Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 108C (2017) 1125–1134

Procedia Computer Science

International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland

Rank miRNA: a web tool for identifying polymorphisms altering miRNA target sites

Stefano Beretta[1], Carlo Maj[2] and Ivan Merelli[3]

[1] Dipartimento di Informatica, Sistemistica e Comunicazione, Università degli Studi di Milano - Bicocca, Italia.
beretta@disco.unimib.it

[2] Genetics Unit, IRCCS Centro S. Giovanni di Dio, Fatebenefratelli, Brescia, Italia.
cmaj@fatebenefratelli.eu

[3] Istituto di Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, Segrate, Italia.
ivan.merelli@itb.cnr.it

Abstract

MicroRNAs (miRNAs) are small non-coding RNA molecules that have an important role in a wide range of biological processes, since they interact with specific mRNAs affecting the expression of the corresponding proteins. The role of miRNA can be deeply influenced by *Single Nucleotide Polymorphisms* (SNPs), in particular in their seed sites, since these variations may modify their affinity with particular transcripts, but they may also generate novel binding capabilities for specific miRNA binding sites or destroy them. Several computational tools for miRNA-target site predictions have been developed, but the obtained results are often not in agreement, making the study the binding sites hard, and the analysis of SNP effects even harder.

For these reasons, we developed a web application called `Rank miRNA`, which allows to retrieve and aggregate the results of three prediction tools, but also to process and compare new input miRNA sequences, allowing the analysis of how variations impact on their function. Therefore, our tool is also able to predict the impact of SNPs (and any other kind of variations) on miRNA-mRNA binding capability and also to find the target genes of (potentially new) miRNA sequences.

We evaluated the performance of `Rank miRNA` on specific human SNPs, which are likely to be involved in several mental disorder diseases, showing the potentiality of our tool in helping the study of miRNA-target interactions.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Science

Keywords: miRNA, List Re-ranking, miRNA-Target Prediction

1877-0509
Peer-review under responsibility of the scientific committee of the International Conference on Computational Science
10.1016/j.procs.2017.05.189

1 Introduction

Numerous *Single Nucleotide Polymorphisms* (SNPs) associated with complex diseases have been identified by *Genome-Wide Association Studies* (GWAS), which are large-scale analysis of specific DNA loci of participants, usually divided in people with a disease (cases) and similar people without diseases (controls), to understand the association of particular genotypes to a specific pathology. However, few of these SNPs have explicit biological functions, since determining the effects of disease-associated SNPs is a complex problem [1, 2].

A possible solution is to correlate the expression of genes/proteins to different genotypes, using a technique called *expression Quantitative Trait Loci* (eQTLs). This approach is particularly suitable to interpret the results of genome-wide association studies, with thousands of variants in association with complex traits and diseases, most of which are non-coding and therefore difficult to correlate to causal genes. While several eQTL studies showed that disease-predisposing variants often affect the expression levels of nearby genes, which are called cis-eQTLs, recent studies have also identified trans-eQTLs [3] showing unpredictable remote consequences of some variants. Moreover, techniques relying on *Chromatin Conformation Capture* can be very useful to explain the functional mechanisms behind these long-range interactions [4].

Relying on these approaches, recent studies indicated that SNPs within the 3'UTR regions of susceptibility genes could affect complex traits/diseases, by affecting the function of microRNAs (or miRNAs). These 3'UTR SNPs are functional candidates to explain the association of particular genotypes with specific diseases, and therefore they are of great interest in genomics and medicine. For example, a large meta-analysis revealed that specific miRNAs are differentially expressed in brain and cerebrospinal fluid of patients with respect to controls [5]. At the same time, it has been demonstrated that SNPs in the sequence of a miRNA largely influence its binding capability to specific mRNA, impacting its regulatory function [6].

Although some methods have been implemented to analyze the impact of these variations on miRNAs-mRNAs bindings (see [7, 8]), they typically rely on a single tool, which analyze how the score of a miRNA-target interaction changes, according to a given genetic variation. Unfortunately, by comparing computational results with experimental validations, we can observe that these tools produce a large number of false positive predictions [9]. This fact suggests that creating a consensus between different approaches can be very useful to reduce these false positive results.

