ETH *zürich*　　　　　　　　　**D** BIOL

Master Thesis

miRNA target enrichment benchmark and application to neuronal differentiation identifies miRNAs putatively involved in maturation

Departement of Systems Biology / D-BIOL, ETH Zürich
Laboratory of Systems Neuroscience, Schratt group

Dr. Germain, P.L., Laboratory of Systems Neuroscience, ETH Zürich, Main Supervisor
Prof. Dr. Schratt, G., Laboratory of Systems Neuroscience, ETH Zürich, Supervisor and Referee
Prof. Dr. Ciaudo, C., Inst. f. Molecular Health Sciences, ETH Zürich, Referee

Submitted by:

Tomás Germade

19.06.2020

Contents

ETH Zürich, Laboratory of Systems Neuroscience

miRNA target enrichment benchmark and application to neuronal differentiation identifies miRNAs putatively involved in maturation

Tomás Germade
Supervised by: Pierre-Luc Germain, Gerhard Schratt
Contributions from: Pierre-Luc Germain, Michael Soutschek, Reetu Ramesh Daswani, Helena Caria Martins

Abstract

MiRNAs can regulate a large host of target mRNAs, each member of which in turn can be regulated by many different miRNAs. Due to the complexity and scope of this network it is often difficult to efficiently validate the functional role of miRNAs in different cellular contexts. Computational prediction of miRNA targets has proven invaluable in this undertaking. One of these approaches aims to identify co-regulatory expression patterns in pre-defined sets of putative miRNA targets. Here we examined one such tool called enrichMiR which incorporates a varied set of statistical tests. We performed a systematic benchmark in order to assess the predictive accuracy achieved by these tests and by supplementary procedures. We found that a simple Fisher's test combined with a tissue-specific list of miRNAs performed best. EnrichMiR was thus applied on a novel Ngn2-mediated neurodifferentiation and maturation dataset and identified two miRNAs (miR-153 and miR-181) implicated in neuronal processes. In a secondary analysis of this dataset, independent of enrichMiR, we discovered circRNAs that based on their expression patterns might likewise have a role in neuronal maturation.

Introduction

miRNAs

The study of larva development in the nematode *Caenorhabditis elegans* resulted in the discovery of two small RNA transcripts of about 22 nt length and capable of gene downregulation. The sequence of those RNAs was found to be complementary to 3'UTR regions of putative targets. Named lin-4 (R. C. Lee, Feinbaum, and Ambros 1993) and let-7 (Reinhart et al. 2000), these small RNAs constituted the first identified microRNAs (miRNAs) (Bartel 2018).

Biogenesis

Canonical miRNA biogenesis involves transcription of a so-called pri-miRNA and two subsequent sequential endonuclease-mediated cleavage events. Each pri-miRNA is composed of a hairpin loop structure that is first cleaved by the Drosha-DGCR8 microprocessor. The resulting ~60 nt long pre-miRNA hairpin is subsequently transported from the nucleus to the cytoplasm by Exportin-5. In the cytoplasm the second cleavage takes place, mediated by the Dicer endonuclease. This releases the loop structure to yield a mature 21-25 nt miRNA duplex.

Non-canonical biogenesis on the other hand mostly bypasses Drosha cleavage, as observed in spliceosome-mediated mirtron generation, the formation of endogenous shRNAs and other chimeric variants (O'Brien et al. 2018).

Inhibition

After being generated, one of the duplex strands strands will be loaded onto an Argonaute (AGO) protein and incorporated into the RNA-induced silencing complex (RISC). Which strand gets loaded depends on the
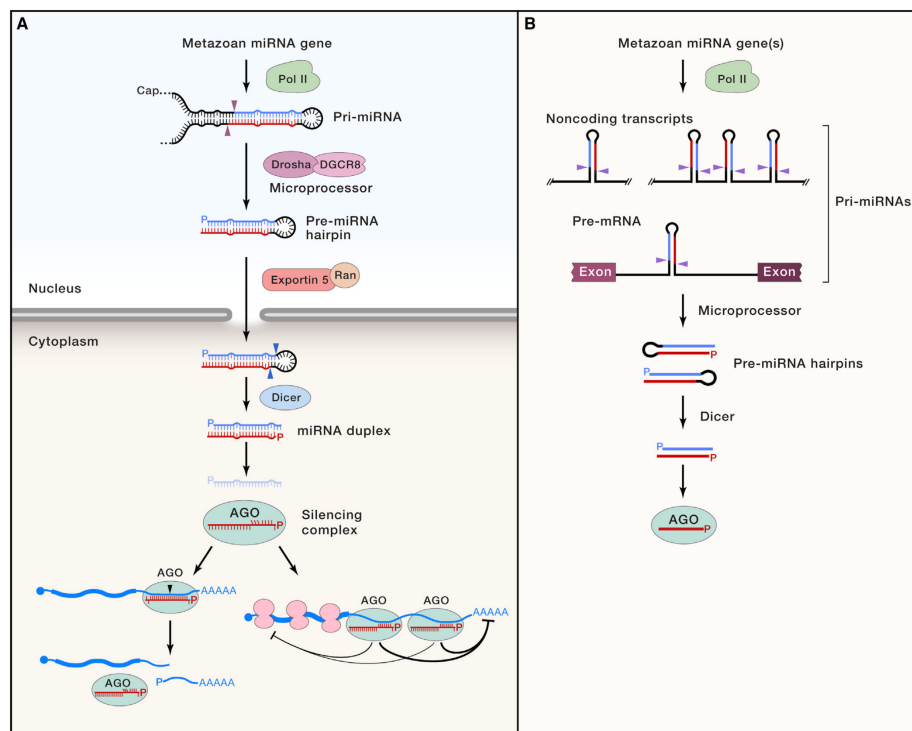
Figure 1: Schematic of canonical miRNA biosynthesis and inhibitory action in animals. (A) Pri-miRNA are first cleaved inside the nucleus by Drosha-DGCR8. The resulting pre-miRNA hairpins are exported into the cytosol where they are cleaved again by Dicer, giving rise to the mature 22 nt miRNA duplex. One of the strands gets loaded into AGO where it serves to target mRNAs based on binding affinity. These strands are referred to as guide strands. Subsequent miRNA-mediated suppression of transcripts can either occur through inhibition of the translation machinery and deadenylation induced mRNA decay or through mRNA slicing. The latter case is rare and requires extensive complementary binding. (B) Pri-miRNA can be found in introns of coding genes, in which case the miRNA in question is often co-transcribed with its host gene, as well as non-coding regions of the genome.

binding affinity of its 5' terminus with the AGO pocket. During the loading process the so-called passenger strand gets expulsed, while the so-called guide strand remains (Schwarz et al. 2003). Such AGO bound miRNAs were shown to be highly stable (Bail et al. 2010) (Guo et al. 2015), revealing half-lives up to 25 hours in certain cells (Kingston and Bartel 2019). A notable exception to this was found in neurons which, depending on their activity levels, showed high rates of miRNA decay (Krol et al. 2010).

MiRNAs bind to mRNAs based on the extent of their sequence complementarity. Depending on the binding affinity, the RISC will induce repression of the mRNA target. This repression can happen in two distinct ways: 1) Extensive sequence complementarity induces slicing of a target (Llave et al. 2002) (Valencia-Sanchez et al. 2006). This process is common in plants and siRNA-mediated repression but rare in animals. 2) The RISC prompts the recruitment of deadenylases which destabilize the target and inhibits translation (Wu, J. Fan, and Belasco 2006). An exception to this mode of action has been found in cells under cell cycle arrest, where miRNAs were implicated in target upregulation (Vasudevan 2012).

Binding

In animals, miRNA binding to target sites involves the so-called miRNA seed sequence, situated in nucleotides 2-7 of its 5' region. This sequence is often extended to nucleotides 2-8. Canonical binding involves either partial or complete pairing of a miRNA seed region to complementary sites on a mRNA target. Functional pairings comprise 8mer, 7mer and 6mer sites, in decreasing order of binding affinity and repression efficiency. Each of these site types features subtypes depending on small variances or additions (Bartel 2018). One example involves the presence of an Adenine (A) on the site across position 1 of the miRNA (at the 5' terminus) (Lewis, Burge, and Bartel 2005). This configuration was found to enhance binding affinities regardless of the specific miRNAs involved. The reason was found to be its binding to AGO (Ma et al. 2005) (Schirle et al. 2020). A second example represents an additional pairing of nucleotides 13-16 in the 3' region of certain miRNAs. This resulted in relatively weak supplementary binding (Grimson et al. 2007) that makes it possible for some miRNAs-mRNA pairs to functionally interact despite seed mismatches (McGeary, Lin, Shi, Bisaria, et al. 2018).
Because of this strong functional dependency on seed-pairing, different miRNA variants sharing the same seed sequence show similar target interaction and are thus are classified as miRNA families. These families don't necessarily share common ancestry but it is thought that they often do (Bartel 2018).

While plant miRNAs tend to bind to the coding region (CDR) of their mRNA targets, in animals most functional binding sites have been identified in 3'UTR regions (Millar and Waterhouse 2005). This being said, many functional interactions have been found in 5'UTR and CDR regions in both flies and different vertebrates (A. Stark et al. 2007) (Schnall-Levin et al. 2010) (Lewis, Burge, and Bartel 2005Lewis, 2005) (Agarwal et al. 2015Agarwal, 2015). A somewhat special case involves so-called miRNA recognition elements (MREs), binding sites exclusive to CDRs bound by extensive 3' base pairing (K. Zhang et al. 2018).

Repression dynamics

The strength of gene suppression conducted by a single microRNA is moderate (Baek et al. 2008). Yet miRNAs have been described to experience a switch-like change in repression efficiency depending on relative target abundances. High repression occurs at low target levels and vice versa (Arvey et al. 2010). The sharpness of such a switch in repression efficiency depends on both interaction strength and relative abundances (Mukherji et al. 2011). The introduction of sponges, able to sequester and silence miRNAs, of either exogenous (Ebert, Neilson, and Phillip A Sharp 2007) or endogenous origin (Thomas B. Hansen et al. 2013), as well as potential compartmentalization can further re-calibrate the ratio of active miRNA and respective targets.

Combined suppression of mRNAs by multiple miRNAs was found to affect suppression strength. Depending on the distance between binding sites, repression can be competitive or cooperative and thus excel individual effects (Doench and Phillip A. Sharp 2004) (Grimson et al. 2007). The same additive repression capabilities have been observed for binding sites outside the 3'UTR in flies (Schnall-Levin et al. 2010).
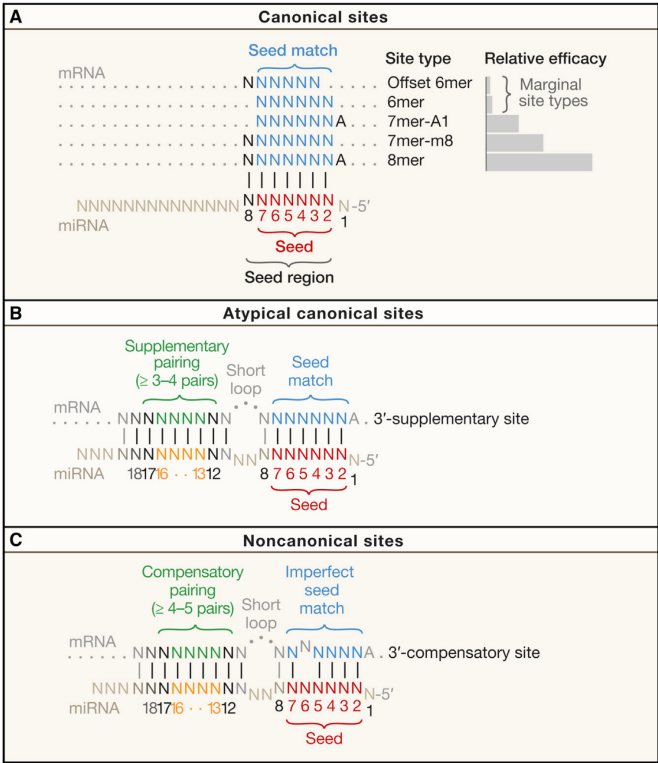
Figure 2: Different functional miRNA-mRNA interactions can be categorized depending on their degree of sequence complementarity and repression strength. (A) Canonical miRNA binding sites involve the so-called seed region at the 5' terminus of a miRNA. It has been found that nucleotides adjacent to the seed can similarly affect binding. An A opposite the first miRNA nucleotide has been shown to interact with the AGO component itself. Usually, the stronger the binding between the transcripts the higher the suppression efficiency. We are thus able to rank the different binding types as follows: 8mer > 7mer-m8 > 7mer-A1 > 6mer > Offset 6mer. (B) It has been discovered that a short region in the 3' terminus of the miRNA can contribute to the binding. (C) These effects are usually weak but can in some cases enable a discernible functional effect of miRNAs with imperfect seed matching.

