miRCancer, HMDD, SM2miR), in which functional terms are directly associated with miRNAs. In the indirect annotations, schema miRNAs are annotated with terms associated with target genes via gene-centered databases (i.e., Gene Ontology, KEGG, WikiPathways, HPO). Then, miRNAs are transformed to their target genes using prediction algorithms (TarPMir, TargetScan, mirTarget, microT-CDS) or experimentally validated targets databases (mirTarBase, TarBase). Functional terms associated with miRNAs, can be grouped by an MEA approach before statistical analysis. Different statistical tests can be applied, SEA and MEA use the same tests to evaluate the enrichment of annotations in the input list with respect to the reference list. Alternatively, threshold-free-based approaches from GSEA tests can be used to analyze the annotations distribution in the entire ranked list. Finally, p-values assigned to each annotation can be used to define over-represented and significant annotations.

Once a set of terms associated with miRNAs exists, a statistical test is applied to determine the most representative annotations in the list. Functional analysis can be classified into three different types [20] (1) Singular Enrichment Analysis (SEA), in which each functional term is evaluated individually using different statistical methods, most commonly the hypergeometric distribution (2) Gene Set Enrichment Analysis (GSEA), which also evaluates independent annotations using a whole set of genes/miRNAs as ranked by certain criteria, e.g., fold change, and computes an enrichment score using Kolmogorov-Smirnov statistics; and (3) Modular Enrichment Analysis (MEA), which takes advantages of the inherent relationship among annotations to define sets of terms that are shared by genes/miRNAs to evaluate their significance.

3. miRNA-Target Gene Annotation and Resources

Target genes for miRNAs are usually discovered by means of sequence-based prediction algorithms or through experimental validation. In this section, we describe some popular prediction tools and

experimentally validated targets databases indicated in Figure 1. We focused on common algorithms used by functional enrichment tools, but more complete reviews on target predictions algorithms are published by Riffo-Campos et al. [21], Peterson et al. [22], and Witkos et al. [23].

3.1. miRNA-Target Gene Prediction Methods

miRNA target prediction is mainly performed based on sequence analysis. These type of algorithms take into account several factors, such as seed pairing and sequence similarity, among miRNA and target mRNAs [24], accessibility of an mRNA [25] AU content [26], GU wobble in the seed match [27], 3' compensatory pairing [28], folding energy [26,29,30], and conservation [31]. Notwithstanding, the binding of a miRNA to its target transcript does not necessarily result in gene expression downregulation. In fact, most observed miRNA binding events, as revealed by crosslinking immunoprecipitation (CLIP) analysis, have little functional consequences [32,33]. Thus, in the context of miRNAs functional analysis, it is important that target prediction algorithms also take into account the effect of target down-regulation by miRNA with RNA-seq data as a method to confirm real functional associations of miRNAs-targets [29]. Table 1 presents a summary of the described prediction tools.

Table 1. Summary of the miRNA target prediction algorithms described.

Tool	Learn Attributes Remark	Organisms	URL	Last Up-Date
TarPmiR	Novel features from CLASH data	<i>Homo sapiens</i>	http://hulab.ucf.edu/research/projects/miRNA/TarPmiR/	2016
TargetScan	Score for mammal predictions	<i>H. sapiens, Mus musculus, Rattus norvegicus, Pan troglodytes, Macaca mulatta, Canis familiaris, M. domestica, Bos taurus, C. elegans, Drosophila melanogaster, Danio rerio, Gallus gallus, Xenopus tropicalis</i>	http://www.targetscan.org	2015
MiRTarget	Functional targets from RNA-seq	<i>H. sapiens, M. musculus, R. norvegicus, C. familiaris, G. gallus</i>	http://mirdb.org/	2019
DIANA microT-CDS	PAR-CLIP data, targets in CDS and 3' UTR	<i>H. sapiens, M. musculus, C. elegans, D. melanogaster</i>	http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=microT_CDS	2013

TarPmiR [34], one of the most complete prediction algorithms, introduced several new miRNA-target binding features by applying four different machine learning methods to CLASH (crosslinking, ligation and sequencing of hybrids) data. Apart from the previously mentioned conventional features, some novel factors introduced in TarPmiR are the consideration of the m/e motif (how different positions in miRNAs match their corresponding target sites positions), length of the target mRNA region, length of the largest consecutive pairs, the difference between the number of paired positions in the seed region and at the miRNA 3' end, number of paired positions at the miRNA 3' end, and position of the largest consecutive pairs relative to the miRNA 5' end and the total number of paired positions.

Another popular algorithm is TargetScan [32] which is mainly based on seed matching and searches for the presence of conserved 8mer, 7mer, and 6mer sites of mRNA that match each miRNA. Sites with mismatches in the seed region but are compensated by conserved 3' pairing, and centered sites are also provided. In the last version of TargetScan, predictions of mammal miRNAs are ranked, based on the predicted efficacy of targeting using cumulative weighted context++ scores of the sites. Context++ is a machine learning model resulting from a study of 26 features, 14 of which were included in the algorithm. These features are: (1) 3'UTR target-site abundance, (2) predicted seed-pairing stability, (3) sRNA position 1, (4) sRNA position 8, (5) site position 8, (6) local AU content, (7) 3'

supplementary pairing, (8) predicted structural accessibility, (9) minimum distance, (10) probability of conserved targeting, (11) ORF length, (12) 3'-UTR length, (13) 3'-UTR offset-6mer sites, and (14) open reading frame (ORF) 8mer sites.

MirTarget [35] is the prediction algorithm that gives rise to MirDB [36], which consists of a support vector machine model trained using public RNA-seq data and miRNA-target databases to identify targeting features characteristic of both miRNA binding and target downregulation. Key aspects in this algorithm include seed conservation, seed match specifically in positions 2–8, base composition in the regions flanking the seed pairing sites, secondary structure, site accessibility, free energy, and location of the site within the 3' UTR. Respectively, MirDB hosts miRNA expression profiles of over 1000 cell lines and presents target prediction data tailored for specific cell models, as well as predictions of miRNA functions by integrative analysis of targets and Gene Ontology data.

DIANA microT-CDS [37] is another algorithm for target prediction, which identifies the most remarkable features extracted from photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) datasets via machine learning techniques. This results in an algorithm with the ability to discover miRNA whose binding location is directly known in both coding sequences (CDS) and 3' UTR based on extended seed matching, distance to the nearest end of CDS or 3' UTR, distance to an adjacent binding site, the free energy of the hybrid, conservation, AU content and 3' UTR accessibility.

The combination of different target prediction methods is a common approach to get more consistent results. For example, Oliveira et al. [38] concluded that the most effective approach was the union, instead of the intersection, of the results from different algorithms to maximize performance, and that several true targets were not identified by these tools alone. Tabas-Madrid, D. et al. [39] proposed two methods to measure the confidence of predicted interactions based on experimentally validated information. These reassigned new scores and statistical confidences for each predicted interaction by nine studied algorithms. In this way, they reduce the selection of interactions to a unique database based on an intuitive score and allow comparing databases between them. Several authors agree with the union of algorithms methods using different approaches [40–43].