Among the most known integration tools, miRGator [10] and ExprTarget [11] try to perform this task by exploiting functional analysis and genome annotation, to better characterize the identified targets, and to allow a more suitable discrimination of the predicted list of targets. Moreover, miRGator also provides a miRNAs expression profile by importing expression experiments from the GEO databank [12]. Analogously, expression profiles are reported in mESAdb [13] and miRex [14]. The MAGIA [15] tool allows both to retrieve predictions as union or intersection of results produced by TargetScan [16, 17], miRanda [18, 19], and PITA [20], and to integrate mRNA expression values with a miRNA expression score, in order to elucidate an inverse correlation, thus hypothesizing new miRNA-target couples.

Anyway, none of these tools deals with the impact of variations in the nucleotide sequence that could affect the miRNA-target pairing. Some tools are available to predict the impact of SNPs in the mRNA sequence, e.g. MirSNP [7] and PolymiRTS [21], but they rely on a single algorithm and do not consider SNPs in the miRNA (but only in the mRNA). Other approaches allow to predict the pairing capability of a user-defined miRNA sequence, but they rely on a single algorithm and do not allow any comparison [22, 23], or other methods exploit machine learning approaches to integrate different predictions [24].

For this reason, we developed a publicly available online database and web application called `Rank miRNA`, which allows to predict the impact of SNPs on the pairing capability of miRNA-mRNA binding sites and also to find the target genes of (potentially new) miRNA sequences. More precisely, `Rank miRNA` allows to retrieve and aggregate in a consensus ranked list, the prediction results of three prediction tools: miRanda [18, 19], TargetScan [16, 17], and RNAhybrid [25]. This task is accomplished starting from a user-defined input miRNA sequence, on which its target interactions are computed with each of the considered tool, and then their results are aggregated into a ranked list, so that new variants of the miRNA sequences can be easily tested. Therefore, the real strength of the web application is the possibility of comparing the (aggregate) prediction results of two given input miRNA sequences, so that the impact of variants in the sequence on the target predictions can be obtained and evaluated.

`Rank miRNA` is freely accessible at `http://155.253.6.106/rankmirna` for testing purposes.

2 Background

MicroRNAs (miRNAs) are small non-coding RNA molecules of ~ 22 nucleotides that primarily mediate post-transcriptional gene silencing processes in animals [26, 27]. MiRNAs inactivate specific mRNAs and interfere with the translation of the encoded proteins [28]. In mammals, miRNAs are predicted to control the activities of $\sim 50\%$ of all the protein-coding genes [29]. As key post-transcriptional regulators, miRNAs have an important role in a wide range of biological processes, including cell proliferation, differentiation, apoptosis, and metabolism [26, 27]. Evidences indicate that miRNAs are also involved in the pathogenesis of complex diseases, such as cancer and mental disorders [30, 29].

Complementarity of bases $2-8$ of the miRNA (called the *seed site*) is important for the miRNA-mRNA binding [31, 32]. MiRNAs are key regulators of gene expression and, therefore, SNPs in their seed sites may create, as well as destroy, miRNA binding sites, and further affect phenotypes and disease susceptibility. Identifying these seed-site SNPs could help in the exploration of the molecular mechanism of gene dysregulation. In addition, genetic variants in miRNA genes may also have important roles by affecting the miRNA maturation, which may be involved in disease susceptibility [33]. Certain polymorphisms in miRNA genes have been found to be associated with various complex diseases, including cancers, mental diseases, cardiomyopathy, and asthma.

It is a challenging problem to identify miRNAs on experimental basis, because of their limited expression, their structural/sequence variability during miRNA maturation process, and also considering the tissue specificity of their control mechanism [34]. Moreover, it must be pointed out that a miRNA can affect the expression of several genes (up to hundreds of different transcripts), probably in a tissue specific manner, and experimental validation is needed to corroborate their biological effect. The reduction of the target protein in miRNA transfected cells can validate the existence of a specific miRNA-target interaction. This technique is very complex and costly to achieve, thus limiting the number of experiments that can be performed.

Therefore, computational predictions represent a very important approach to screen possible targets to be experimentally tested. More precisely, the interactions between a miRNA and its mRNA target sites can be considered from a thermodynamic, probabilistic, and evolutionary (or sequence-based) point of view. Several computational tools [35] for miRNA-target sites prediction have been developed in the last years, using one or more of the aforementioned aspects.