Experimental investigation

The verification of these complex interactions and confirmation of their regulatory action can be addressed with experimental means. Direct miRNA-mRNA interaction may be studied using techniques such as AGO-Clip (Argonaute-crosslinking and immunoprecipitation) (Chi et al. 2009) (Grosswendt et al. 2014) and CLASH (crosslinking, ligation, and sequencing of hybrids) (Kudla et al. 2011). The effects of a miRNA on its targets may be studied using luciferase reporter assays (Nicolas 2011) or RNA-seq after perturbation of miRNA levels, e.g. via knock-out, overexpression or antagomirs. Any evidence produced by the latter may be corroborated by the system's reversion to an original state upon neutralizing the perturbation, e.g. by a rescue or the manipulation of UTR binding sites via the introduction of mutations, knock-out and overexpression.

Target prediction

A way to overcome these limitations, are bioinformatic predictions to narrow experimental queries to fewer candidates. This can be achieved by scanning miRNAs and mRNAs for features that have previously been experimentally associated with functional effects, such as: seed sequence complementarity, binding thermodynamics, UTR accessibility, target site conservation and target site abundance, just to name a few.

Another possibility is to train Machine Learning (ML) algorithms on datasets harbouring experimentally confirmed miRNA-mRNA relationships. This can be done either in a supervised fashion, when combined with a pre-defined feature selection as listed above, or unsupervised fashion, when features are discovered independently, based on raw data. The latter can also be used for the discovery of features that can then be applied to supervised methods (Grimson et al. 2007) (Agarwal et al. 2015). Depending on the quality and quantity of training datasets such methods have been shown to excel the predictive power of some experimental approaches (Agarwal et al. 2015). In order for such ML methods to extrapolate generalizable patterns, specific feature combinations need to be sufficiently abundant in their respective training datasets.

There have also been cases where results based on multiple differing approaches have been combined in order to increase overall prediction inference (Oliveira et al. 2017). For an up-to-date review on different machine learning strategies as well as existing target prediction tools and databases see Schäfer and Ciaudo 2020.

TargetScan

One example of a miRNA binding site prediction tool is TargetScan (current version: 7.2). This tool was first published in 2003 (Lewis, Shih, et al. 2003) and received its last major update in 2015 (Agarwal et al. 2015).

TargetScan is an example for the initial application of an unsupervised ML algorithm to select features. These are subsequently applied to train a supervised ML algorithm in order to generate a predictive model. The selected features were partly based on features derived from predecessor models and partly based on features identified by combined step-wise linear regression and successive Akaike Information Criterion (AIC) evaluation. The training data constituted 74 microarrays of single-miRNA HeLa cell transfections. The resulting model is composed of 14 features, including minimum distance from 3'UTR ends, local AU composition, 3' supplementary pairing, predicted seed-pairing stability, target-site abundance, 3'UTR length, ORF length, predicted structural accessibility, the number of offset-6mer sites in the 3'UTR and 8mer sites in the ORF, the nucleotide identity of position 8 of the target, the nucleotide identity of positions 1 and 8 of the miRNA, and site conservation. No interaction terms between given features was allowed, in order to limit model complexity. Additionally, based on findings that indicate unsignificant repression of non-canonical binding sites, only 3'UTR seed-pairing was considered (Agarwal et al. 2015). Binding site predictions were verified using various test datasets. TargetScan was able to outperform all other established tools at the time (Agarwal et al. 2015). Lately, methods such as MiRTarget (v4.0) (W. Liu and Irudayaraj 2019), deepTarget (B. Lee et al. 2016) and a biochemical model generated by McGeary et al. McGeary, Lin, Shi, Pham, et al. 2019 based on dissociation rates (KD) as well as approaches that combine various tools were able to outperform TargetScan in terms of explained dataset variance (Schäfer and Ciaudo 2020).

Target enrichment tools

A different approach for the identification of actively regulating miRNA candidates is employed by miRNA target enrichment tools. Based on mRNA expression profiles and pre-defined miRNA target annotations, various statistical tests can be applied to detect co-regulatory patterns in target sets. This serves to identify miRNAs possibly involved in the formation of such patterns. Several methods have been implemented in the past, each of them implementing different types of analysis.

FAME constructs biparite graphs where miRNAs and targets are separated into different sets and are connected via edges that represent TargetScan-derived context scores. It then generates random degree-preserving permutations of that graph to assess how likely it is that the topology found in the initial network is inherent in the network structure (Ulitsky, Laurent, and Shamir 2010). miTEA employs a so-called mmHG statistic that assesses similarity between the top of two ranked lists, e.g. targets of a given miRNA and a list of genes representing an expression profile. Subsequently, the likelihood of the observed element densities at their positions is calculated, assuming random distribution (Steinfeld et al. 2013). Other tools involve tests such as gene set enrichment analysis (GSEA) and over-representation analysis (ORA) (Backes et al. 2016), hypergeometric tests (Y. Fan et al. 2016) (Dongen, Abreu-Goodger, and Enright 2008) (Vlachos et al. 2016) and hierarchical gene co-expression signature enrichments (Gu and Xuan 2013). There is currently no clear consensus as to the best way of implementing enrichment statistics.

Project scope

In this project we implemented an in-house miRNA target enrichment tool called enrichMiR. It incorporates various statistical tests in order to examine a possible co-regulation of miRNA target sets in differentially expressed mRNA profiles. To evaluate the inference accuracy of the different tests and input options, we conducted a systematic benchmark on 5 datasets constituting 33 miRNA perturbation experiments. A simple Fisher's test was found to perform best, demonstrating consistently high metric scores. Moreover, tissue-specific restriction of miRNA resulted in an additional improvement. Having selected the most beneficial measures, we applied enrichMiR on a novel Ngn2-mediated neurodifferentiation and maturation dataset generated in our lab. We were thus able to identify several miRNA candidates, two of which (miR-153 and miR-210) were confirmed to be involved in neuronal processes.
Further investigation of transcriptional dynamics in the Ngn2 neurodifferentiation dataset resulted in the identification of a set of circRNAs possibly implicated in neuronal development.

Results

A novel tool for miRNA target enrichment analysis

Our aim was to examine transcriptome profiles in order to discover miRNAs whose regulatory function could impact neuronal development in humans. For this purpose, we chose to utilize an in-house computational tool named enrichMiR. EnrichMiR has the capability to run various statistical tests on differential mRNA expression signatures to infer miRNAs that either 1) cause said differential signatures or 2) would be capable of doing so. To achieve this, enrichMiR depends on the input of several objects: a differential mRNA expression profile, a pre-defined library of individual miRNAs and their putative targets, and optionally a species-specific list of expressed miRNAs. The statistical methods incorporated into enrichMiR at the time of this report range from Fisher's tests over cumulative distribution tests to linear regressions, each of them composed of diverse implementations. Their particular choice and procedural implementation predated the current project and will not be discussed in depth here. Nevertheless, a brief overview can be found under in Tables 2 and 3, where each test is defined by the inputs it requires as well as a brief description of the underlying approach. The information a given test extracts from a supplied DEA is either binary (up- or downregulation) or continuous (differential expressions). The information extracted from the miRNA target annotation can either entail binary target gene sets, non-binary binding site numbers or TS derived context scores.

Table 1: The benchmarking datasets have been selected to harbour both upregulated (transfection, TF) and down-regulated (knock-out, KO) effects. Each of them derives from a different tissue. The Bartel datasets contribute the largest number of experiments. The Number of DE (differentially expressed) genes contains the numbers of both unseparated and exon-specific profiles (in brackets). We applied a threshold of FDR < 0.05. More details on each of the datasets and DEAs can be found under Materials & Methods.

	Experimental design	Organism	Celltypes	Nr. of experiments	Nr. of DE genes	Perturbation time [d]
Bartel	TF	Human	HeLa	16	1629 (193)	1
			HEK293	12	482 (101)	1
Rat	TF	Rat	hippocampal neurons/glia	2	21 (0)	5
Jeong	KO	Human	SNU-638	1	6473 (277)	several
Amin	KO	Mouse	CAD	1	3741 (297)	4 (E12.5)
Cherone	KO	Mouse	iN	1	7726 (1561)	0
					5525 (1003)	1

A benchmark of the various miRNA target enrichment tests

To test the prediction accuracy of each statistical analysis method, we chose to conduct a systematic benchmark based on datasets with known miRNA perturbations. We assessed if we could successfully predict the correct affected miRNA family and which method produced the most accurate prediction. In addition, we tested two measures that could potentially increase inference accuracy of each method. First, enrichMiR can optionally be supplied with a tissue-specific miRNA expression profile in order to restrict the otherwise species-specific set of miRNAs to be tested. Secondly, we supplied each method with differential expression (DE) profiles specific to reads from processed transcripts which aren't visible in intronic/unspliced reads. This, so we argued, would partially filter out the downstream cellular response to miRNA perturbation and allow for the direct effect of miRNAs to be more distinct (Zeisel et al. 2011) (Gaidatzis et al. 2015) (see Materials & Methods).

Based on the performance of the methods, we created three different combinations, each comprised of three individual methods. They were generated by p-value aggregation, once using the Fisher and once the geometric mean approach. They are included in the following benchmark.

Table 2: Individual statistical tests contained in enrichMiR. They were measured against each other in our benchmark. Signal type refers to how DEA results are incorporated into the respective tests: 'set' signifies a binomial variable (either up-/downregulated or not) while 'continuous' signifies the incorporation of expression data. Annotation input refers to the information that is utilized from the supplied list annotation: 'genesets' means that transcripts with at least 1 putative binding site are considered, 'nb sites' stands for non-binary sites, meaning that each binding site is considered individually and 'scores' refers to the context scores e.g. provided by TargetScan.

Test	Signal type	Annotation input	Description
EN	set	genesets	ORA of target genes among set (e.g. downregulated)
wEN	set	nb sites	ORA of binding sites corrected for UTR length
michael	set	nb sites	ORA of binding sites
regmir	set	scores	Regularized constrained logistic regression on scores
regmirb	set	genesets	Regularized constrained logistic regression (binary)
aREAmir	continuous	scores	Score-weighted analytic rank enrichment analysis on signed log-FDRs
aREAmir2	continuous	scores	Modification of aREAmir correcting for pleiotropy
modScores	continuous	scores	Linear regression of logFCs on target prediction scores
modSites	continuous	nb sites	Linear regression of logFCs on number of binding sites
KS	continuous	genesets	Kolmogorov-Smirnov test on logFCs
KS2	continuous	genesets	Modification of KS-test

Test	Signal type	Annotation input	Description
MW	continuous	genesets	Mann-Whitney / Wilcoxon test on logFCs
GSEA	continuous	genesets	Gene set enrichment analysis (permutation test)

Table 3: Based on the initial performances of individual tests we aggregated them into combined tests.

Test	Combination	Aggregation method
combFish.1	michael, wEN and modScore	Fisher
combGeom.1	michael, wEN and modScore	Geometric mean
combFish.2	michael, aREAmir and KS2	Fisher
combGeom.2	michael, aREAmir and KS2	Geometric mean
combFish.3	michael, aREAmir and regmirb	Fisher
combGeom.3	michael, aREAmir and regmirb	Geometric mean

We selected five datasets, derived from different organisms and tissues, and including a total of 33 miRNA perturbations, have been selected to serve as benchmark datasets based on their single-miRNA perturbation structure. For the purpose of simplification, we refer to the different datasets as *Bartel*, *Rat*, *Jeong*, *Amin* and *Cherone* based on the authors of their respective publications. A more detailed summary of each dataset can be found in Table 1 and the Materials & Methods section.

EnrichMiR analysis depends on the supplementation of a list linking miRNA families (ie. miRNA sharing the same seed sequences) to their putative targets. The annotation we used for this purpose was acquired from TargetScan (TS) (see Materials & Methods), although the supplementation of alternative target annotations is equally serviceable. Individual miRNA targets were recorded on a per gene level. We further restricted annotations to include predicted 8mer and 7mer binding sites. With the aim of introducing varied degrees of noise to the datasets, we performed multiple enrichMiR runs per experiment, each supplied with a differently permuted TS annotation. For each experiment, the random permutations were limited to the set of target genes assigned to the perturbed miRNA particular to that experiment. Also, the randomly selected target genes were exchanged for non-target genes of the same annotation. The permutation percentages we chose were 0 (original), 20, 35 and 50. Each permutation was replicated 3 times. For instance, a miR-7 duplex transfection experiment derived from the Bartel dataset underwent 10 enrichMiR runs: once with the original annotation and three times each for the three different permutation degrees. Each permutation randomly exchanged miR-7 target genes with non-miR-7 target genes.