3.2. Validated miRNA-Target Gene Resources

miRTarBase [44] and DIANA-TarBase [45], are two main resources that centralize validated miRNA-target information. An overview of them is presented in Table 2.

Table 2. Summary of the reviewed databases with experimentally validated miRNA targets genes.

Tool	Curation	Target-miRNA	Organisms	URL	Last Update
miRTarBase	11,021 articles, 331 CLIP-seq datasets	479,340	32	http://mirtarbase.cuhk.edu.cn	2020
DIANA-TarBase	1208 articles, 353 datasets, 34 methods	665,843	18	https://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=tarbasev8	2017

miRTarBase is one of the largest databases of experimentally validated miRNA-target interactions. The last version, published in 2020, contains almost 480,000 validated interactions extracted from manual curation from a corpus of 11,021 articles. This database not only includes targets, but also regulators of miRNAs to investigate the up- and down-regulation of miRNAs. Targets are classified by the experimental technique used in the validation and whether the evidence is weak or strong. Reporter assays, Western blotting, and qRT-PCR qualify as strong evidence, whereas high throughput techniques, e.g., CLIP-seq, microarray, pulsed stable isotope labeling by amino acids in cell culture (pSILAC), etc., are considered weak evidence. Additionally, disease information from HMDD is incorporated into this database.

Within the DIANA suite, TarBase reached its eighth version in 2017, and includes over 670,000 unique miRNA-target pairs. Nearly 1200 manually curated publications and more than

350 high-throughput datasets support miRNAs-target genes evidence in the database. TarBase divides the targets into low- and high-throughput techniques. Some of the most present methods in the database are reporter assays, Western blotting, qPCR, proteomics, biotin miRNA tagging, sequencing data and microarrays, among others. Additionally, disease information from miR2Disease [46] is included.

4. Functional Annotation Resources for miRNAs

Functional annotations for miRNAs can be directly assigned by literature curation or inferred using information from target genes. In this section, we provide a brief overview of the main biological information sources used in enrichment analysis for both miRNAs and genes.

4.1. miRNA-Based Annotation Resources

Databases that contain functional information of miRNAs compile annotations obtained via manual literature searches or text-mining algorithms plus an expert curation process to check evidence levels. There are many resources focused on associating diseases to miRNAs, while few of them relate them to concrete biological functions or entities, such as transcription factors, drugs, or epigenetic modifiers. We briefly describe some of these databases used in miRNAs enrichment tools.

miRCancer is a reference database which includes human miRNA-cancer associations from empirical evidence; built via text-mining of more than 26,000 PubMed articles, and currently contains 9080 relationships among 57,984 miRNAs and 196 cancers. The algorithm is based on 75 rules, which represent the common sentence structures typically used to state miRNA expressions in cancers. All the annotations are confirmed manually after automatic extraction.

HMDD is focused on establishing human miRNA—disease associations evidence-based on manual curation. Currently, more than 32,281 experimentally supported miRNA—disease links, covering 1102 miRNA genes and 850 diseases from 17,412 papers are contained in the database. All miRNAs included are standardized to match the miRBase nomenclature and diseases are classified and normalized on the basis of Disease Ontology [47] and (Medical Subject Headings) MeSH; associations are categorized in six evidence codes.

In contrast to miRCancer and HMDD, MNDR integrates experimental and predicted ncRNA-disease associations from manual literature curation and 10 other resources for 11 different mammalian organisms. More than one million of ncRNA-disease entries, including 6301 miRNAs, 39,880 lncRNAs, 20,256 circRNAs, 10,894 piRNAs and 521 snoRNAs with over 1600 diseases, are stored in the database. Diseases are mapped to Disease Ontology and MeSH terms. The associations are classified into three evidence types: Strong experimental evidence, weak experimental evidence and prediction algorithm and miRNAs following the miRBase nomenclature.

PhenomiR [48] provides data from 542 studies which investigate the deregulation of miRNA expression in diseases and biological processes as a systematic, manually curated resource. miRNAs are mapped to miRBase, and diseases are annotated according to the Online Mendelian Inheritance in Man (OMIM) [49] Morbid Map. SM2miR [50] is another useful resource which contains information about research on drugs that affect miRNAs expression levels.

4.2. Gene-Based Annotation Resources

Gene Ontology and KEGG are probably the most common databases that store functional annotations used in enrichment analysis. Although, as mentioned previously, Gene Ontology includes miRNAs under their terminology, it is still widely used in indirect miRNA enrichment approaches. Gene Ontology provides a vocabulary for categorizing biological processes, cellular components and molecular functions. Besides the ontology itself, the consortium also provides annotations for several organisms, with evidence-based statements relating a specific gene product to specific ontology terms.

The KEGG database is a very popular and well-established resource that originally focused on metabolic pathways, but currently includes 18 different databases classified into four main categories, namely, systems, genomics, chemicals and health. KEGG is widely known due to its interactive

pathways and network maps. In addition to its relevance for basic research, in the last few years, the database began to move towards biomedicine applications integrating human diseases, drugs and other health-related substances.

Being part of the GO consortium, Panther (protein analysis through evolutionary relationships) [51] was born as a classification system of proteins and their genes into families and subfamilies based on their sequence orthology. This classification, along with a tool suite, allows to perform different functional analyses based on Gene Ontology and their inherent pathways annotations. The pathways are complemented with information from Reactome [52], and their own GOSlim version (a subset of the GO ontology with broader terms) is available. While Panther covers several organisms, Reactome only applies to human data and aims to annotate validated information regarding genes, drugs, small molecules, catalysts and regulators throughout more than 1800 pathways organized in a multilevel hierarchical network that could be collapsed into 26 super-pathways, e.g., hemostasis or muscle contraction. Another source of pathway annotation is WikiPathways [53], which was characterized by a crowdsourcing curation and presents more than 2600 pathways. This classification also implements different ways to train new users and ensure quality terms and proper evidence tracking.

Disease to gene databases are also well-established, for example, Online Mendelian Inheritance in Man (OMIM) or Disease Ontology and Human Phenotype Ontology (HPO) [54] are widely used resources. OMIM focuses on human genes and genetic disorders and traits, highlighting the nature of their variations and the resulting phenotypes. Currently, OMIM has over 24,600 entries with approximately 16,000 genes and 8600 phenotypes. The Disease Ontology includes over 9069 disease terms which are interconnected semantically with other databases, such as OMIM. Likewise, the Human Phenotype Ontology provides the most comprehensive normalized vocabulary in order to carry out deep phenotyping in the rare diseases field. Given the heterogeneity of rare diseases, this ontology was adopted by many organizations, i.e., database of genomic variation and phenotype in humans using ensemble resources (DECIPHER) [55] and Orphanet.