Among the most known prediction tools for miRNA-targets recognition are miRanda [18, 19], TargetScan [17, 16], and RNAhybrid [25]. miRanda performs three sequential steps: (i) sequence

matching to find the maximal local complementarity between the mature miRNA and the putative target site; (ii) free energy calculation to estimate the strength of the potential RNA duplex; (iii) filtering of predicted targets on the basis of evolutionary conservation. TargetScan is based on two hypothesis: (i) highly conserved miRNAs are more involved in regulation, and (ii) membership in large miRNA families leads to a higher number of existing targets. After the matching step (allowing wobble pairs and stopping at the first mismatch encountered), thermodynamical evaluation of the RNA duplex is performed. Finally, RNAhybrid predicts the target genes based on free energy calculation: collects the energetic most favorable structures, normalizes them, and then uses estimated p-values to determine the significance of each predicted binding site.

3 Method

We developed a web application that tries to integrate the predictions of the aforementioned tools on miRNA-target interactions of the given input miRNA sequences. This feature consists of running the three tools on a given input sequence of a miRNA, and then integrating the lists of predicted target genes with a procedure based on list re-ranking. Figure 1 presents an overview of the result aggregation procedure employed in the `Rank miRNA` tool.

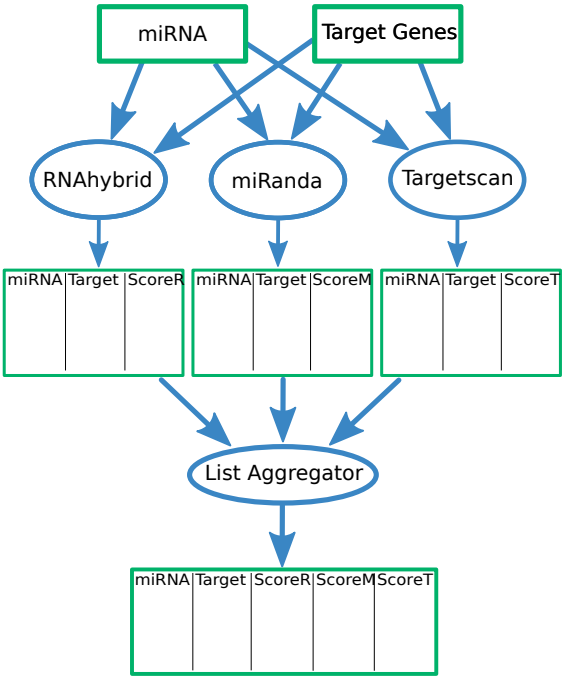

Figure 1: Workflow of the rank aggregation procedure. Starting from miRNA and genes sequences, the predictions of their interactions are computed with TargetScan, miRanda, and RNAhybrid. Then, the (ranked) result lists are aggregated employing a median-based procedure.

The integration of the results obtained with different prediction algorithms is performed by taking advantage of a previously published approach based on list re-ranking [9], which overcomes the mere intersection or union of the results provided by some of the previously

algorithms. The method employed in this work is a further improvement, obtained by replacing the re-ranking algorithm, from the simple *Borda* count to the *median* rank aggregation.

More formally, let $E = \{e_1, e_2, \cdots, e_n\}$ be the set of elements to be ranked, and let $S_k = \{\sigma_1, \sigma_2, \cdots, \sigma_k\}$ be a set of k different ranks of the elements in E, that is, a (ordered) permutation of its elements. We denote by $\sigma_i(e_j)$, with $1 \leq i \leq k$ and $1 \leq j \leq n$, the position of element e_j (i.e. its rank) in the ordered permutation σ_i, and $\sigma_i(e_j) < \sigma_i(e_z)$ means that σ_i ranks the element e_i above element e_z.

In our approach, we employ an heuristic approach to aggregate the different ranking lists, based on the median. More precisely, the new method computes the median position (rank) of each interaction between miRNA and target obtained from the different prediction algorithms, and creates the "best" miRNA-target list. Formally, given the set $S_k = \{\sigma_1, \sigma_2, \cdots, \sigma_k\}$ of the different rankings of the elements of E, we compute $\mu'(e_i = median(\sigma_1(e_i), \sigma_2(e_i), \cdots, \sigma_k(e_i))$, with $1 \leq i \leq n$, and we order the permutation μ' to obtain the final aggregate rank μ.