To assess the relative performance of each statistical analysis method integrated in enrichMiR, we used four different scoring metrics. Each of them depends on the definition of a true miRNA family (or seed sequence). Detailed information on true miRNA family assignments can be found under the Materials & Methods subsection. The first two scores are True-Positive (TP) and False-Positive (FP) numbers at FDR < 0.05. Since we only ever assigned one single true miRNA family per sample, the TP numbers coincide with the TP rate (TPR) per sample. For the same reason, we could use the number of FP directly as it is tightly linked to the false discovery rate ($FDR = FP/(TP + FP)$). The third score we termed detPPV (positive predictive value at detection level). It reflects the relative FDR based rank of the true to non-true miRNA families. Its maximum value is 1 (when the true family achieves the highest rank) and decreases with the number of non-true families that surpass the true family in rank. The detPPV score is useful in that it can highlight methods which produce good ranking of families but poorly calibrated estimates of significance. The fourth score was termed log10QDiff and conveys the $-log10(FDR)$ differences between the true family and the highest ranking non-true family, where the rankings again are based off of FDR values. This term can be negative if a non-true family achieves a higher rank than the true family.

Most tests appear relatively robust to ample noise in the annotation

The enrichMiR results for each experiment across the different permutation levels of the TS annotation (0, 20, 25 and 50 %) revealed a high robustness to noise. Across all statistical tests, detPPV scores displayed larger variance across the experiments than within, i.e. across their permutation percentages. Figure 3 shows a heatmap of the mean detPPV values across all analysis methods subdivided into the different TS permutations. None of the methods detected the true miRNA family when applied to experiments pertaining to the *Jeong* and *Cherone* datasets, as deducible from Figures 3, 5 and 6, each highlighting different scores.
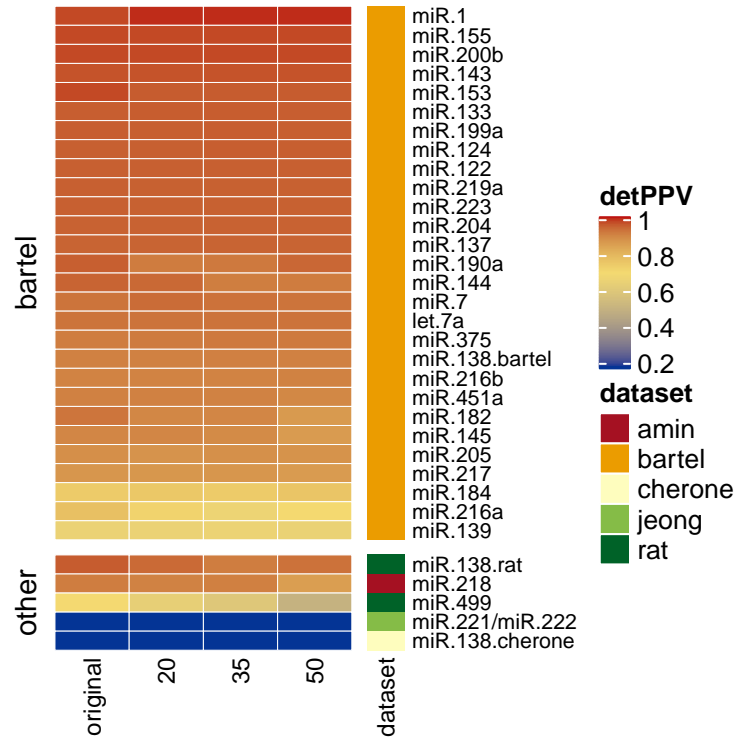

Figure 3: Heatmap of detPPV scores across experiments and miRNA target list permutations. One of the scores we used to evaluate the performance of the enrichMiR tests we called detPPV. This score represents the relative rank of the true miRNA family relative to non-true families. Its values range from 0 to 1. With few exceptions the enrichMiR scores for each sample either slightly decreased with each level of permutation or stayed constant. The score differences between the treatments were more striking, revealing their degree of difficulty regarding enrichMiR analysis. In the Jeong and Cherone datasets detection of the true family largely failed.

High robustness to noise introduced by TS permutation was also observable across the statistical methods. Here again, score variances were higher across methods than permutations. Figure 4 gives an idea of the detPPV behaviours of each applied enrichMiR method when separated into scores derived from *Bartel* and all other datasets.

A simple Fisher-based method performs best

Methods such as *wEN*, *EN*, *combFish.1*, *combGeom.1* and *michael* showed high detPPV scores, reflecting consistently high true miRNA family rankings. All of them except for *michael* simultaneously displayed high numbers of FP results relative to other methods. Even though these methods were highly sensitive they were also relatively unspecific. The *michael* method on the other hand proved to be both sensitive and specific: the number of reported FP miRNA families was relatively low while true miRNA family detection
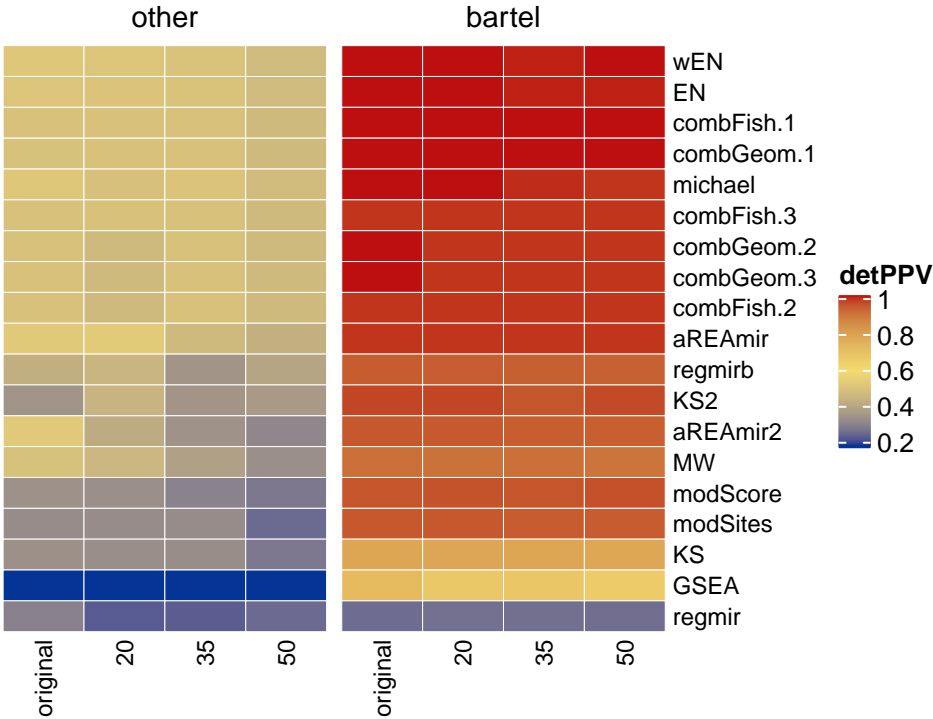
Figure 4: Heatmap of detPPV scores across methods and miRNA target list permutations. The samples are split up according to their membership to either Bartel or other datasets. The displayed scores are mean detPPV values across samples. In all cases except for regmir the performance of enrichMiR analysis methods was higher for samples stemming from Bartel. In all but a few cases the scores either gradually declined or remained constant with increased degrees of miRNA target list permutation (x-axis). See Tables 1 and 2 for summaries of the enrichMiR methods.

consistently high. In the *Jeong* dataset it was the only method to ever detect the the correct miRNA family, partially doing so even with the introduction of additional noise. Figure 5 displays the TPR and FP results for each method subdivided into the different datasets. Here, the combined methods 2 and 3 (*combFish & combGeom*) as well as *KS2* tended to achieve relatively high sensitivity and specificity with the latter being affected by the noise introduction to a relatively high degree.
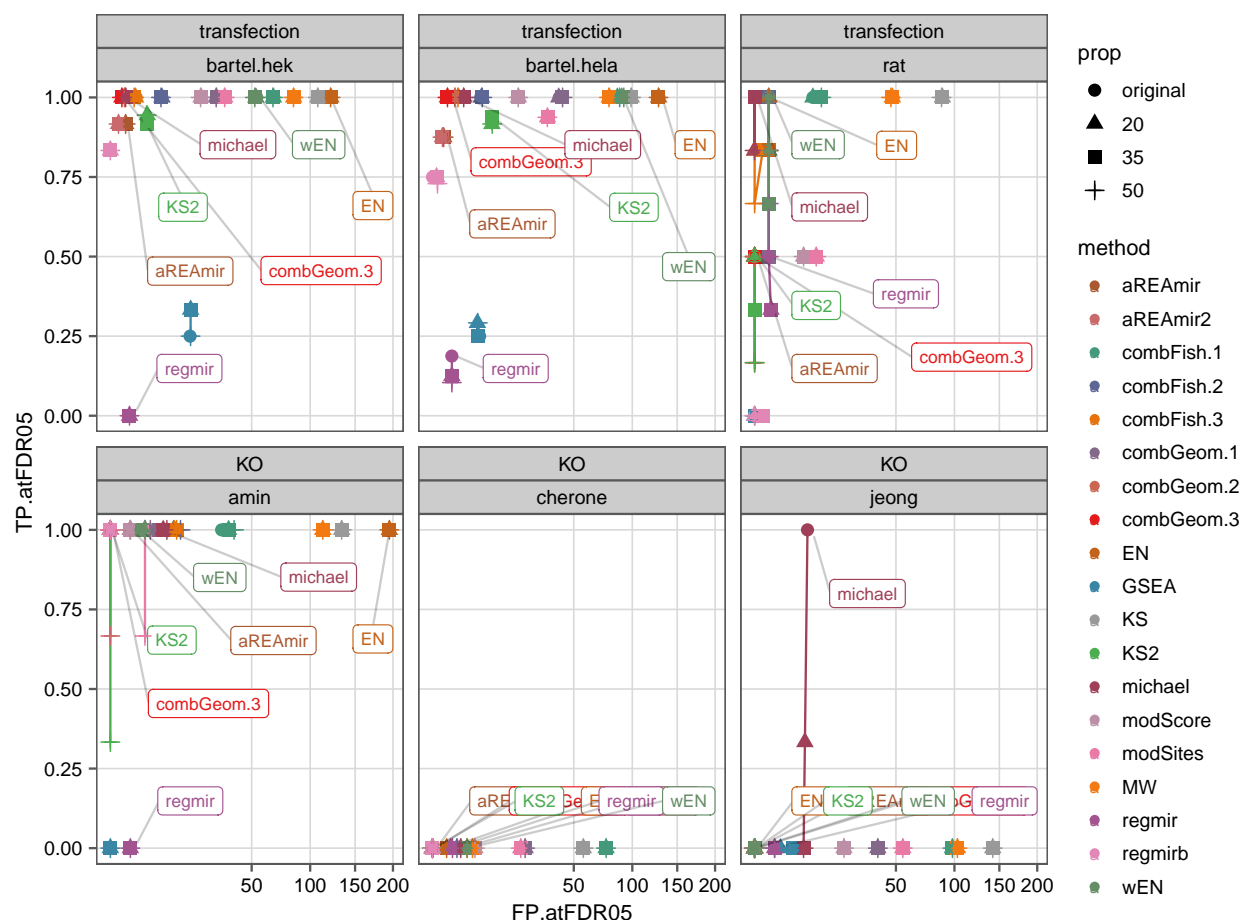

Figure 5: Comparing False-Positive (FP) and True-Positive (TP) discoveries for each enrichMiR method by their experimental design. The different shapes of the datapoints reflect the mean values achieved under the different levels of miRNA target list permutation we performed. True family detection failed for all tests in both the Cherone and the Jeong datasets, with a notable exception being the michael method.

The differences between the FDR values assigned to the true miRNA family and the highest ranking non-true miRNA family can be observed in Figure 6. Small differences were obtained from *Jeong*, *Cherone* and *Rat* datasets. In the *Bartel* and *Amin* datasets the *combFish* methods achieved high scores, together with the *EN*, *wEN* and *modScore* methods. The only method that consistently produced negative log10QDiff scores was *regmir*, meaning that it rarely ranked the true miRNA family at the highest position.