Information from these databases can also be used to annotate miRNAs-centered resources by linking them via target genes. For instance, miRPathDB [56] uses Gene Ontology, KEGG, Reactome and WikiPathways.

5. Tools for miRNA Functional Enrichment Analysis

In this section, we provide a review of popular tools used for miRNA functional enrichment analysis, focused on annotation sources, available organisms, workflow, bias handling and statistical methods. Table 3 contains a summary of the revised tools specifying the type of annotation, functional analysis method, target gene sources, available annotations and supported organisms. In addition, supplementary material (File S1) contains an overview of results from all the reviewed tools using a list of 26 dysregulated miRNAs in serum exosomes from glioblastoma (GBM) patients [57].

Table 3. Summary of the revised miRNA functional enrichment analysis tools.

Tool	Annotation/Bias Handling	Method	Targets	Sources of Annotations	Organism
miRNet	Indirect, Direct/Empirical sampling	SEA	Validated, predicted	GO, KEGG, Reactome, TAM [58]	<i>H. sapiens</i> , <i>M. musculus</i> , <i>R. norvegicus</i> , <i>B. taurus</i> , <i>Sus scrofa</i> , <i>G. gallus</i> , <i>D. melanogaster</i> , <i>C. elegans</i> , <i>D. rerio</i> , <i>Schistosoma mansoni</i>
GeneCodis	Indirect/Empirical sampling, co-annotation	MEA, SEA	Validated	DoRothEA [59], miRTarBase, GO, KEGG Pathways, MGI [60], Panther, Reactome, WikiPathways, CTD [61], HPO, LINCS [62], OMIM, PharmGKB [63]	<i>D. melanogaster</i> , <i>D. rerio</i> , <i>H. sapiens</i> , <i>M. musculus</i> , <i>R. norvegicus</i> .

Table 3. Cont.

Tool	Annotation/Bias Handling	Method	Targets	Sources of Annotations	Organism
miEAA	Indirect, Direct/None	SEA, GSEA	Validated, redicted	GO, HMDD, KEGG, miRandola [64], miRBase, miRPathDB, miRTarBase, miRWalk [65], MNDR, NPInter [66], RNALocate [67], SM2miR, TAM, TissueAtlas [68], TransmiR [69], Literature	<i>H. sapiens</i> , <i>M. musculus</i> , <i>R. norvegicus</i> , <i>Arabidopsis thaliana</i> , <i>B. taurus</i> , <i>C. elegans</i> , <i>D. melanogaster</i> , <i>D. rerio</i> , <i>G. gallus</i> , <i>S. scrofa</i>
MIENTURNET	Indirect/None	SEA	Validated, predicted	KEGG, Reactome, WikiPathways, Disease Ontology	<i>H sapiens</i> , <i>M. musculus</i> , <i>R. norvegicus</i> , <i>C. elegans</i> , <i>D. melanogaster</i> , <i>D. rerio</i>
TAM	Direct/Mask cancer and unspecific terms	SEA	-	Literature	<i>H. sapiens</i>
miTALOS	Indirect/Background specificity	SEA	Validated, predicted	KEGG, WikiPathways, Reactome	<i>H. sapiens</i> , <i>M. musculus</i>
miRSystem	Indirect/Empirical sampling	SEA	Validated, predicted	KEGG, GO, BioCarta [70], Pathway Interaction Database [71], Reactome.	<i>H. sapiens</i> , <i>M. musculus</i>
DIANA miRPath	Indirect/Empirical sampling	SEA	Validated, predicted	GO, KEGG	<i>H. sapiens</i> , <i>M. musculus</i> , <i>R. norvegicus</i> , <i>D. melanogaster</i> , <i>C. elegans</i> , <i>G. gallus</i> , <i>D. rerio</i> .

5.1. MiRNet

MiRNet [72] offers a unique way of exploring the interactome of miRNAs. Its database includes 11 different miRNA resources and allows the use of custom data. The interactome includes miRNAs and their validated and/or predicted targets, plus, elements like xeno-miRs, transcription factors, epigenetic modifiers, SNPs, pseudogenes, ncRNAs, diseases, and small compounds. All of this information is available for 10 different model organisms. To use the application users can introduce any of the mentioned elements alongside a gene expression data table from mRNA or miRNA experiments. The analysis will report pairwise tables and networks of connections among miRNAs or genes based on all the different elements selected.

From the network, a functional SEA can be applied with the hypergeometric test, including direct and indirect annotations depending on the collection selected. If this is based on genes, an empirical sampling of the test p-values is incorporated along with annotations from GO, KEGG, Reactome or Diseases, meanwhile, if miRNAs are chosen no empirical sampling is applied because direct annotations from TAM are available, allowing over-representation in tissues, diseases, miRNA functions, miRNA clusters, miRNA families and miRNA transcription factors. This dual implementation is motivated by bias in the indirect approach of miRNA enrichment. The empirical sampling follows what is described by Bleazar et al. [18] and the inclusion of direct annotations follows a solution proposed by Godard and van Eyll [19]. The empirical p-values could be re-estimated with another 1000 permutations by resubmitting the functional analysis, but a limitation is that they can only be obtained by using the full set of genes in the network and results are only reported if p-values are below 0.001. If the hypergeometric test is applied, no p-value cut-off exists; therefore, all the related terms are shown. MiRNet is available programmatically via R's package, web application programming interface (API) and at <https://www.mirnet.ca>.

5.2. GeneCodis

GeneCodis [73] is a popular functional enrichment tool first presented in 2007 [74] as one of the first applications for modular enrichment analysis. It allows enrichment analysis of single annotations, but its main advantage is the extraction of sets of annotations associated with the same set of genes to evaluate statistical significance. It currently supports 15 annotation sources for biological processes,

pathways, regulatory elements and drugs. In the last update, GeneCodis incorporated a functionality of miRNAs enrichment analysis via indirect annotation for 5 of 15 available organisms. Target genes with strong evidence were retrieved from the last version of miRTarBase. Additionally, from a list of genes, transcription factors or CpG sites, this database checks which miRNAs are significantly represented.

Depending on the organism selected, different annotations can be used. In both types of enrichment analysis, the significance of the co-annotation or individual term is obtained using either the hypergeometric test or the chi-square tests, and p-values are corrected by false discovery rate (FDR) or using a permutations-based approach. GeneCodis allows users to customize a background set of genes or miRNAs; by default, all annotated genes are considered. Two bias corrections are available, namely, with permutations in concordance with the Blazard et al. bias solution, and also, in the MEA methodology, a similar grouping of miRNAs targets in co-occurring annotations was proposed by Godard and van Eyll. This tool also implements different visualization capabilities to explore the results beyond standard graphs and tables, such as term clustering that is based on gene sharing by combining principal component analysis and t-SNE (t-distributed Stochastic Neighbor Embedding). GeneCodis is available at <https://genecodis.genyo.es>.