Consider the following example in which we have three different rankings of the elements $E = \{a, b, c, d\}$, that is: $\sigma_1 = \langle a, b, c, d\rangle$, $\sigma_2 = \langle b, a, d, c\rangle$, and $\sigma_3 = \langle b, a, c, d\rangle$. Then, we compute $\mu'(a) = median\{1, 2, 2\} = 2$, $\mu'(b) = median\{1, 1, 2\} = 1$, $\mu'(c) = median\{3, 3, 4\} = 3$, and $\mu'(d) = median\{3, 4, 4\} = 4$, so that the resulting (ordered) ranking is $\mu = \langle b, a, c, d\rangle$.

This re-ranking algorithm guarantees that the new list is as close as possible to all the individual ordered lists provided by the considered prediction tools, according to the *Spearman* footrule distance. This latter measure is the sum of the absolute distances between the ranks of all the unique elements from all the original lists, and the aggregated one. Formally, given a set of elements $E = \{e_1, e_2, \cdots, e_n\}$ and two rankings σ and μ of its elements, the footrule distance F is defined as:

$$F(\sigma, \mu) = \sum_{i=1}^{n} |\sigma(e_i) - \mu(e_i)|$$

Recalling the previous example, $F(\sigma_1, \mu) = |1 - 2| + |2 - 1| + |3 - 3| + |4 - 4| = 2$, that is, the first two elements are switched, changing one position for each of them.

The median rank aggregation approach provides better results with respect to those obtained with the *Borda* count approach, still allowing to order the final miRNA-target list. Moreover, it guarantees a 3-approximation of the *Spearman* footrule optimal aggregation, with respect to the 5-approximation of the *Borda* approach.

As shown in Figure 1, we apply this median-based procedure to aggregate the lists of results obtained with the three different tools for predicting miRNA-target interactions, namely TargetScan, miRanda, and RNAhybrid.

In addition to the possibility of computing the aggregated results for a given input miRNA sequence, we also developed a procedure to compare the predictions on two of them. The idea is to allow the possibility of evaluating the impact of variants in the nucleotide sequence of a miRNA on its target genes. In fact, to help in the identification of putative SNPs that can influence the binding of miRNA to the target 3'UTR of a gene, we designed an integrative approach which is able to estimate the impact of SNPs that are located in miRNA target sites. More precisely, to compare the predictions on the two miRNA sequences, say s_1 and s_2, `Rank miRNA` runs the same procedure described above to aggregate the results of the three considered tools on a single sequence, and then compares the aggregated result lists, that is, μ_{s_1} and μ_{s_2}. To do this, starting from the aggregated result lists computed for each of the two sequences μ_{s_1} and μ_{s_2}, `Rank miRNA` extracts:

- interactions present only in μ_{s_1}: target interactions of s_1 that *disappear* in μ_{s_2};
- interactions present only in μ_{s_2}: new target interactions *created* by s_2;

- differences of predictions in μ_{s_1} w.r.t. the same ones in μ_{s_2}: *variations* of the scores of the target interactions present in both the lists.

In details, the idea is to compare similar miRNA sequences, like those having a SNP variation, in order to assess the differences in the predicted target genes. For this reason, we focused on the target interactions present in only one list, that is, those only found when computing the results of s_1 that *disappear* in s_2, and the ones that are *created* by the second sequence s_2 but are not present among the results of s_1. Moreover, we also highlight the *variations* of the results obtained considering s_1 w.r.t. the ones obtained for the s_2 sequence. In this case we quantify the *variations* of the scores of the considered tools: Targetscan, miRanda, and RNAhybrid.

We implemented the `Rank miRNA` web application using Java servlet, and embedded the code into JSP pages, which allow a dynamic visualization of the results as well as an easy way to interact with them.

4 Experimental Evaluation

In order to test the capability of `Rank miRNA` in evaluating the effect of specific miRNAs variants over the regulation of gene expression, by considering the alterations of miRNAs-mRNAs bindings, we searched for known examples of SNPs impacting the regulatory effect of miRNAs.

It should be noticed that the frequencies of SNPs in putative seed-matching regions are lower compared to the overall SNPs distribution in the reference human genome, suggesting a possible negative selection and, thus, a potential functional biological role of such genetic variants [36].