Tissue-specific restriction of miRNAs improves inferences

Apart from comparing the scores between the different statistical tests we also set out to investigate the impact of restricting the set of miRNAs for which target enrichment should be considered. We ran all enrichMiR analyses once without restriction, in which case we supplied an extensive TargetScan-derived collection of
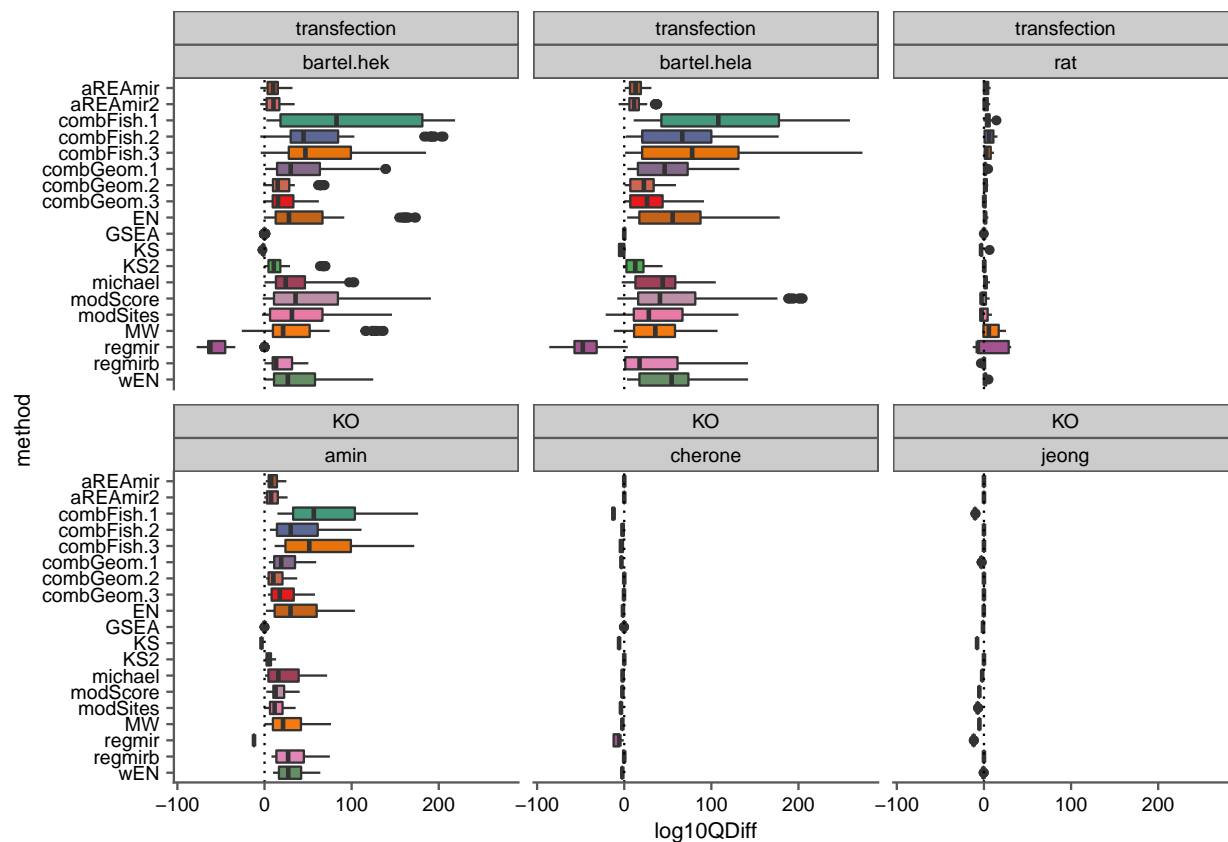
Figure 6: Comparison of log10QDiff scores between the different enrichMiR methods for each dataset. The log10QDiff score represents the distance between the true miRNA family and the highest ranking non-true family in terms of FDR value. The Cherone and Jeong datasets again didn't perform well, with many FDR values being equal to 1. The Rat dataset showed similarly low distances while still achieving high true family discovery rates (see Figure 5). In all other datasets the highest log10QDiff scores were recorded among the combined methods, in particular the ones combined using the Fisher method.

species-specific (but not tissue-specific) miRNA, and once with restriction, in which case we supplied a list of miRNAs expressed in each respective tissue. The origin of these resources is described in the Materials & Methods section.For this test the *Jeong* dataset was excluded due to the lack of tissue-specific miRNA expression data.

Application of tissue-specific miRNA profiles improved the scores (see Figure 7). The TPR was increased significantly (p-value of 1.4e-06) and the number of FPs reduced highly significantly (p-value of 5.4e-181) over all datasets. When examining TPR scores for each dataset separately, we found no significant effect for datasets *Rat* (p-value of 0.3), *Amin* (p-value of 0.4) and *Cherone* (p-value of 1). The same analysis for FP scores revealed a highly significant effect for all datasets except *Rat* (p-value of 0.3) (see Figure 14). The effect of this measure on a per sample basis and resolved by permutation level is illustrated in Figure 7.
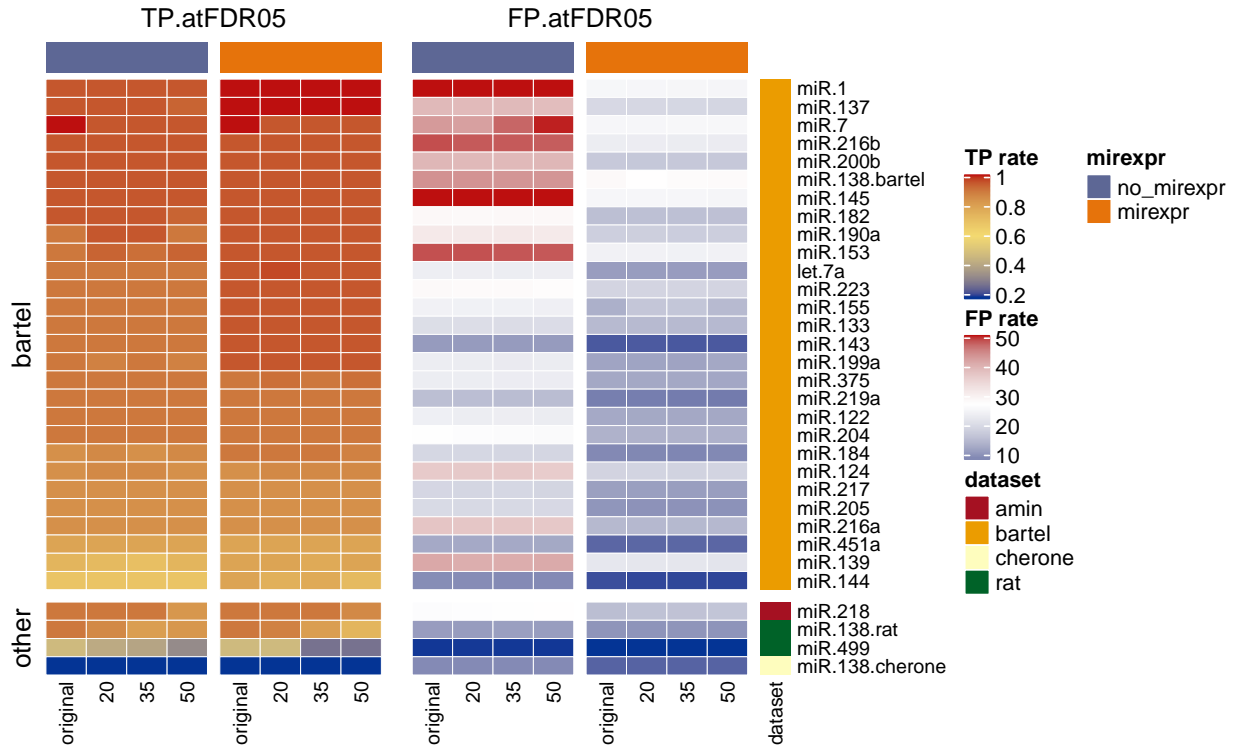

Figure 7: Heatmaps showing the difference of True-Positive rate (TPR) and False-Positives (FPs) across treatments with and without supplementation of miRNA expression (mirexpr) profiles to enrichMiR. In all cases we either saw an increase in TPR or no change at all when comparing enrichMiR analyses performed without and with mirexpr profiles. The same comparison revealed large decreases in FPs in an overwhelming majority of cases. Both, increases in TPR and decreases in FPs were significant.

Exon-specific DEA inputs increase test specificity and decrease sensitivity

In a next step, we aimed to investigate the effect of running enrichMiR on exon-specific mRNA DEA profiles. The reasoning for this was to partly filter out indirect miRNA perturbation effects inside the cell such as compensatory regulation of mRNA transcription. We would thus exclude the differential effects on unspliced immature transcripts and focus on mature mRNA instead (see Materials & Methods).

Substitution of unseparated DE profiles with exonic transcript-derived profiles resulted in a highly significant decrease in both TPR (p-value of 1.0e-171) and FP (p-value of 1.7e-187). Separate examination per dataset revealed a significant decrease in TPR for all datasets except *Jeong* (p-value of 0.2) and *Cherone* (p-value of 1). The same analysis for FP scores revealed a highly significant decrease for all datasets (see Figure 15).

ETH Zürich, Laboratory of Systems Neuroscience

The effect of DE profile substitution on a per sample basis and resolved by permutation level is illustrated in Figure 9.

A new differentiation protocol yields high-purity, physiologically mature neurons in vitro

Traditional Ngn2-mediated neuronal differentiation protocols yield neurons in vitro, but these require a time investment of several weeks to mature. This has hampered the reproducible establishment of mature neuronal features such as synaptic spines and activity within a manageable timeframe. The inclusion of astrocytes to the cultures has moreover limited the applicability of -omics methods, since signals derived from these larger cells would mix and possibly screen those derived from neurons. To counteract these issues, Michael Soutschek from the Schratt lab combined existing methods in order to create a novel Ngn2-based neurodifferentiation protocol capable of producing rapidly-maturing, high-purity neurons in the absence of astrocytes.
This novel protocol was applied to human induced Pluripotent Stem Cells (hiPSC) allowing for RNA-seq and proteomic profiling at different intervals of neuron maturation. The measurements were taken at days 0, 9, 15, 21, 27, 33, 40. The resulting transciptome datasets include small and long RNA-seq profiles. All described steps were concluded before this project was initiated.

Transcriptomic profiling of Ngn2 differentiation pinpoints the dynamics of in vitro neuralization and maturation

The Ngn2 differentiation dataset was comprised of 3 replicates for each recorded timepoint. A Principal Component Analysis (PCA) not only revealed a clear trajectory of the different timepoints correlating to neuronal maturation trends but also a systematic deviation of replicate 1 samples (see Figure 10). We were able to observe a rapid shut down of markers for pluripotency as well as an activation of markers for neural identity over time (see Figure 11).
We thus performed different DEAs: 1) over each consecutive pair of recorded timepoints, 2) of each timepoint relative to the undifferentiated stem cell state (day 0) and 3) of each timepoint relative to the first recorded timepoint after induction of neuralization (day 9). In each DEA we accounted for the deviations observed in replicates 1 ($\sim rep1 + stage$ vs. $\sim rep1$). The respective numbers of differentially expressed genes for each DEA can be viewed under Table 5.

Additionally, we extracted circular RNA (circRNA) structures from our long RNA-seq read alignments. We applied the circRNA quantification tools of DCC[1] (v0.4.7) (Cheng, Metge, and Dieterich 2016) and CIRCexplorer2[2] (v2.3.8) (X.-O. Zhang et al. 2016). Both quantifications showed a high degree of overlap, detecting transcript numbers of 1161 and 1025 respectively when excluding putative intergenic circRNAs and after filtering out transcripts lacking at least 10 reads in at least 3 samples. All following procedures were based on the quantification received from CIRCexplorer2. We then aimed to detect circRNAs that showed a different expression pattern across timepoints to their respective host genes. To this end we combined relative expression patterns of independently smoothed linear transcripts and corresponding circRNAs and applied K Nearest Neighbour (KNN) based pattern clustering. This way we were able to identify a set of circRNA that suggested independent regulation to their host genes and serve as candidates for functional roles in neurodevelopment. All clustering results can be viewed under Figure 16. Results obtained over neuronal stages (excluding day 0) can be viewed under Figure 12. All results were obtained after applying a log fold change (logFC) threshold of 1 and FDR threshold of 0.05.