5.3. MiEAA

The miRNA Enrichment Analysis and Annotation tool (miEAA) [75] implements GSEA and SEA for miRNAs. It consists of a comprehensive database of more than 40 different collections obtained with direct and indirect annotations of miRNAs. The main sources are 15 databases that allow users to explore associations with pathways, diseases, miRNA nature and classification, drugs, functions, cells and tissues, targets, and transcription factors. To avoid bias in miRNA enrichment analysis, a majority of sources providing direct annotations were collected. The statistical method used in the SEA is the Fisher's exact test, whereas, GSEA uses an un-weighted variant of the algorithm which corresponds to a Kolmogorow-Smirnow test. The static GSEA running sum plots shows a simulated background distribution computed by randomly permuting the input list 100 times and traversing the running sum for each random permutation. Furthermore, six different procedures are available for multiple tests correction.

MiEAA uses a list of precursors or mature miRNAs as input, where the user can select from the enrichment methodologies. In the case of SEA, a background set can be uploaded, otherwise, all annotated miRNAs/precursors are used as a reference. To perform GSEA the input must be sorted by some criterion, such as fold change. Once submitted, the results gather all the categories into a single table. For the top 100, sorted by p-value, a word cloud and heatmaps of miRNA versus annotation are created, whereas interactive enrichment graphs are shown for GSEA. MiEAA is available programmatically via API and at: <https://ccb-compute2.cs.uni-saarland.de/mieaa2>.

5.4. MIENTURNET

MIENTURNET (miRNA enrichment turned network) [76] was published as a tool to study miRNA-target interactome. Every miRNA and target gene is extracted from TargetScan and those experimentally validated from miRTarBase. The six organisms shared between these databases are included in the tool. Four sources of annotations available for functional analysis performance, specifically, KEGG, Reactome, WikiPathways, and, for human data only, Disease Ontology. These are linked to genes, indicating that over-representation analysis is performed over the miRNAs targets and designating it as an indirect annotation approach.

As input, a list of genes or miRNAs can be used. In both cases, this will trigger an enrichment analysis of miRNAs or targets, respectively, by querying, TargetScan and miRTarBase simultaneously, which in turn provide two interactomes of miRNA-gene pairs, one is based on predicted interactions and the other on validated interactions. Targets can be filtered by the type of evidence for miRTarBase and two prediction scores, the cumulative weighted context++ and the probability of conserved targeting. Functional analysis can be performed using all the targets of up to 10 miRNAs from the interactomes.

The significance of terms is addressed with the hypergeometric test, and p-values are corrected using Benjamini-Hochberg FDR. As a background, by default, the whole genome is used; otherwise, the input genes are used. MIENTURNET is available at <http://userver.bio.uniroma1.it/apps/mienturnet>.

5.5. TAM

TAM [58] is a dedicated human miRNA functional analysis tool whereby, through manual curation of more than 9000 papers, a database was created to characterize around 1200 distinct miRNAs. These are associated in a total of 1238 miRNA sets distributed throughout 547 diseases, 158 biological functions, 166 transcription factors, 6 tissues and 211 clusters and 151 families of miRNAs. Interestingly, the nomenclatures of disease and functional terms were normalized regarding other well-established ontologies (ICD-10-CM, Disease Ontology, MeSH, OMIM, HPO and Gene Ontology). Another unique feature of this tool is that for almost all miRNA-disease associations, miRNAs are classified into two groups, namely, (1) up-miRNA, i.e., miRNAs that are up-regulated in disease conditions, or exhibit disease-promoting function according to the phenotype from gene permutation assays, and conversely, (2) down-miRNA, curated in the same way.

TAM is an example of miRNA functional analysis using exclusively direct annotations. Taking into consideration the bias regarding the knowledge of miRNAs separate from the direct annotations, this tool provides an option to mask cancer and non-standard miRNA terms when performing analyses. TAM offers two types of enrichment processes, a SEA with the hypergeometric test, and notably, a comparison of query miRNAs signature (up and down-regulated) with the disease-based signatures stored in the database. TAM is available at <http://www.lirmed.com/tam2>.

5.6. MiTALOS

MiTALOS [77] is a dedicated functional analysis tool for miRNA in which Preusse et al. integrated three key aspects. First, from CLIP-seq studies of StarBase [78], they extracted high-quality pairs miRNA-target, but also considered predicted pairs from TargetScan and miRanda [79]. Second, they captured tissue-specific gene expression from the latest version of EBI Expression Atlas. Lastly, they included three major pathway databases: KEGG, WikiPathways and Reactome. As a result, this tool incorporates the specificity of expression signatures of miRNAs and target transcripts in different tissues to improve the functional analysis of miRNAs. All miRNA information is relative to *H. sapiens* and *M. musculus*.

The approach of miTALOS is an indirect annotation, but thanks to a dynamic database tissue filter, the results provide insight and may overcome the described bias. In detail, they establish that the expression of miRNAs and pathways is tissue-specific, therefore miRNAs, genes and pathways can be discarded in the face of a functional analysis if they are not expressed in the selected tissue. Terms significance is obtained via Fisher's exact test and corrected using the Benjamini-Hochberg procedure. MiTALOS is available at <http://mips.helmholtz-muenchen.de/mitalos>.

5.7. MiRSystem

MiRSystem [80] is a dedicated source to perform miRNA functional enrichment analysis. Its miRNAs database is built from miRBase and allows two species, *H. sapiens* and *M. musculus*. Regarding the functional annotations, five databases are collected: KEGG, Gene Ontology, BioCarta, Pathway Interaction Database, and Reactome.

This tool uses the indirect annotation approach and miRNAs are transformed to target genes by combining different prediction algorithms and two experimentally validated databases. A list of targets is obtained after prediction by a minimum number of algorithms. By default, the experimentally validated targets are added. The test incorporated is the hypergeometric test which is complemented by calculating the observed/expected (O/E) ratio (i.e., observed genes divided by expected genes under a term) and an empirical p-value using a permutations test. Similar to other tools, the calculation of an empirical p-value is a method to overcome bias. A weighted enrichment analysis also was

developed, which is a unique method compared to the other tools. If miRNAs are introduced along with a numeric value, e.g., the expression, the enriched pathways are scored according to the expression of the associated miRNAs. The weight for one miRNA is calculated by dividing its absolute expression value by the absolute sum of the expression values of all input miRNAs. Thereafter, the pathway ranking score is obtained by totaling the weight of its miRNA times its enrichment $2\log(p\text{-value})$. MiRSystem is available at <http://mirsystem.cgm.ntu.edu.tw>.

5.8. DIANA miRPath

miRPath [81] is part of the DIANA framework and links miRNAs to Gene Ontology and KEGG. miRNAs are converted to predicted targets by TargetScan or DIANA microT-CDS and/or experimentally validated targets stored in DIANA TarBase. These associations are available for seven model organisms.