However, by considering the Single Nucleotide Polymorphism database (dbSNP build 137), we identified few variants within miRNA seed regions (see Table 1) known to have a biological impact on neurological disease developments.

miRNA id.	SNP id.	Chr.	Position	Strand	Seed Variation
hsa-miR-146a-3p	rs2910164	5	159912418	+	CU[C/G]UGAA
hsa-miR-221-5p	rs113054794	X	45605666	-	CCU[G/U]GCA
hsa-miR-128-1-5p	rs117812383	2	136422988	+	GGGGCC[G/A]
hsa-miR-145-3p	rs190323149	5	148810267	+	GAUU[C/U]CU
hsa-miR-124-3p	rs34059726	20	61809907	+	AA[G/U]GCAC

Table 1: Tested miRNA. The first column reports the id. of the considered miRNA, while the remaining five columns correspond to the tested SNPs, with their details: id., chromosome, position, strand, and the nucleotide variation. The coordinates (chromosome, position, and strand) are referred to the Human Genome version 19 (*hg19*).

These variants can significantly affect the regulation of gene expression exerted by miRNAs and, thus, are potential candidates for gene association studies. For instance, SNP rs2910164 in hsa-miR-146a-3p was found to be associated with Alzheimer, and interestingly a specific genotype of this SNP (GG) is correlated to higher expression of inflammatory proteins [37].

Relying on these test cases, we tested the capabilities of `Rank miRNA` in identifying the different binding affinity of miRNAs, considering the number of novel binding sites created and destroyed. Moreover, for those miRNA-target associations still present in both the lists, we evaluated the changes in the binding affinity.

Rank miRNA

- jobID: 1485945089911
- miRNA Sequence1: CCUCUGAAAUUCAGUUCUUCAG
- miRNA Sequence2: CCUGUGAAAUUCAGUUCUUCAG
- organism: hg19

Download Full Sequence1 Result file (CSV format)

Download Full Sequence2 Result file (CSV format)

Interactions Changed

Interactions for Sequence1 changed when considering Sequence2.

Download Interactions-modified Result file (CSV format)

Search:

Showing page 1 of 2

Rank	New_Rank	Name	Gene_Name	Chrom	Strand	Start	End	Miranda_Start	Miranda_End	Miranda_Score	TScan_Start	TScan_End	TScan_Score	RNAhybrid_Start	RNAhybrid_Score
29	17 ↑	NM_001126181	NRGN	chr11	+	124615621	124617102	235	259	-25.26 (-25.26)	null	null	null (null)	236	-27.4 (-27.4)
29	17 ↑	NM_006176	NRGN	chr11	+	124615621	124617102	235	259	-25.26 (-25.26)	null	null	null (null)	236	-27.4 (-27.4)
50	830 ↓	NM_001079818	ITGA6	chr2	+	173368981	173371181	89	109	-26.3 (-21.79)	102	108	-0.196 (null)	null	null (null)
56	119 ↓	NM_002547	OPHN1	chrX	-	67262186	67268265	635	656	-28.61 (-24.1)	649	655	-0.1 (null)	null	null (null)
61	595 ↓	NM_003137	SRPK1	chr6	-	35800811	35803080	879	900	-26.85 (-22.34)	893	899	-0.04 (null)	null	null (null)
63	105 ↓	NM_001193421	TSHZ2	chr20	+	51873103	52111869	62836	62858	-29.52 (-25.01)	62851	62858	-0.018 (null)	null	null (null)
63	105 ↓	NM_173485	TSHZ2	chr20	+	51873103	52111869	62836	62858	-29.52 (-25.01)	62851	62858	-0.018 (null)	null	null (null)
78	830 ↓	NM_000210	ITGA6	chr2	+	173366608	173371181	2462	2482	-26.3 (-21.79)	2475	2481	0.002 (null)	null	null (null)
84	459 ↓	NM_198560	LHFPL4	chr3	-	9540045	9543894	2499	2522	-27.31 (-22.8)	2515	2521	0.042 (null)	null	null (null)
133	94 ↑	NM_015995	KLF13	chr15	+	31664503	31670102	4435	4456	-25.5 (-26.14)	null	null	null (null)	null	null (null)

Display 10 **records per page**

Previous 1 2 Next

Figure 2: Example of results provided by `Rank miRNA`. Here, the table showing the interactions of hsa-miR-146a-3p found in both the variants of SNP rs2910164 is reported. Arrows in the second column highlight if the new interaction obtained for the second sequence has a lower (down) or higher (up) rank w.r.t. that of the first sequence.