---

[1]https://github.com/dieterich-lab/DCC
[2]https://github.com/YangLab/CIRCexplorer2
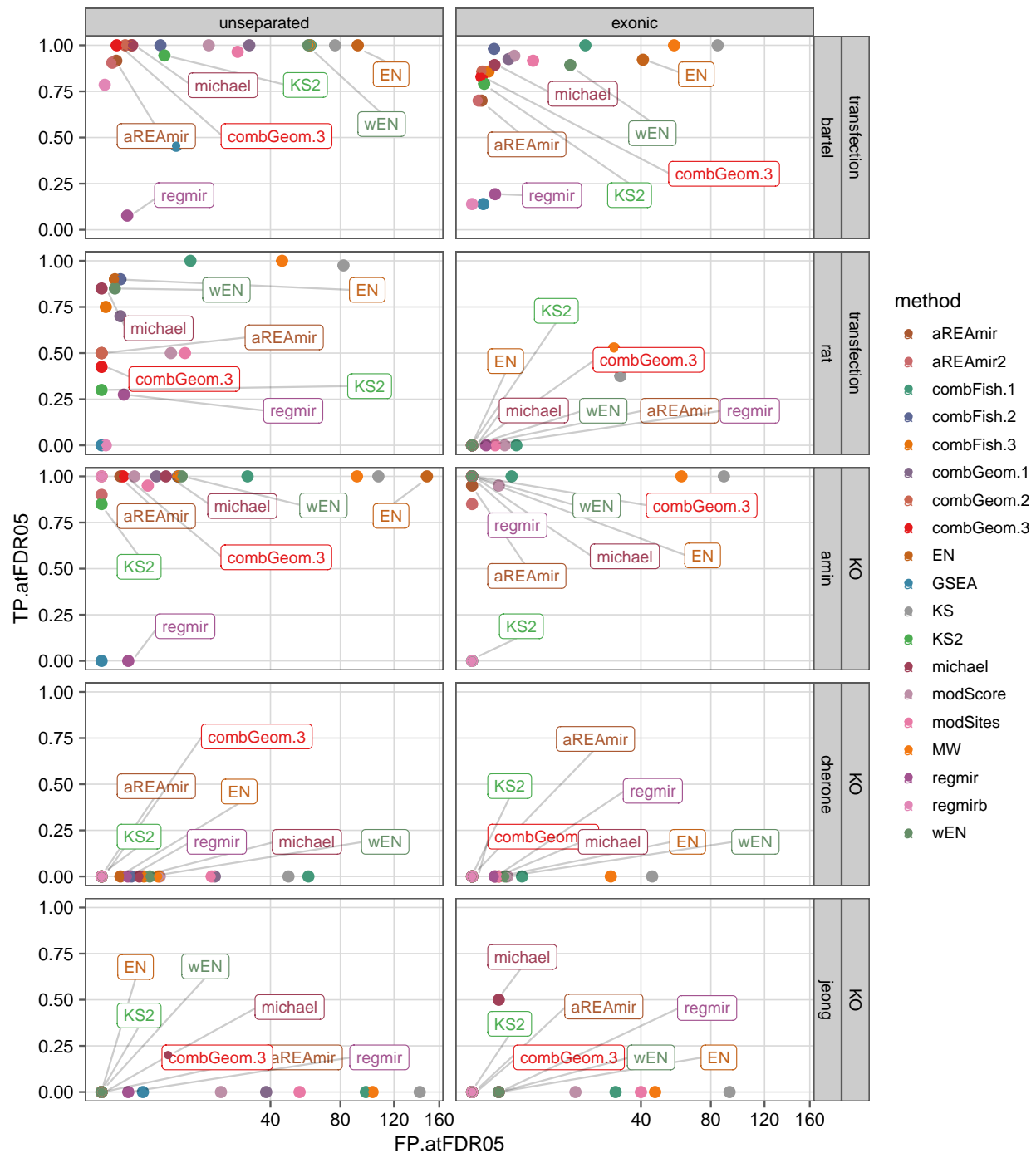
Figure 8: Comparison of FPs and TPR for each enrichMiR method by dataset and exon specificity of DEA profiles. The datapoints represent score means across TS permutations. Across all datasets the TP rates achieved by the individual methods significantly decreased. Notable exceptions could be seen in the Amin dataset where regmir experienced a large increase and the Jeong dataset where michael more moderately increased as well. In the Rat dataset we observed a complete failure of any method except MW and KS to detect any TPs. At the same time the overall FPs decreased significantly as well.
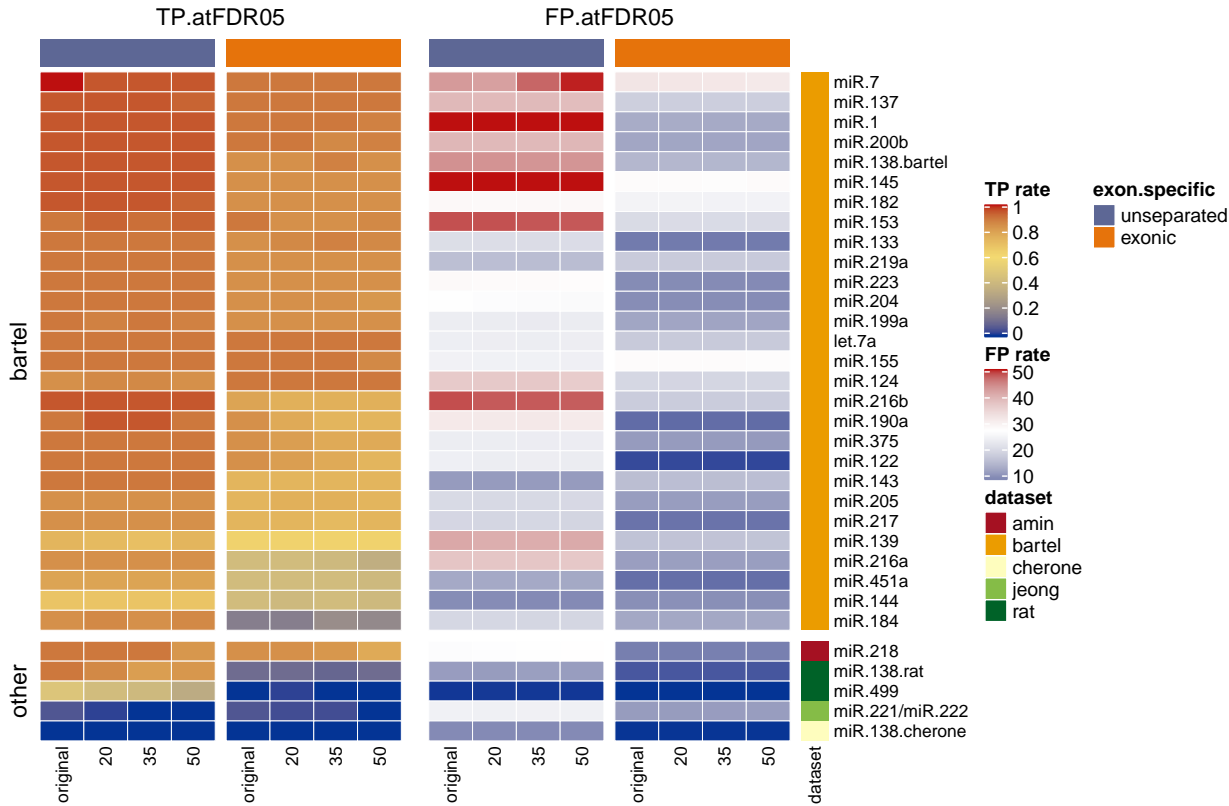
Figure 9: Heatmaps showing TPR and FPs across experiments upon running ernrichMiR with and without exon-specific DEA profiles. The values of each treatment are subdivided into the different levels of miRNA target list permutations. Across all treatments we observed a significant decrease in TPR and FPs.
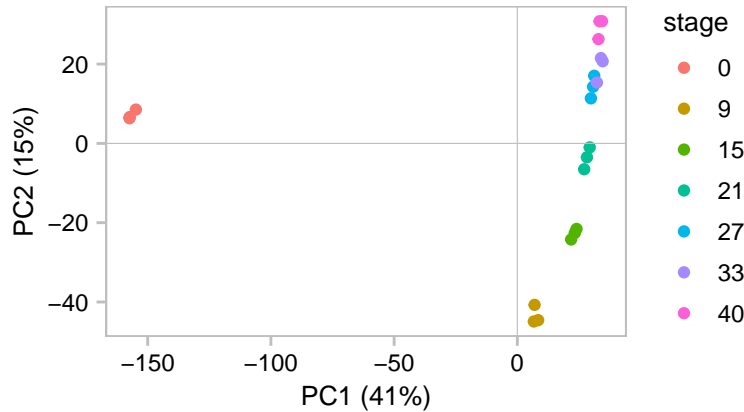
Figure 10: A Principal Component Analysis (PCA) over the different stages and replicates of our Ngn2-mediated neurodifferentiation dataset reveals a large initial change in expression patterns when neuralizing and a subsequent trajectory of further maturation. Each PC captures one trajectory almost fully, with the initial neuralization step representing the most dominant component.
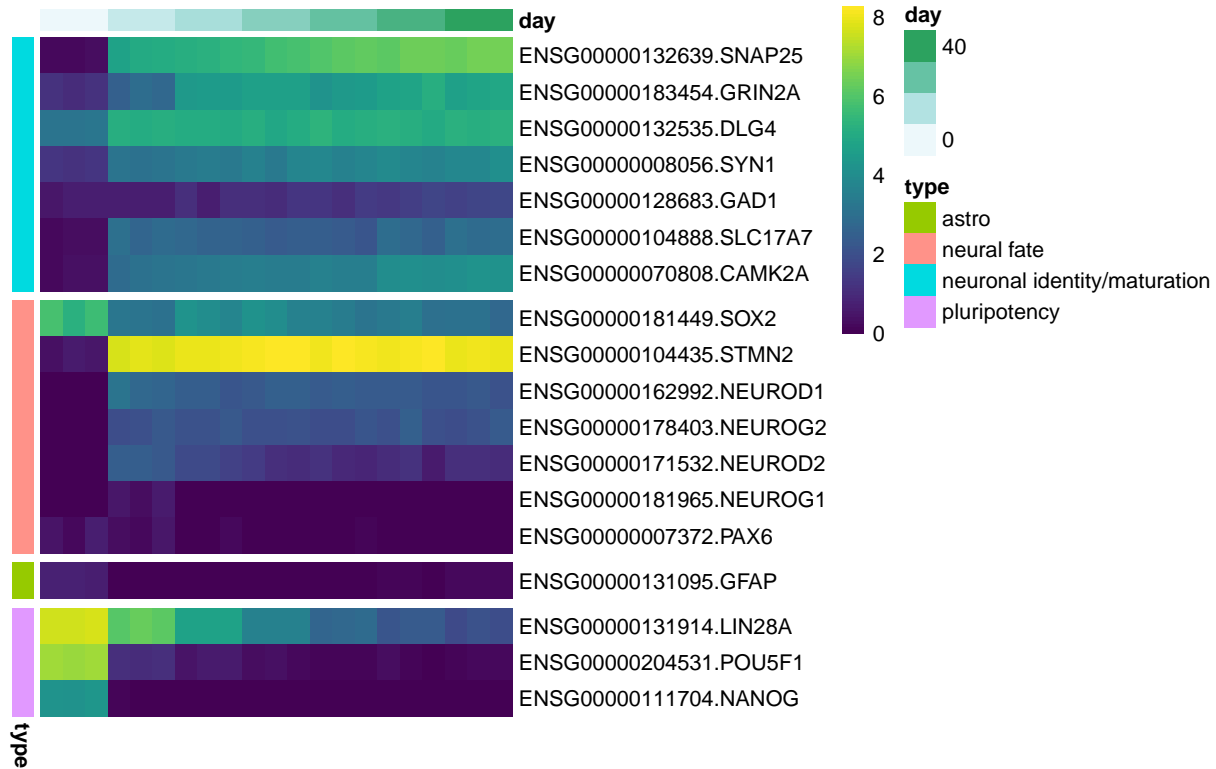
Figure 11: Transcriptional changes of different marker genes provide evidence for a divergence from a pluripotent state as well as neuronal differentiation and maturation.
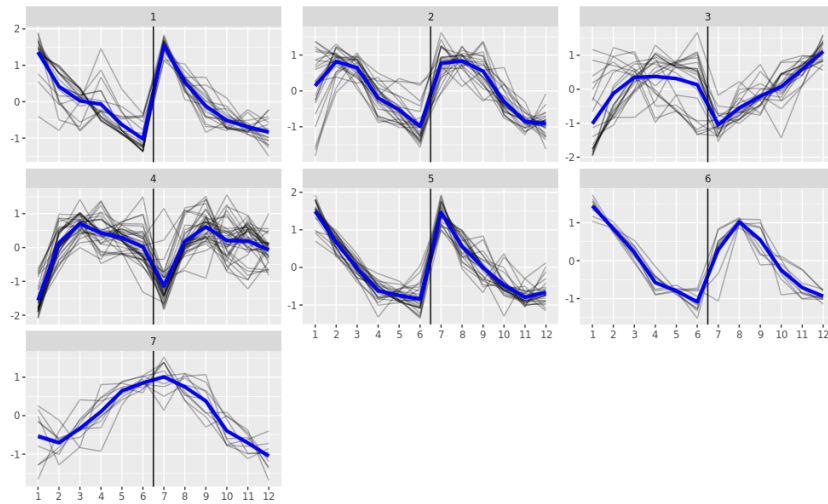
Figure 12: Combined expression patterns of circular and linear counterparts across neuronal maturation stages of Ngn2 programmed hiPSC cells. Smoothed relative expression patterns were combined for visual analysis. Each subplot represents a cluster of genes similarly regulated. To the left of the vertical line are the linear and to the right the circular transcript expressions. Circular and linear RNA derived from the same gene often show co-expression. A divergent expression pattern is evidence for targeted circRNA regulation and thus a functional role in a given context. We were thus primarily interested in clusters 3 and 7. Among these the most differentially expressed circRNAs were BAZ2B and RTN4 respectively. A more detailed list can be seen in Figure 16.