Before performing enrichment analysis for each introduced miRNA, its predicted targets can be filtered according to each algorithm's own score system and then collapsed into a single list either by the intersection of a minimum number of associated miRNAs or by the union of all targets. miRPath fits into the indirect annotation category, but it comes with statistical implementations for bias assessment. To obtain p-values, miRPath can use Fisher's exact test and the DAVID's EASE score [82] optionally, p-values can be corrected via Benjamini-Hochberg's FDR. These values undergo permutation testing derived from an adaptation of the sampling algorithm presented by Blaezard et al., resulting in empirical p-value. Finally, a method consisting of a meta-analysis statistic was developed to enable the identification of pathways controlled by multiple miRNAs by examining each one individually and subsequently combining the result probabilities and test statistics. Additionally, functional analysis can be performed in the opposite direction, i.e., in the reverse search module, to discover over-representation of miRNAs from a set of targets of a single GO term or KEGG pathway. DIANA miRPath is available at: <http://snf-515788.vm.oceanos.grnet.gr>.

6. Conclusions and Discussion

There are numerous databases for miRNAs that provide various types of data, including predicted and experimental evidence of miRNA and target gene associations. However, the availability of miRNAs functional information gained importance over the last few years as an essential step to interpret high-throughput experiments and decipher the biological processes in which they are involved. In this context, several miRNA enrichment analysis tools were recently introduced, and are becoming very useful resources in miRNA research. These tools can combine both miRNA target genes plus gene-based annotations databases and/or direct miRNA functional annotations databases.

Initially, miRNAs functional analysis methods were based on SEA using target genes information, but different efforts are being realized to provide high-quality direct annotations for miRNAs. In the indirect approach, it is important to use well-validated target genes, while the use of direct annotations, miRNA-based databases, requires a well-established and normalized vocabulary. In order to decipher specific pathways and functions, it is essential the specificity of target prediction algorithms. Therefore, target validation information or integration of other data sources, such as gene expression, is very convenient in functional analysis. Tools like GeneCodis, use uniquely target genes validated by miRTarBase offering MEA, but some others stand out because they incorporate direct annotations, such as TAM. Differently, miEEA implements GSEA and both types of annotations. These varying implementations, in addition to statistical tests and supported organisms, offer several alternatives to perform miRNA enrichment analysis. It is known that each study requires a different setup, whatsoever, motivated by the bias in the indirect approach, tools that have miRNAs direct annotations should be preferred whenever possible. Nonetheless, it could be argued that direct annotations are still lacking and are not widely implemented in miRNAs functional analysis tools. Thus, if miRNAs direct annotations do not meet the research specificity level and gene-based annotations are necessary, the bias handling method in the indirect workflow must be noticed. Empirical sampling is the most common method to assess the bias, while the clustering of annotations in GeneCodis and the background

specificity of a selected tissue in tools like MiTalos are less implemented. In the case of TAM, although it uses direct annotations, an option to mask cancer and unspecific terms are available when the miRNAs list under study is not related to the pathology.

This review provided an overview of the most widely used resources for miRNA functional analysis, remarking upon the main features in terms of the type of annotation, statistical test, organism or enrichment analysis method. We expect this review to be useful in selecting the most appropriate resource depending on the experimental context.

Supplementary Materials: Available online at <http://www.mdpi.com/2218-273X/10/9/1252/s1>, File S1: Tables S1–S9. Tables S1–S8: Results from each tool for a 26 miRNAs signature in glioblastoma patients. Table S9: Table with tools and URLs not included from mentioned resources.

Funding: This work was partially supported by the Junta de Andalucía through grants PI-0173-2017 and CV20.36723.

Conflicts of Interest: The authors of this review are responsible for GeneCodis. The authors declare no further conflict of interest.

References

1. Lee, R.C.; Feinbaum, R.L.; Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **1993**, *75*, 843–854. [[CrossRef](#)]
2. Alles, J.; Fehlmann, T.; Fischer, U.; Backes, C.; Galata, V.; Minet, M.; Hart, M.; Abu-Halima, M.; Grässer, F.A.; Lenhof, H.-P.; et al. An estimate of the total number of true human miRNAs. *Nucleic Acids Res.* **2019**, *47*, 3353–3364. [[CrossRef](#)] [[PubMed](#)]
3. Rupaimoole, R.; Slack, F.J. MicroRNA therapeutics: Towards a new era for the management of cancer and other diseases. *Nat. Rev. Drug Discov.* **2017**, *16*, 203–222. [[CrossRef](#)]
4. Kozomara, A.; Birgaoanu, M.; Griffiths-Jones, S. miRBase: From microRNA sequences to function. *Nucleic Acids Res.* **2019**, *47*, D155–D162. [[CrossRef](#)]
5. Backes, C.; Fehlmann, T.; Kern, F.; Kehl, T.; Lenhof, H.-P.; Meese, E.; Keller, A. miRCarta: A central repository for collecting miRNA candidates. *Nucleic Acids Res.* **2018**, *46*, D160–D167. [[CrossRef](#)]
6. Fromm, B.; Domanska, D.; Høye, E.; Ovchinnikov, V.; Kang, W.; Aparicio-Puerta, E.; Johansen, M.; Flatmark, K.; Mathelier, A.; Hovig, E.; et al. MirGeneDB 2.0: The metazoan microRNA complement. *Nucleic Acids Res.* **2020**, *48*, D132–D141. [[CrossRef](#)] [[PubMed](#)]
7. Fehlmann, T.; Backes, C.; Kahraman, M.; Haas, J.; Ludwig, N.; Posch, A.E.; Würstle, M.L.; Hübenthal, M.; Franke, A.; Meder, B.; et al. Web-based NGS data analysis using miRMaster: A large-scale meta-analysis of human miRNAs. *Nucleic Acids Res.* **2017**, *45*, 8731–8744. [[CrossRef](#)]
8. Huang, Z.; Shi, J.; Gao, Y.; Cui, C.; Zhang, S.; Li, J.; Zhou, Y.; Cui, Q. HMDD v3.0: A database for experimentally supported human microRNA-disease associations. *Nucleic Acids Res.* **2019**, *47*, D1013–D1017. [[CrossRef](#)]
9. Xie, B.; Ding, Q.; Han, H.; Wu, D. miRCancer: A microRNA-cancer association database constructed by text mining on literature. *Bioinformatics* **2013**, *29*, 638–644. [[CrossRef](#)] [[PubMed](#)]
10. Cui, T.; Zhang, L.; Huang, Y.; Yi, Y.; Tan, P.; Zhao, Y.; Hu, Y.; Xu, L.; Li, E.; Wang, D. MNDR v2.0: An updated resource of ncRNA-disease associations in mammals. *Nucleic Acids Res.* **2018**, *46*, D371–D374. [[CrossRef](#)]
11. Guo, Z.; Kuang, Z.; Wang, Y.; Zhao, Y.; Tao, Y.; Cheng, C.; Yang, J.; Lu, X.; Hao, C.; Wang, T.; et al. PmiREN: A comprehensive encyclopedia of plant miRNAs. *Nucleic Acids Res.* **2020**, *48*, D1114–D1121. [[CrossRef](#)] [[PubMed](#)]
12. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **2019**, *47*, D330–D338. [[CrossRef](#)] [[PubMed](#)]
13. Kanehisa, M.; Sato, Y.; Furumichi, M.; Morishima, K.; Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **2019**, *47*, D590–D595. [[CrossRef](#)] [[PubMed](#)]
14. Huntley, R.P.; Kramarz, B.; Sawford, T.; Umrao, Z.; Kalea, A.; Acquaah, V.; Martin, M.J.; Mayr, M.; Lovering, R.C. Expanding the horizons of microRNA bioinformatics. *RNA* **2018**, *24*, 1005–1017. [[CrossRef](#)]
15. Huntley, R.P.; Sitnikov, D.; Orlic-Milacic, M.; Balakrishnan, R.; D'Eustachio, P.; Gillespie, M.E.; Howe, D.; Kalea, A.Z.; Maegdefessel, L.; Osumi-Sutherland, D.; et al. Guidelines for the functional annotation of microRNAs using the Gene Ontology. *RNA* **2016**, *22*, 667–676. [[CrossRef](#)]