The web site computes the predictions of both the input sequences, using the three tools described above, i.e. Targetscan, miRanda, and RNAhybrid. Then, it visualizes three tables with the achieved results. In this experimental setup, we decided to report the first 2000 high-ranked miRNA-mRNA interactions, i.e. those having the lowest median value, which are used for the comparison and split into the aforementioned tables. The motivation for this choice is that the RNAhybrid tool tends to report several results, that could lead in some cases to an overestimation of the predictions. For this reason, we selected from each aggregated list those best interactions, and we compared the difference between them, to highlight the results that are most affected by the SNP variants.

The main result table highlights genes that are targeted by both the input sequences (Figure 2), showing the changes in the list, according to the common schema in which a "down arrow" means a worst affinity, while an "up arrow" a better affinity.

Moreover, genes that, according to the achieved predictions, are targeted by the first (wild type) miRNA but not by the second (variated) sequence (destroyed binding sites), are identified. Vice versa, also genes that are not targeted by the wild type miRNA, but that seems to be controlled by the variated form (new binding sites) are reported.

Results achieved on the test datasets show that SNPs in the seed region of miRNAs have a huge impact in their ability to bind specific target genes (see Table 2), as testified by the large number of different genes that each miRNA is able to bind with and without considering the genomic variation.

By comparing the results with those in the literature, we can say that `Rank miRNA` is highly sensitive and covers most experiments confirmed SNPs that affect the miRNA function [37].

This makes `Rank miRNA` suitable to be combined by researchers' with specific GWAS or eQTL positive data sets to identify the putative miRNA-related SNPs from traits/diseases associated variants.

miRNA id.	Num. Interaction		
	Created	Disappeared	Modified
hsa-miR-146a-3p	1981	1981	19
hsa-miR-221-5p	1982	1982	18
hsa-miR-128-1-5p	1521	1521	479
hsa-miR-145-3p	1926	1926	74
hsa-miR-124-3p	1991	1991	74

Table 2: Top-ranked 2000 results achieved in the comparison of the tested miRNAs. The first column reports id of the considered miRNA. The second, third, and fourth columns represent the number of miRNA-target interactions that are created, disappear, and are modified, respectively, due to the presence of the SNP.

5 Conclusions

In this work we presented `Rank miRNA`, a web application able to analyze the impact of genomic variations on miRNA-target interactions. Indeed, our tool offers the possibility to test the potential interactions of new sequences and to compare the results obtained with two input miRNA sequences. This allows the study of how variations, including but not limited to SNPs, impact on the miRNA regulation capabilities. Our approach differs from other approaches since it offers the possibility to combine target predictions from different tools, which allows the reduction of false positive results. These combined features represent a real improvement since, to our knowledge, no other tool allows to perform the same operations on a given miRNA sequence(s).

This will make the study of the effects of specific nucleotide variations (e.g. SNPs) on the target interactions easier. This is very important, since the identification of mRNAs binding alterations as a consequence of genetic variations can help in the detection of important biochemical pathways involved in the analyzed phenotype. In particular, a number of dysregulated miRNAs have been reported to be associated with Alzheimer's disease.

Now, we are working on the integration of other tools (such as PITA) in `Rank miRNA` to make its predictions more robust, and we are also testing new algorithms to improve the list aggregation procedure. For the future, we are planning to extend the set of operations done by our tool and to improve its performance for the computation of the results and also to allow the possibility of retrieving (aggregated) interactions by specifying a miRNA or Target identifier.

References

[1] Ivan Merelli, Andrea Calabria, Paolo Cozzi, Federica Viti, Ettore Mosca, and Luciano Milanesi. Snpranker 2.0: a gene-centric data mining tool for diseases associated snp prioritization in gwas. *BMC Bioinformatics*, 14(1):S9, 2013.

[2] Harm-Jan Westra, Marjolein J Peters, Tõnu Esko, Hanieh Yaghootkar, Claudia Schurmann, and et.al. Systematic identification of trans-eqtls as putative drivers of known disease associations. *Nature genetics*, 45(10):1238–1243, 10 2013.