Table 4: Enriched GO terms for miR-153 and miR-181 detected in our Ngn2 neurodifferentiation dataset by enrichMiR.

Term		Annotated	Significant	Enrichment	PValue	FDR	seed	miRNA
GO:0097485	neuron projection guidance	261	44	4.44	3.00e-07	0.000125	UGCAUAG	hsa-miR-153-3p
GO:0048812	neuron projection morphogenesis	627	77	3.24	1.70e-06	0.000510	UGCAUAG	hsa-miR-153-3p
GO:0048667	cell morphogenesis involved in neuron differentiation	562	69	3.24	9.70e-06	0.002170	UGCAUAG	hsa-miR-153-3p
GO:0031175	neuron projection development	934	93	2.62	1.20e-04	0.015500	UGCAUAG	hsa-miR-153-3p
GO:0050767	regulation of neurogenesis	768	75	2.57	3.59e-04	0.034000	UGCAUAG	hsa-miR-153-3p
GO:0045664	regulation of neuron differentiation	627	63	2.65	6.13e-04	0.048200	UGCAUAG	hsa-miR-153-3p
GO:0070571	negative regulation of neuron projection regeneration	15	7	7.95	3.37e-04	0.042700	ACAUUCA	hsa-miR-181
GO:0021954	central nervous system neuron development	74	19	4.37	3.97e-04	0.045400	ACAUUCA	hsa-miR-181

EnrichMiR identifies 2 miRNAs in the Ngn2 differentiation dataset whose expression patterns suggest regulative action

Based on the benchmark results we decided to select *michael* as our method of choice, achieving consistently high scores, high sensitivity and relatively high specificity. The *michael* method is based on a Fisher's test that assumes the number of miRNA binding sites per gene or transcript as categorical variables of its contingency table. We also determined that miRNA restriction lead to significant increases in both sensitivity and specificity. The substitution of unseparated DEA with exon-specific ones however increased specificity at the cost of sensitivity. We thus decided to only adopt the former measure. We ran enrichMiR using the *michael* method and our smallRNA expression data for miRNA tissue-specificity on the Ngn2 neurodifferentiation DE profile. In order to increase the probability of detecting miRNAs potentially involved in neuronal maturation rather than in the departure from stem cell identity, we focused on the last two sequential time transitions: days 33 vs. 27 and days 40 vs. 33. EnrichMiR identified four miRNA family candidates. They are shown in Table 6 together with their subdivision into individual miRNAs.

We conducted a GO term enrichment analysis on the sets of targets of each of the identified miRNA families (based on TargetScan) and filtered the results by terms describing different neural processes. Out of the initial four candidates, two familes showed an enrichment (FDR < 0.05) in such terms: miR-153 and miR-181. These enriched terms included among others regulation of neuron differentiation and neuron projection guidance (see Table 4).

Next, we examined whether the co-regulatory patterns among the targets of the four identified miRNAs were reflected in the expression of the miRNA themselves. After looking up respective DEA profiles across the latest thee timepoints in our Ngn2 differentiation-derived small RNA dataset, we found varying behaviours. While miR-153-3p was found upregulated for both late transitions, all miR-181 variants showed initial upregulation (days 33 vs. 40) and subsequent downregulation (days 40 vs. 33). However, all differential expressions were non-significant. MiR-153-3p expression patterns thus agree with the signatures discovered by enrichMiR, showing target downregulations while the miRNA levels increase. MiR-181 patterns on the other hand imply simultaneous downregulation of both targets and miRNAs thus weakening the hypothesis of direct miR-181 regulation in a context of neuronal maturation.

Discussion

Summary

The Schratt lab created a dataset of human induced Pluripotent Stem Cell (hiPSC) neuronal differentiation and maturation based on Ngn2 programming. It features small and large RNA-seq data. Our goal was to investigate this dataset in order to find miRNA candidates that could have implications in neurodevelopment. For this purpose we decided to adopt an in-house computational tool called enrichMiR. EnrichMiR can employ various statistical tests or methods to identify miRNAs that could or do lead to an observed change in transcriptional signatures when comparing one cell state to another. Before applying enrichMiR on the Ngn2 neurodifferentiation dataset we investigated the following questions: 1) Which method of statistical analysis is most suited for the task? 2) Does, and if so in what way, a restriction of miRNAs to only those

expressed in a given tissue, as opposed to a given organism, affect enrichMiR results? 3) Does, and if so in what way, mRNA expression data derived solely from exonic reads affect enrichMiR results? To address these questions we performed a systematic benchmark and summarized the results using various complementary scores. The methods that performed well across the benchmark datasets were then applied to the Ngn2 neurodifferentiation dataset, yielding miRNAs putatively involved in neuronal maturation.

Benchmark

Examination of the benchmarking scores revealed that a simple method applying Fisher's test on the number of binding sites outperformed many more complex variants. The method of choice, termed *michael* method, distinguishes itself from other overrepresentation tests (e.g. *EN* and *wEN*) by defining the categorical variables of the categories for which overrepresentation was to be assessed as the number of miRNA binding sites per transcript or gene. The *michael* method consistently performed well across benchmark datasets in terms of sensitivity and specificity. In the *Jeong* dataset it was able to find the Positive miRNA where all other methods couldn't. However, different methods will be preferable depending on dataset quality and experimental question. For example, tests like Mann-Whitney (*MW*), two-sided Kolmogorov-Smirnov (*KS*) and simple enrichment (*EN*) displayed low specificity but are very sensitive. On the other hand, tests like Master Regulator Inference Analysis (*aREAmir*) and aggregated one-sided Kolmogorov-Smirnov (*KS2*) proved less sensitive but more specific. Method combinations, either using Fisher or geometric mean p-value aggregation, were in some cases able to mitigate drawbacks of individual methods and improve score performance.

Upon restricting the range of miRNA to be considered as possible enrichMiR outputs to tissue-specific ones, we observed an expected improvement in overall scores for most methods. This measure significantly improved both sensitivity and specificity of methods. In contrast, while the use of exon-specific DEA profiles expectedly increased specificity, it did so at the cost of sensitivity. A reason for this drop could be the low coverage of reads assigned to unprocessed transcripts and the increased complexity of the model fit, which decrease the power of DEA. Regardless, the specificity gained in datasets where large quantities of significant results are identified might justify a loss in sensitivity.

Even though we tested several method combinations that were aggregated using different strategies (*combFish* & *combGeom* methods), most combinations remain unexplored. Combining methods that each do well in particular types of datasets and when applied in conjunction with different measures, such as with and without exon-specific DEAs, might raise enrichMiR's inference accuracy. Moreover, all our tests were confined to the discovery of a single Positive miRNA family, which although simplifying scoring to a modest degree, was not required. We thus intend to incorporate datasets with more complex perturbation profiles.

Applying enrichMiR on our Ngn2-mediated neurodifferentiation dataset

Based on the insights gained from our benchmark, we ran enrichMiR on the Ngn2 neurodifferentiation dataset implementing the *michael* statistical analysis method and miRNA restriction to those expressed in human neurons. EnrichMiR identified several miRNAs with possible implications in neuronal maturation. Two of them, miR-153-3p and miR-181, showed enrichments of several GO terms associated with neuronal processes. In the case of miR-153-3p miRNA expression trends reflected the expression changes of their targets as identified by enrichMiR. For miR-181 on the other hand we detected trends opposite to their expected direction. Although the trends in miRNA expression were non-significant in both cases, this weakens the notion that miR-181 might have a functional role in a context of neuronal maturation.

That being said, it has been previously shown that non-coding transcripts, such as circRNAs, are capable of miRNA sequestration (Thomas B. Hansen et al. 2013), resulting in their partial neutralization. It has been suggested that such miRNA retention can rapidly be reversed (Rybak-Wolf et al. 2015a) (Thomas B Hansen et al. 2011), thus increasing miRNA-mediated repression at equal or possibly lower total abundances. CircRNAs in particular have been found to be relatively abundant in neuronal contexts, especially in synapses (Rybak-Wolf et al. 2015a). Whether or not such mechanisms apply in this case remains to be discovered.

A look into the published literature has revealed various neuronal associations for both described miRNA candidates. MiR-153 was found to prevent a premature switch from neurogenesis to gliogenesis (Tsuyama et al. 2015). It was further implicated in motor neuron growth and branching during development in zebrafish by inhibiting Snap25, a component of the SNARE complex involved in vesicle exocytosis (Wei et al. 2013). The same component was associated with mental disorders such as schizophrenia and bipolar disorder (Gray et al. 2010). MiR-153 was also shown to suppress alpha-Synuclein (SNCA), whose accumulation precipitates Parkinson disease (Doxakis 2010). Other studies demonstrated its implication in the promotion of adult neurogenesis in aged mice (Qiao et al. 2020), and mouse models of autism (Y.-H. You et al. 2019) by regulating the Notch and JAK-STAT signaling pathways respectively.

As for miR-181, it was implicated in the inhibition of axon and dendrite growth in fish retina tissue (Carrella, D'Agostino, et al. 2015) (Carrella, Barbato, et al. 2015) as well as rat hippocampal neurons (Y. Liu et al. 2013) via suppression of the MAPK signalling pathway and the associated CREB transcription factor. Application of binding site prediction tools revealed its role in AMPA receptor inhibition, leading to reduced dendritic spine formation and firing activity in the nucleus accumbens (Saba et al. 2012). Its implication in dopaminergic reward systems has been corroborated by findings involving natural positive affect (Wingo et al. 2017) and drugs such as amphetamine (Zhao et al. 2016) and cocaine (Nestler 2001) (Chandrasekar and Dreyer 2009) (K.-W. Lee et al. 2006). The study of schizophrenia has confirmed its suppressive effect on the AMPA receptor and revealed effects on the calcium sensor gene VSNL1 in the superior temporal gyrus (Beveridge et al. 2008). Schizophrenia and other disorders, such as PTSD (Zhou et al. 2014) and corticosterone induced depression in mice (Dwivedi et al. 2015), share a common pattern of miR-181 upregulation in the brain and downregulation in the blood (Wingo et al. 2017). Stroke studies have further shown that miR-181 suppression using an antagomir resulted in protection from ischemic damage both when administered before (Ouyang et al. 2012) and after induced stroke, enabling for motor function recovery (Xu et al. 2015).

Analysis of circRNA dynamics in our Ngn2-mediated neurodifferentiation dataset

We analysed the expression behaviour of circRNAs relative to linear counterparts that shared a common host gene inside our Ngn2 neurodifferentiation dataset over the course of the different recorded timepoints. By applying KNN clustering, we were able to identify sets of circRNAs that showed expression patterns independent of linear counterparts. Such differential expression suggests that the identified circRNAs hold importance in a neurodevelopmental context. However, these instances were uncommon, with most circRNAs mirroring their linear variants.

It has been previously shown that certain circRNA variants are implicated in neuronal development (Venø et al. 2015) and specifically in synapses (X. You et al. 2015) (Rybak-Wolf et al. 2015b) as well as neurological diseases (Shao and Chen 2016). However, of the circRNAs that we discovered and that showed highest differential expression, none have yet been found to be implicated in neurodevelopmental processes.

enrichMiR

We plan to test the performance of enrichMiR when supplied with an alternative miRNA binding site prediction tool to TargetScan. Our lab generated a new search algorithm that operates on a quantitative model of miRNA binding affinity adapted from the Bartel group. Individual miRNAs will no longer be categorized based on seed sequence congruence only but based on logarithmic dissociation rates (K_D) (McGeary, Lin, Shi, Pham, et al. 2019). To what degree this alternative categorization will impact the results obtained by enrichMiR remains to be seen.