16. Ziemann, M.; Kaspi, A.; El-Osta, A. Evaluation of microRNA alignment techniques. *RNA* **2016**, *22*, 1120–1138. [[CrossRef](#)]
17. Bisgin, H.; Gong, B.; Wang, Y.; Tong, W. Evaluation of Bioinformatics Approaches for Next-Generation Sequencing Analysis of microRNAs with a Toxicogenomics Study Design. *Front. Genet.* **2018**, *9*, 22. [[CrossRef](#)]
18. Bleazard, T.; Lamb, J.A.; Griffiths-Jones, S. Bias in microRNA functional enrichment analysis. *Bioinformatics* **2015**, *31*, 1592–1598. [[CrossRef](#)]
19. Godard, P.; van Eyll, J. Pathway analysis from lists of microRNAs: Common pitfalls and alternative strategy. *Nucleic Acids Res.* **2015**, *43*, 3490–3497. [[CrossRef](#)]
20. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **2009**, *37*, 1–13. [[CrossRef](#)]
21. Riffo-Campos, Á.L.; Riquelme, I.; Brebi-Mieville, P. Tools for Sequence-Based miRNA Target Prediction: What to Choose? *Int. J. Mol. Sci.* **2016**, *17*, 1987. [[CrossRef](#)] [[PubMed](#)]
22. Peterson, S.M.; Thompson, J.A.; Ufkin, M.L.; Sathyanarayana, P.; Liaw, L.; Congdon, C.B. Common features of microRNA target prediction tools. *Front. Genet.* **2014**, *5*, 23. [[CrossRef](#)] [[PubMed](#)]
23. Witkos, T.M.; Koscianska, E.; Krzyzosiak, W.J. Practical Aspects of microRNA Target Prediction. *Curr. Mol. Med.* **2011**, *11*, 93–109. [[CrossRef](#)]
24. Brennecke, J.; Stark, A.; Russell, R.B.; Cohen, S.M. Principles of MicroRNA—Target Recognition. *PLoS Biol.* **2005**, *3*, e85. [[CrossRef](#)]
25. Kertesz, M.; Iovino, N.; Unnerstall, U.; Gaul, U.; Segal, E. The role of site accessibility in microRNA target recognition. *Nat. Genet.* **2007**, *39*, 1278–1284. [[CrossRef](#)]
26. Grimson, A.; Farh, K.K.-H.; Johnston, W.K.; Garrett-Engele, P.; Lim, L.P.; Bartel, D.P. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Mol. Cell* **2007**, *27*, 91–105. [[CrossRef](#)] [[PubMed](#)]
27. Doench, J.G. Specificity of microRNA target selection in translational repression. *Genes Dev.* **2004**, *18*, 504–511. [[CrossRef](#)]
28. Garcia, D.M.; Baek, D.; Shin, C.; Bell, G.W.; Grimson, A.; Bartel, D.P. Weak seed-pairing stability and high target-site abundance decrease the proficiency of *lscy-6* and other microRNAs. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1139–1146. [[CrossRef](#)]
29. Enright, A.J.; John, B.; Gaul, U.; Tuschl, T.; Sander, C.; Marks, D.S. MicroRNA targets in *Drosophila*. *Genome Biol.* **2003**, *5*, R1. [[CrossRef](#)]
30. Yousef, M.; Jung, S.; Kossenkov, A.V.; Showe, L.C.; Showe, M.K. Naïve Bayes for microRNA target predictions—machine learning for microRNA targets. *Bioinforma. Oxf. Engl.* **2007**, *23*, 2987–2992. [[CrossRef](#)]
31. John, B.; Enright, A.J.; Aravin, A.; Tuschl, T.; Sander, C.; Marks, D.S. Human MicroRNA Targets. *PLoS Biol.* **2004**, *2*, e363. [[CrossRef](#)] [[PubMed](#)]
32. Agarwal, V.; Bell, G.W.; Nam, J.-W.; Bartel, D.P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **2015**, *4*, e05005. [[CrossRef](#)]
33. Wang, X. Composition of seed sequence is a major determinant of microRNA targeting patterns. *Bioinformatics* **2014**, *30*, 1377–1383. [[CrossRef](#)] [[PubMed](#)]
34. Ding, J.; Li, X.; Hu, H. TarPmiR: A new approach for microRNA target site prediction. *Bioinformatics* **2016**, *32*, 2768–2775. [[CrossRef](#)] [[PubMed](#)]
35. Liu, W.; Wang, X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biol.* **2019**, *20*, 18. [[CrossRef](#)] [[PubMed](#)]
36. Chen, Y.; Wang, X. miRDB: An online database for prediction of functional microRNA targets. *Nucleic Acids Res.* **2020**, *48*, D127–D131. [[CrossRef](#)]
37. Paraskevopoulou, M.D.; Georgakilas, G.; Kostoulas, N.; Vlachos, I.S.; Vergoulis, T.; Reczko, M.; Filippidis, C.; Dalamagas, T.; Hatzigeorgiou, A.G. DIANA-microT web server v5.0: Service integration into miRNA functional analysis workflows. *Nucleic Acids Res.* **2013**, *41*, W169–W173. [[CrossRef](#)]
38. Oliveira, A.C.; Bovolenta, L.A.; Nachtigall, P.G.; Herkenhoff, M.E.; Lemke, N.; Pinhal, D. Combining Results from Distinct MicroRNA Target Prediction Tools Enhances the Performance of Analyses. *Front. Genet.* **2017**, *8*, 59. [[CrossRef](#)]
39. Tabas-Madrid, D.; Muniategui, A.; Sánchez-Caballero, I.; Martínez-Herrera, D.J.; Sorzano, C.O.S.; Rubio, A.; Pascual-Montano, A. Improving miRNA-mRNA interaction predictions. *BMC Genom.* **2014**, *15* (Suppl. 10), S2. [[CrossRef](#)]