[3] Ivan Kel, Zisong Chang, Nadia Galluccio, Margherita Romeo, Stefano Beretta, Luisa Diomede, Alessandra Mezzelani, Luciano Milanesi, Christoph Dieterich, and Ivan Merelli. Spire, a modular pipeline for eqtl analysis of rna-seq data, reveals a regulatory hotspot controlling mirna expression in c. elegans. *Mol. BioSyst.*, 12:3447–3458, 2016.

[4] Ivan Merelli, Fabio Tordini, Maurizio Drocco, Marco Aldinucci, Pietro Liò, and Luciano Milanesi. Integrating multi-omic features exploiting chromosome conformation capture data. *Frontiers in Genetics*, 6:40, 2015.

[5] Yong-Bo Hu, Chun-Bo Li, Ning Song, Yang Zou, Sheng-Di Chen, Ru-Jing Ren, and Gang Wang. Diagnostic value of microrna for alzheimer's disease: A systematic review and meta-analysis. *Frontiers in Aging Neuroscience*, 8:13, 2016.

[6] Jian Lu and Andrew G. Clark. Impact of microrna regulation on variation in human gene expression. *Genome Research*, 22(7):1243–1254, 2012.

[7] Chenxing Liu, Fuquan Zhang, Tingting Li, Ming Lu, Lifang Wang, Weihua Yue, and Dai Zhang. Mirsnp, a database of polymorphisms altering mirna target sites, identifies mirna-related snps in gwas snps and eqtls. *BMC Genomics*, 13(1):661, 2012.

[8] Ivan Merelli, Pietro Liò, and Luciano Milanesi. Nuchart: An r package to study gene spatial neighbourhoods with multi-omics annotations. *PLOS ONE*, 8(9), 09 2013.

[9] Dario Corrada, Federica Viti, Ivan Merelli, Cristina Battaglia, and Luciano Milanesi. mymir: a genome-wide microrna targets identification and annotation tool. *Briefings in Bioinformatics*, 12(6):588, 2011.

[10] Seungyoon Nam, Bumjin Kim, Seokmin Shin, and Sanghyuk Lee. mirgator: an integrated system for functional annotation of micrornas. *Nucleic Acids Research*, 36(suppl_1):D159, 2007.

[11] Eric R. Gamazon, Hae-Kyung Im, Shiwei Duan, Yves A. Lussier, Nancy J. Cox, M. Eileen Dolan, and Wei Zhang. Exprtarget: An integrative approach to predicting human microrna targets. *PLOS ONE*, 5(10):1–8, 10 2010.

[12] Ron Edgar, Michael Domrachev, and Alex E Lash. Gene expression omnibus: Ncbi gene expression and hybridization array data repository. *Nucleic Acids Research*, 30(1):207–210, 01 2002.

[13] Koray D. Kaya, Gökhan Karakülah, Cengiz M. Yakıcıer, Aybar C. Acar, and Özlen Konu. mesadb: microrna expression and sequence analysis database. *Nucleic Acids Research*, 39(suppl_1):D170, 2010.

[14] Andrzej Zielezinski, Jakub Dolata, Sylwia Alaba, Katarzyna Kruszka, Andrzej Pacak, Aleksandra Swida-Barteczka, Katarzyna Knop, Agata Stepien, Dawid Bielewicz, Halina Pietrykowska, Izabela Sierocka, Lukasz Sobkowiak, Alicja Lakomiak, Artur Jarmolowski, Zofia Szweykowska-Kulinska, and Wojciech M Karlowski. mirex 2.0 - an integrated environment for expression profiling of plant micrornas. *BMC Plant Biology*, 15:144, 2015.

[15] Gabriele Sales, Alessandro Coppe, Andrea Bisognin, Marta Biasiolo, Stefania Bortoluzzi, and Chiara Romualdi. Magia, a web-based tool for mirna and genes integrated analysis. *Nucleic Acids Research*, 38(suppl_2):W352, 2010.

[16] Robin C. Friedman, Kyle Kai-How Farh, Christopher B. Burge, and David P. Bartel. Most mammalian mrnas are conserved targets of micrornas. *Genome Research*, 19(1):92–105, 2009.

[17] Benjamin P. Lewis, I hung Shih, Matthew W. Jones-Rhoades, David P. Bartel, and Christopher B. Burge. Prediction of mammalian microrna targets. *Cell*, 115(7):787 – 798, 2003.