Importantly, its application on transcript-level data (rather than a single 3'UTR per gene, as is the case for TargetScan) will allow us to not only differentiate between varying 3'UTR regions of mRNA transcripts derived from a single gene but also investigate binding sites residing in CDR and 5'UTR regions. Moreover, detection of binding sites on non-coding RNAs such as circRNAs will be rendered possible. Although translation appears to be necessary for miRNAs to downregulate their target (Biasini et al. 2020), binding to

Table 5: Number of differentially expressed transcripts per RNA-seq dataset over the different recorded stages of Ngn2 neurodifferentiation. DEA.all is the comparison between all neuronal stages (stages 9, 15, 21, 27, 33 and 40) and the initial undifferentiated state (day 0). DEA.neuronal is the comparison between all later neuronal stages (stages, 15, 21, 27, 33 and 40) to the first neuronal stage (stage 9). All other DEAs represent sequential pair-wise analyses over all stages.

	longRNA.DE.nr	smallRNA.DE.nr
DEA.all	21863	456
DEA.neuronal	16485	188
DEA.9v0	18827	279
DEA.15v9	6389	4
DEA.21v15	3328	0
DEA.27v21	4616	0
DEA.33v27	2627	0
DEA.40v33	1031	0

non-coding transcripts can still have an impact on miRNA availability (Thomas B. Hansen et al. 2013).

We ultimately seek to improve enrichMiR's capability to detect miRNAs associated to various mRNA expression signatures. This might allow for the discovery of miRNA contributing to or alleviating complex disorders such as Autism Spectrum Disorder (ASD) (Hicks and Middleton 2016) and schizophrenia (K. L. Stark et al. 2008) (Hu et al. 2019). In cases where these candidates are found to not directly affect a cell's transciptomic landscape, we may gain new means and angles from which to affect it.

Materials & Methods

Ngn2-mediated neuronal differentiation and maturation datasets

The Schratt group created a novel neuronal differentiation and maturation protocol based on Ngn2 transcription factor programming. Human induced Pluripotent Stem Cells (hiPSC) were cultured in vitro in the absence of astrocytes. The resulting mature neurons exhibited functional synaptic spines. Confirmation of neuronal identity and the degree of maturation has been and is still being undertaken, including morphological analysis, pre- and post-synaptic marker staining and RNA marker analysis.
Small and long RNAseq and proteomic profiling were performed at constant intervals at days 0, 9, 15, 21, 27, 33, 40 of differentiation. The number of differentially expressed genes yielded by DEA can be viewed under Table 5.

Benchmark datasets

To test the performance of enrichMiR's statistical methods, we required RNAseq datasets generated upon perturbation of one or more known miRNA(s). To make sure both up- and down-regulation of miRNA targets could be handled by the enrichMiR statistical analysis methods, we included miRNA transfections as well as knockouts to our benchmark set. We chose 5 different datasets that fulfilled our requirements, together offering a total of 33 miRNA perturbation experiments. We abbreviated these datasets as follows: *Bartel*, *Rat*, *Jeong*, *Amin*, *Cherone*.

Assignment of true miRNA families

Each of the experiments belonging to one of the 5 benchmark datasets corresponds to a single miRNA family perturbation and each of them was assigned a single true miRNA seed sequence. This true miRNA family was used for the scoring of the enrichMiR results.

Since guide strand sequences of the perturbation miRNAs were available for the *Bartel* experiments and in other cases only a single strand was listed in our species-specific TargetScan library, we selected either the 3p or 5p strand to constitute the true family seed. In the case of miR-138 perturbation experiments, which were available in different datasets, we chose true miRNA families as follows: 1) In the *Bartel* dataset we chose based on the available guide strand sequence. 2) In the Rat dataset we chose based on the only option that gave significant results across the different statistical tests. 3) In the *Cherone* dataset we chose based on the respective *Bartel* sequence. In the latter experiments there was no instance in which the true miRNA family was discovered.

Bartel dataset

McGeary, Lin, Shi, Bisaria, et al. 2018 generated an extensive collection of 29 single miRNA duplex transfections performed on 2 different human celltypes: 17 miRNA transfections were performed on HeLa and 12 on HEK293 cells. The experiments were divided into 5 batches in a way that each treatment was replicated twice. RNA extraction took place 24 h after transfection. A summarized list of experiments and exact sequences of their transfected miRNA duplexes is available in the Supplemental Material. We used all the available samples except for the lsy-6 transfections, since this miRNA is not expressed in human cells and therefore not listed in the TargetScan annotation of miRNA targets that we used for human celltypes (see subsection miRNA target annotations). McGeary et al. chose to omit control samples in their experimental design, and used raw foldchanges (without statistical testing) to other transfection experiments. To obtain more accurate dysregulations, we instead created controls by taking the median expression values across the transfection samples of each batch. This approach could lead to a bias in the control sample for genes that were affected similarly in more than half of the applied treatments (miRNA transfections). Large upregulations in a minority of samples would be discernible after calculating the log2 fold changes if in a majority of samples the same genes were highly suppressed by their treatment miRNA. We indeed identified such genes and confirmed their enrichment for binding sites of several of the miRNAs. They were excluded from differential expression profiles of all samples. The identification of biasing genes as well as control sample assembly are supplied online[3].

Rat dataset

The Schratt group generated a yet unpublished dataset of miR-499 and miR-138 duplex transfections on rat cell-cultures consisting of hippocampal neurons and glia. Transfections of treatment miRNA and negative duplexes were replicated thrice. Transfections were done on DIV7 (days in vitro) and extraction was done on DIV12.

Jeong dataset

Jeong et al. 2016 generated individual miR-221 and miR-222 knockouts as well as a combined double-knockouts of human SNU-638 gastric cancer cells. Since both miRNAs are paralogous we only chose to include the miR-221/222 double-knockout (DKO). Both WT and DKO samples were replicated twice. RNA extraction was performed several days after transfection.

Amin dataset

Amin et al. 2015 generated individual miR-218-1 and miR-218-2 knockout mice as well as a combined double-knockout mice. Since both miR-218-1 and miR-218-2 are paralogous we only chose to include the miR-218 DKO. DKO samples were replicated twice. For the WT control samples we chose 2 which conformed with the former, consisting of unpaired reads. RNA extraction was performed on day 4 neurospheres and spinal cord dissections (motoneurons) deriving from E12.5 mice.

---

[3]https://github.com/ETHZ-INS/enrichMiR_benchmark

Cherone dataset

Cherone, Jorgji, and Burge 2019 generated miR-138-1/miR-138-2 DKO of mouse CAD cells. RNA extractions were performed on day 0 and day 1 after initiation of neuronal differentiation. WT and DKO samples were replicated twice. RNA extraction was performed after several days.

RNA-seq mapping & quantification

RNA-seq reads were mapped using the STAR alignment software[4] either verion 2.7.3a or verion 2.6.1b for mouse samples, due to genome compatibility issues (Dobin et al. 2013). We used the following Ensembl reference genomes: GRCh38.p13 for human samples (*Bartel & Jeong*), Rnor6.0 for the *Rat* sample and GRCm38.p5 for mouse samples (*Amin & Cherone*). The detailed STAR alignment options for all datasets except *Rat* can be found online[5].

Quantification of unseparated samples with transcript resolution was done using the featureCounts tool. Separation of exonic reads from combined intronic and exonic samples was equally performed using featureCounts combined with subsequent R processing. Both unseparated and exonic quantification results were saved into SummarizedExperiment objects in R. All scripts harbouring the performed processing steps for each dataset are available online[6].

Expression thresholds & Differential expression analysis (DEA)

In all datasets we excluded all transcripts that did not have at least 10 counts in at least 2 samples. We performed differential expression analysis (DEA) in R using negative binomial generalized log-linear models provided by the EdgeR RNA-seq analysis package[7] (Robinson, McCarthy, and Smyth 2010). Whenever possible the linear models were kept simple due to the low replicate number of most of our datasets ($\sim treatment$ vs. ~ 1). We included an additional annotated variable for the *Bartel*, *Rat* and *Jeong* datasets to account for the large batch effects we discovered after Principal Component Analysis (PCA). The DEA results were saved within SummarizedExperiment objects[8]. All involved R scripts and .rds files are available online[9].

Exon-specific DEA

We separately quantified 1) purely split exonic reads, 2) reads overlapping exons, excluding split reads and 3) reads overlapping the full transcript, excluding split reads. The first quantification represents fully processed mature mRNA. The second and third were subtracted in order to isolate reads from purely unprocessed pre-mRNA transcripts. Having joined both processed and unprocessed datasets, we performed combined DEA. We implemented a negative binomial log-linear model as described above, additionally accounting for the combined effect of miRNA perturbation (treatment) and mRNA maturation stage (processed/unprocessed): $\sim readtype * treatment$ vs. $\sim readtype + treatment$. Depending on the dataset we accounted for batch effects as described above. We were thus able to isolate the relative effect of mature mRNA for each differential analysis (Zeisel et al. 2011) (Gaidatzis et al. 2015).

MiRNA target annotations

miRNA target annotations for different species were taken from TargetScan (Release 7.2) (Agarwal et al. 2015). We downloaded tables found under "Summary Counts, all predictions"[10], containing information about conserved and unconserved target sites. Since there is no TargetScan version based on an alignment

---

[4]https://github.com/alexdobin/STAR
[5]https://github.com/ETHZ-INS/enrichMiR_benchmark
[6]https://github.com/ETHZ-INS/enrichMiR_benchmark
[7]https://bioconductor.org/packages/release/bioc/html/edgeR.html
[8]https://www.bioconductor.org/packages/release/bioc/html/SummarizedExperiment.html
[9]https://github.com/ETHZ-INS/enrichMiR_benchmark
[10]http://www.targetscan.org/cgi-bin/targetscan/data_download.vert72.cgi

to a rat genome reference, we used the TargetScanMouse[11] site for both mouse and rat datasets to ensure higher UTR similarities. TargetScanHuman[12] provided us with targets for our human samples. In each case only entries with corresponding species ID were considered for further analysis. Compilations of miRNAs and their seed sequences were downloaded under "miR Family"[13] on the respective TargetScan sites. These compilations served as species-specific miRNA profiles.

Tissue-specific miRNA expression profiles

enrichMiR test performances were assessed with and without the supplementation of tissue-specific miRNA expression profiles. Profiles for human HeLa and HEK293 cells were acquired through the microRNAome R package[14] (McCall et al. 2017). The neuronal expression profile for rat was acquired from Rajman et al. 2017. The mouse profile was acquired from He et al. 2012. Samples of all brain tissues were combined except for the ones derived from the cerebellum.

GO term enrichment

GO term enrichment analysis was performed using the Goseq gene ontology analyser[15] R package (Young et al. 2010) over a custom implementation stemming from an in-house package. This method is able to correct for biases based transcript lengths. We ran an analysis for each separate miRNA family identified by enrichMiR. We supplied a background gene list, composed of all transcripts found in our Ngn2 neurodifferentiation long RNA dataset, and for each iteration the putative target genes of the respective family. For minimum and maximum set sizes we used the default parameters of 10 and 1000.

Session information

```
## R version 3.6.1 (2019-07-05)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 18.04 LTS
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/openblas/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/libopenblasp-r0.2.20.so
##
## locale:
##  [1] LC_CTYPE=C.UTF-8       LC_NUMERIC=C           LC_TIME=C.UTF-8
##  [4] LC_COLLATE=C.UTF-8     LC_MONETARY=C.UTF-8    LC_MESSAGES=C.UTF-8
##  [7] LC_PAPER=C.UTF-8       LC_NAME=C              LC_ADDRESS=C
## [10] LC_TELEPHONE=C         LC_MEASUREMENT=C.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
##  [1] grid      stats4    parallel  stats     graphics  grDevices utils
##  [8] datasets  methods   base
##
## other attached packages:
##  [1] org.Hs.eg.db_3.10.0        GO.db_3.10.0
##  [3] AnnotationDbi_1.48.0       goseq_1.38.0
##  [5] geneLenDataBase_1.22.0     BiasedUrn_1.07
##  [7] kableExtra_1.1.0           ComplexHeatmap_2.2.0
##  [9] ggrepel_0.8.1              RUVSeq_1.20.0
## [11] edgeR_3.28.1               limma_3.42.2
## [13] EDASeq_2.20.0              ShortRead_1.44.3
## [15] GenomicAlignments_1.22.1   SummarizedExperiment_1.16.1
## [17] DelayedArray_0.12.2        Rsamtools_2.2.3
```