40. Shirdel, E.A.; Xie, W.; Mak, T.W.; Jurisica, I. NAViGaTing the micronome—Using multiple microRNA prediction databases to identify signalling pathway-associated microRNAs. *PLoS ONE* **2011**, *6*, e17429. [[CrossRef](#)]
41. Andrés-León, E.; González Peña, D.; Gómez-López, G.; Pisano, D.G. miRGate: A curated database of human, mouse and rat miRNA-mRNA targets. *Database J. Biol. Databases Curation* **2015**, *2015*, bav035. [[CrossRef](#)]
42. Le, T.D.; Zhang, J.; Liu, L.; Li, J. Ensemble Methods for MiRNA Target Prediction from Expression Data. *PLoS ONE* **2015**, *10*, e0131627. [[CrossRef](#)] [[PubMed](#)]
43. Quillet, A.; Saad, C.; Ferry, G.; Anouar, Y.; Vergne, N.; Lecroq, T.; Dubessy, C. Improving Bioinformatics Prediction of microRNA Targets by Ranks Aggregation. *Front. Genet.* **2020**, *10*, 1330. [[CrossRef](#)] [[PubMed](#)]
44. Huang, H.-Y.; Lin, Y.-C.-D.; Li, J.; Huang, K.-Y.; Shrestha, S.; Hong, H.-C.; Tang, Y.; Chen, Y.-G.; Jin, C.-N.; Yu, Y.; et al. miRTarBase 2020: Updates to the experimentally validated microRNA—Target interaction database. *Nucleic Acids Res.* **2020**, *48*, D148–D154. [[CrossRef](#)] [[PubMed](#)]
45. Karagkouni, D.; Paraskevopoulou, M.D.; Chatzopoulos, S.; Vlachos, I.S.; Tastsoglou, S.; Kanellos, I.; Papadimitriou, D.; Kavakiotis, I.; Maniou, S.; Skoufos, G.; et al. DIANA-TarBase v8: A decade-long collection of experimentally supported miRNA—Gene interactions. *Nucleic Acids Res.* **2018**, *46*, D239–D245. [[CrossRef](#)]
46. Jiang, Q.; Wang, Y.; Hao, Y.; Juan, L.; Teng, M.; Zhang, X.; Li, M.; Wang, G.; Liu, Y. miR2Disease: A manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res.* **2009**, *37*, D98–D104. [[CrossRef](#)]
47. Schriml, L.M.; Mitraka, E.; Munro, J.; Tauber, B.; Schor, M.; Nickle, L.; Felix, V.; Jeng, L.; Bearer, C.; Lichtenstein, R.; et al. Human Disease Ontology 2018 update: Classification, content and workflow expansion. *Nucleic Acids Res.* **2019**, *47*, D955–D962. [[CrossRef](#)]
48. Ruepp, A.; Kowarsch, A.; Schmidl, D.; Buggenthin, F.; Brauner, B.; Dunger, I.; Fobo, G.; Frishman, G.; Montrone, C.; Theis, F.J. PhenomiR: A knowledgebase for microRNA expression in diseases and biological processes. *Genome Biol.* **2010**, *11*, R6. [[CrossRef](#)]
49. Amberger, J.S.; Bocchini, C.A.; Scott, A.F.; Hamosh, A. OMIM.org: Leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res.* **2019**, *47*, D1038–D1043. [[CrossRef](#)]
50. Liu, X.; Wang, S.; Meng, F.; Wang, J.; Zhang, Y.; Dai, E.; Yu, X.; Li, X.; Jiang, W. SM2miR: A database of the experimentally validated small molecules' effects on microRNA expression. *Bioinform. Oxf. Engl.* **2013**, *29*, 409–411. [[CrossRef](#)]
51. Mi, H.; Muruganujan, A.; Huang, X.; Ebert, D.; Mills, C.; Guo, X.; Thomas, P.D. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* **2019**, *14*, 703–721. [[CrossRef](#)] [[PubMed](#)]
52. Jassal, B.; Matthews, L.; Viteri, G.; Gong, C.; Lorente, P.; Fabregat, A.; Sidiropoulos, K.; Cook, J.; Gillespie, M.; Haw, R.; et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* **2020**, *48*, D498–D503. [[CrossRef](#)] [[PubMed](#)]
53. Slenter, D.N.; Kutmon, M.; Hanspers, K.; Riutta, A.; Windsor, J.; Nunes, N.; Mélius, J.; Cirillo, E.; Coort, S.L.; Digles, D.; et al. WikiPathways: A multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res.* **2018**, *46*, D661–D667. [[CrossRef](#)] [[PubMed](#)]
54. Köhler, S.; Carmody, L.; Vasilevsky, N.; Jacobsen, J.O.B.; Danis, D.; Gourdine, J.-P.; Gargano, M.; Harris, N.L.; Matentzoglou, N.; McMurry, J.A.; et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res.* **2019**, *47*, D1018–D1027. [[CrossRef](#)]
55. Firth, H.V.; Richards, S.M.; Bevan, A.P.; Clayton, S.; Corpas, M.; Rajan, D.; Vooren, S.V.; Moreau, Y.; Pettett, R.M.; Carter, N.P. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.* **2009**, *84*, 524–533. [[CrossRef](#)]
56. Kehl, T.; Kern, F.; Backes, C.; Fehlmann, T.; Stöckel, D.; Meese, E.; Lenhof, H.-P.; Keller, A. miRPathDB 2.0: A novel release of the miRNA Pathway Dictionary Database. *Nucleic Acids Res.* **2020**, *48*, D142–D147. [[CrossRef](#)]
57. Ebrahimkhani, S.; Vafaee, F.; Hallal, S.; Wei, H.; Lee, M.Y.T.; Young, P.E.; Satgunaseelan, L.; Beadnall, H.; Barnett, M.H.; Shivalingam, B.; et al. Deep sequencing of circulating exosomal microRNA allows non-invasive glioblastoma diagnosis. *NPJ Precis. Oncol.* **2018**, *2*, 28. [[CrossRef](#)]
58. Li, J.; Han, X.; Wan, Y.; Zhang, S.; Zhao, Y.; Fan, R.; Cui, Q.; Zhou, Y. TAM 2.0: Tool for MicroRNA set analysis. *Nucleic Acids Res.* **2018**, *46*, W180–W185. [[CrossRef](#)]