[18] Anton J. Enright, Bino John, Ulrike Gaul, Thomas Tuschl, Chris Sander, and Debora S. Marks. Microrna targets in drosophila. *Genome Biology*, 5(1):R1, 2003.

[19] Bino John, Anton J Enright, Alexei Aravin, Thomas Tuschl, Chris Sander, and Debora S Marks. Human microrna targets. *PLOS Biology*, 2(11), 10 2004.

[20] Michael Kertesz, Nicola Iovino, Ulrich Unnerstall, Ulrike Gaul, and Eran Segal. The role of site accessibility in microrna target recognition. *Nat Genet*, 39(10):1278–1284, 10 2007.

[21] Anindya Bhattacharya, Jesse D. Ziebarth, and Yan Cui. Polymirts database 3.0: linking polymorphisms in micrornas and their target sites with human diseases and biological pathways. *Nucleic Acids Research*, 42(D1):D86, 2013.

[22] Nathan Wong and Xiaowei Wang. mirdb: an online resource for microrna target prediction and functional annotations. *Nucleic Acids Research*, 43(D1):D146, 2014.

[23] Claudia Coronnello and Panayiotis V Benos. Comir: combinatorial microrna target prediction tool. *Nucleic Acids Research*, 41(Web Server issue):W159–W164, 07 2013.

[24] S. Beretta, M. Castelli, Y. Martínez, L. Muñoz, S. Silva, L. Trujillo, L. Milanesi, and I. Merelli. A machine learning approach for the integration of mirna-target predictions. In *2016 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP)*, pages 528–534, Feb 2016.

[25] Marc Rehmsmeier, Peter Steffen, Matthias Höchsmann, and Robert Giegerich. Fast and effective prediction of microrna/target duplexes. *RNA*, 10(10):1507–1517, 2004.

[26] Lin He and Gregory J. Hannon. Micrornas: small rnas with a big role in gene regulation. *Nat Rev Genet*, 5(7):522–531, 07 2004.

[27] Joshua T. Mendell. Micrornas: Critical regulators of development, cellular physiology and malignancy. *Cell Cycle*, 4(9):1179–1184, 2005. PMID: 16096373.

[28] David P Bartel. Micrornas: Genomics, biogenesis, mechanism, and function. *Cell*, 116(2):281 – 297, 2004.

[29] Jacek Krol, Inga Loedige, and Witold Filipowicz. The widespread regulation of microrna biogenesis, function and decay. *Nat Rev Genet*, 11(9):597–610, 09 2010.

[30] Fen-Biao Gao. Posttranscriptional control of neuronal development by microrna networks. *Trends in Neurosciences*, 31(1):20–26, 2017/02/07.

[31] Julius Brennecke, Alexander Stark, Robert B Russell, and Stephen M Cohen. Principles of microrna-target recognition. *PLOS Biology*, 3(3), 02 2005.

[32] Eric C. Lai. Micro rnas are complementary to 3[prime] utr sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet*, 30(4):363–364, 04 2002.

[33] Bríd M. Ryan, Ana I. Robles, and Curtis C. Harris. Genetic variation in microrna networks: the implications for cancer research. *Nat Rev Cancer*, 10(7):523–523, 07 2010.

[34] Z. Guo, M. Maki, R. Ding, Y. Yang, B. Zhang, and L. Xiong. Genome-wide survey of tissue-specific microrna and transcription factor regulatory networks in 12 tissues. *Scientific Reports*, 4, 2014. cited By 36.

[35] Dong Yue, Hui Liu, and Yufei Huang. Survey of computational algorithms for microrna target prediction. *Current Genomics*, 10(7):478–492, 11 2009.

[36] Shiwei Duan, Shuangli Mi, Wei Zhang, and M Eileen Dolan. Comprehensive analysis of the impact of snps and cnvs on human micrornas and their regulatory genes. *RNA biology*, 6(4):412–425, Sep–Oct 2009.

[37] Lili Cui, You Li, Guoda Ma, Yan Wang, Yujie Cai, Shengyuan Liu, Yanyan Chen, Jia Li, Yuliu Xie, Gen Liu, Bin Zhao, and Keshen Li. A functional polymorphism in the promoter region of microrna-146a is associated with the risk of alzheimer disease and the rate of cognitive decline in patients. *PLOS ONE*, 9(2):1–8, 02 2014.