59. Garcia-Alonso, L.; Holland, C.H.; Ibrahim, M.M.; Turei, D.; Saez-Rodriguez, J. Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res.* **2019**, *29*, 1363–1375. [[CrossRef](#)]
60. Bult, C.J.; Blake, J.A.; Smith, C.L.; Kadin, J.A.; Richardson, J.E.; The Mouse Genome Database Group; Anagnostopoulos, A.; Asabor, R.; Baldarelli, R.M.; Beal, J.S.; et al. Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* **2019**, *47*, D801–D806. [[CrossRef](#)]
61. Davis, A.P.; Grondin, C.J.; Johnson, R.J.; Sciaky, D.; McMorran, R.; Wiegers, J.; Wiegers, T.C.; Mattingly, C.J. The Comparative Toxicogenomics Database: Update 2019. *Nucleic Acids Res.* **2019**, *47*, D948–D954. [[CrossRef](#)] [[PubMed](#)]
62. Stathias, V.; Turner, J.; Koleti, A.; Vidovic, D.; Cooper, D.; Fazel-Najafabadi, M.; Pilarczyk, M.; Terry, R.; Chung, C.; Umeano, A.; et al. LINCS Data Portal 2.0: Next generation access point for perturbation-response signatures. *Nucleic Acids Res.* **2020**, *48*, D431–D439. [[CrossRef](#)] [[PubMed](#)]
63. Whirl-Carrillo, M.; McDonagh, E.M.; Hebert, J.M.; Gong, L.; Sangkuhl, K.; Thorn, C.F.; Altman, R.B.; Klein, T.E. Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* **2012**, *92*, 414–417. [[CrossRef](#)] [[PubMed](#)]
64. Russo, F.; Di Bella, S.; Vannini, F.; Berti, G.; Scoyni, F.; Cook, H.V.; Santos, A.; Nigita, G.; Bonnici, V.; Laganà, A.; et al. miRandola 2017: A curated knowledge base of non-invasive biomarkers. *Nucleic Acids Res.* **2018**, *46*, D354–D359. [[CrossRef](#)]
65. Dweep, H.; Gretz, N. miRWalk2.0: A comprehensive atlas of microRNA-target interactions. *Nat. Methods* **2015**, *12*, 697. [[CrossRef](#)]
66. Teng, X.; Chen, X.; Xue, H.; Tang, Y.; Zhang, P.; Kang, Q.; Hao, Y.; Chen, R.; Zhao, Y.; He, S. NPInter v4.0: An integrated database of ncRNA interactions. *Nucleic Acids Res.* **2019**, *48*, D160–D165. [[CrossRef](#)]
67. Zhang, T.; Tan, P.; Wang, L.; Jin, N.; Li, Y.; Zhang, L.; Yang, H.; Hu, Z.; Zhang, L.; Hu, C.; et al. RNALocate: A resource for RNA subcellular localizations. *Nucleic Acids Res.* **2017**, *45*, D135–D138. [[CrossRef](#)]
68. Ludwig, N.; Leidinger, P.; Becker, K.; Backes, C.; Fehlmann, T.; Pallasch, C.; Rheinheimer, S.; Meder, B.; Stähler, C.; Meese, E.; et al. Distribution of miRNA expression across human tissues. *Nucleic Acids Res.* **2016**, *44*, 3865–3877. [[CrossRef](#)]
69. Tong, Z.; Cui, Q.; Wang, J.; Zhou, Y. TransmiR v2.0: An updated transcription factor-microRNA regulation database. *Nucleic Acids Res.* **2019**, *47*, D253–D258. [[CrossRef](#)]
70. Nishimura, D. BioCarta. *Biotech Softw. Internet Rep.* **2001**, *2*, 117–120. [[CrossRef](#)]
71. Schaefer, C.F.; Anthony, K.; Krupa, S.; Buchoff, J.; Day, M.; Hannay, T.; Buetow, K.H. PID: The Pathway Interaction Database. *Nucleic Acids Res.* **2009**, *37*, D674–D679. [[CrossRef](#)] [[PubMed](#)]
72. Chang, L.; Zhou, G.; Soufan, O.; Xia, J. miRNet 2.0: Network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic Acids Res.* **2020**, *48*, W244–W251. [[CrossRef](#)] [[PubMed](#)]
73. Tabas-Madrid, D.; Nogales-Cadenas, R.; Pascual-Montano, A. GeneCodis3: A non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res.* **2012**, *40*, W478–W483. [[CrossRef](#)] [[PubMed](#)]
74. Carmona-Saez, P.; Chagoyen, M.; Tirado, F.; Carazo, J.M.; Pascual-Montano, A. GENECODIS: A web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol.* **2007**, *8*, R3. [[CrossRef](#)]
75. Kern, F.; Fehlmann, T.; Solomon, J.; Schwed, L.; Backes, C.; Meese, E.; Keller, A. miEAA 2.0: Integrating multi-species microRNA enrichment analysis and workflow management systems. *Nucleic Acids Res.* **2020**, *48*, W521–W528. [[CrossRef](#)] [[PubMed](#)]
76. Licursi, V.; Conte, F.; Fiscon, G.; Paci, P. MIENETURNET: An interactive web tool for microRNA-target enrichment and network-based analysis. *BMC Bioinform.* **2019**, *20*, 545. [[CrossRef](#)]
77. Preusse, M.; Theis, F.J.; Mueller, N.S. miTALOS v2: Analyzing Tissue Specific microRNA Function. *PLoS ONE* **2016**, *11*, e0151771. [[CrossRef](#)]
78. Li, J.-H.; Liu, S.; Zhou, H.; Qu, L.-H.; Yang, J.-H. starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* **2014**, *42*, D92–D97. [[CrossRef](#)]
79. Betel, D.; Koppal, A.; Agius, P.; Sander, C.; Leslie, C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* **2010**, *11*, R90. [[CrossRef](#)]

80. Lu, T.-P.; Lee, C.-Y.; Tsai, M.-H.; Chiu, Y.-C.; Hsiao, C.K.; Lai, L.-C.; Chuang, E.Y. miRSystem: An Integrated System for Characterizing Enriched Functions and Pathways of MicroRNA Targets. *PLoS ONE* **2012**, *7*, e42390. [[CrossRef](#)]
81. Vlachos, I.S.; Zagganas, K.; Paraskevopoulou, M.D.; Georgakilas, G.; Karagkouni, D.; Vergoulis, T.; Dalamagas, T.; Hatzigeorgiou, A.G. DIANA-miRPath v3.0: Deciphering microRNA function with experimental support. *Nucleic Acids Res.* **2015**, *43*, W460–W466. [[CrossRef](#)] [[PubMed](#)]
82. Hosack, D.A.; Dennis, G.; Sherman, B.T.; Lane, H.C.; Lempicki, R.A. Identifying biological themes within lists of genes with EASE. *Genome Biol.* **2003**, *4*, R70. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).