---

[11] http://www.targetscan.org/mmu_72/
[12] http://www.targetscan.org/vert_72/
[13] http://www.targetscan.org/cgi-bin/targetscan/data_download.vert72.cgi
[14] https://bioconductor.org/packages/release/data/experiment/html/microRNAome.html
[15] https://www.bioconductor.org/packages/release/bioc/html/goseq.html

```
## [19] GenomicRanges_1.38.0      GenomeInfoDb_1.22.0
## [21] Biostrings_2.54.0         XVector_0.26.0
## [23] IRanges_2.20.2            S4Vectors_0.24.3
## [25] BiocParallel_1.20.1       Biobase_2.46.0
## [27] BiocGenerics_0.32.0       matrixStats_0.55.0
## [29] cowplot_1.0.0             dplyr_0.8.4
## [31] data.table_1.12.8         RColorBrewer_1.1-2
## [33] viridis_0.5.1             viridisLite_0.3.0
## [35] DT_0.12                   plgINS_0.1.5
## [37] ggplot2_3.2.1
##
## loaded via a namespace (and not attached):
##   [1] R.utils_2.9.2          tidyselect_1.0.0       RSQLite_2.2.0
##   [4] htmlwidgets_1.5.1      TSP_1.1-9              Rtsne_0.15
##   [7] DESeq_1.38.0           munsell_0.5.0          codetools_0.2-16
##  [10] withr_2.2.0            colorspace_1.4-1       knitr_1.28
##  [13] rstudioapi_0.11        labeling_0.3           GenomeInfoDbData_1.2.2
##  [16] hwriter_1.3.2          bit64_0.9-7            farver_2.0.3
##  [19] pheatmap_1.0.12        vctrs_0.2.3            xfun_0.12
##  [22] BiocFileCache_1.10.2   gclus_1.3.2            R6_2.4.1
##  [25] clue_0.3-57            seriation_1.2-8        locfit_1.5-9.4
##  [28] bitops_1.0-6           assertthat_0.2.1       scales_1.1.0
##  [31] SEtools_1.1.9          nnet_7.3-13            gtable_0.3.0
##  [34] rlang_0.4.5            genefilter_1.68.0      GlobalOptions_0.1.1
##  [37] splines_3.6.1          rtracklayer_1.46.0     lazyeval_0.2.2
##  [40] acepack_1.4.1          GEOquery_2.54.1        checkmate_2.0.0
##  [43] yaml_2.2.1             reshape2_1.4.4         GenomicFeatures_1.38.2
##  [46] backports_1.1.5        Hmisc_4.3-1            tools_3.6.1
##  [49] gplots_3.0.3           Rcpp_1.0.3             plyr_1.8.6
##  [52] base64enc_0.1-3        progress_1.2.2         zlibbioc_1.32.0
##  [55] purrr_0.3.3            RCurl_1.98-1.1         prettyunits_1.1.1
##  [58] rpart_4.1-15           openssl_1.4.1          GetoptLong_0.1.8
##  [61] cluster_2.1.0          magrittr_1.5           openxlsx_4.1.4
##  [64] circlize_0.4.8         randomcoloR_1.1.0.1    aroma.light_3.16.0
##  [67] hms_0.5.3              evaluate_0.14          xtable_1.8-4
##  [70] XML_3.99-0.3           jpeg_0.1-8.1           gridExtra_2.3
##  [73] shape_1.4.4            compiler_3.6.1         biomaRt_2.42.0
##  [76] tibble_2.1.3           KernSmooth_2.23-16     V8_3.0.1
##  [79] crayon_1.3.4           R.oo_1.23.0            htmltools_0.4.0
##  [82] mgcv_1.8-31            Formula_1.2-3          tidyr_1.0.2
##  [85] geneplotter_1.64.0     DBI_1.1.0              dbplyr_1.4.2
##  [88] MASS_7.3-51.5          rappdirs_0.3.1         Matrix_1.2-18
##  [91] readr_1.3.1            R.methodsS3_1.8.0      gdata_2.18.0
##  [94] pkgconfig_2.0.3        registry_0.5-1         foreign_0.8-76
##  [97] xml2_1.2.2             foreach_1.4.8          annotate_1.64.0
## [100] webshot_0.5.2          SRAdb_1.48.2           rvest_0.3.5
## [103] stringr_1.4.0          digest_0.6.25          rmarkdown_2.1
## [106] htmlTable_1.13.3       dendextend_1.13.4      curl_4.3
## [109] gtools_3.8.2           rjson_0.2.20           lifecycle_0.1.0
## [112] nlme_3.1-145           jsonlite_1.6.1         askpass_1.1
## [115] pillar_1.4.3           lattice_0.20-40        httr_1.4.1
## [118] survival_3.1-8         glue_1.4.1             zip_2.0.4
## [121] png_0.1-7              iterators_1.0.12       bit_1.1-15.2
## [124] stringi_1.4.6          blob_1.2.1             DESeq2_1.26.0
## [127] latticeExtra_0.6-29    caTools_1.18.0         memoise_1.1.0
```

Supplementary Material

· **Enclosed** A declaration of originality
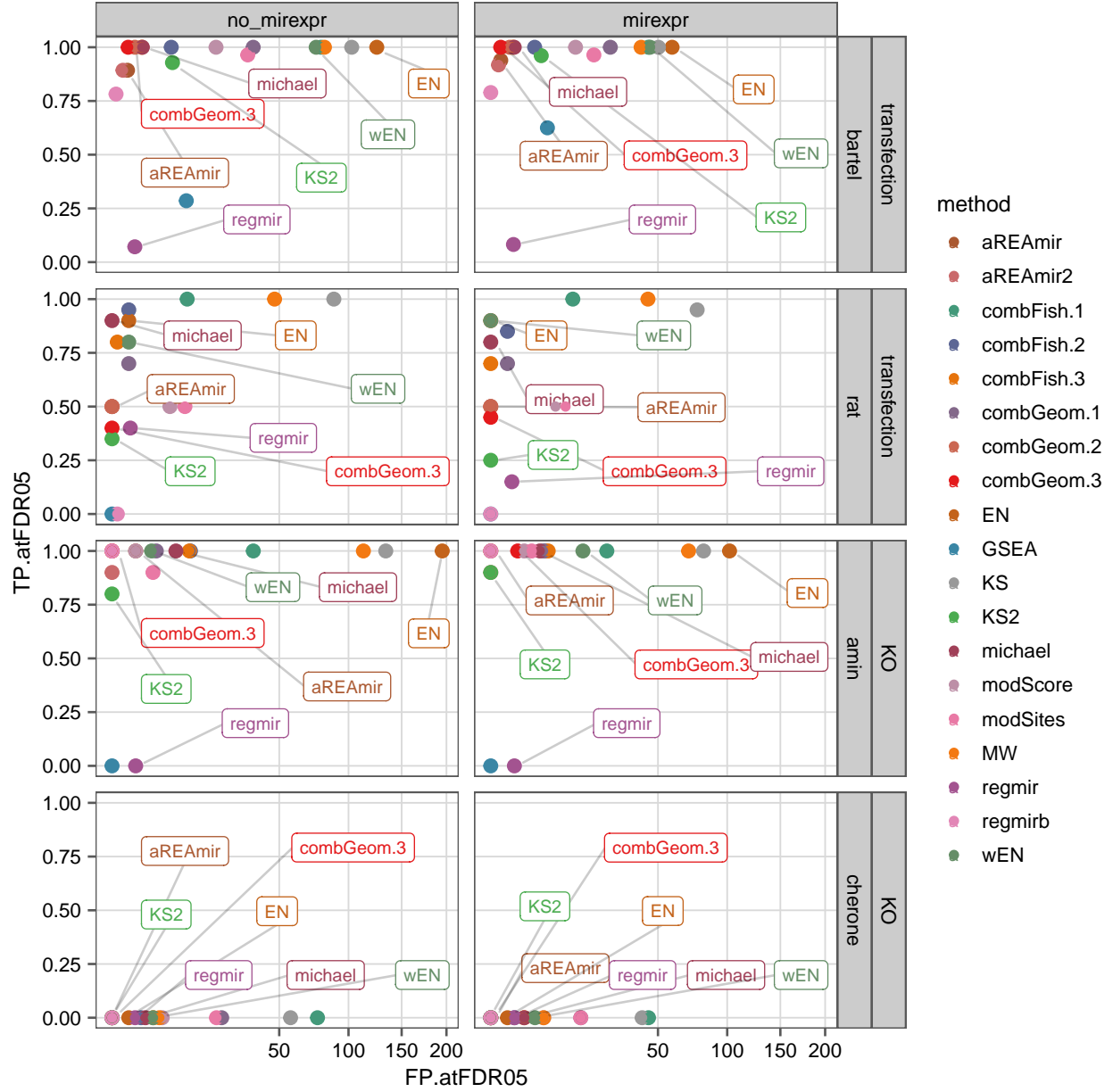
Figure 13: Comparison of FPs and TPR for each enrichMiR method by supplementation of miRNA expression profiles. Each datapoint of a single color reflects the mean FPs and TPR across the samples of a respective dataset and across TS perturbations. In most cases the FP numbers decreased upon supplementation of miRNA expression data (mirexpr). The largest differences in TPR could be observed for the GSEA and KS2 methods, both increasing. Both FP and TPR changes were significant.
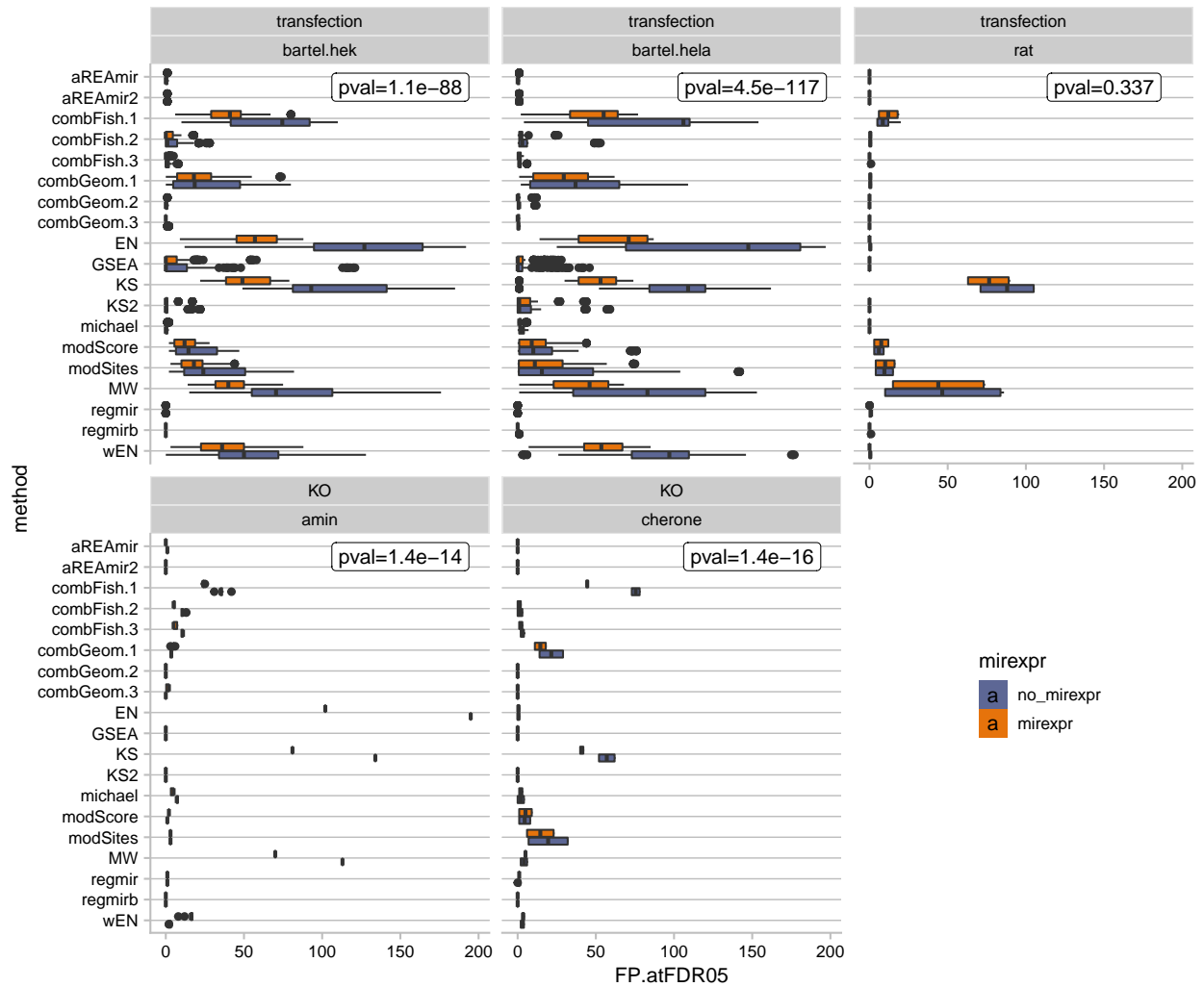
Figure 14: Comparison of False-Positive (FP) numbers between enrichMiR analyses with and without supplementation of miRNA expression (mirexpr) profiles for all methods. The FP values (x-axis) are split according to their dataset membership and experimental design. For each dataset the corresponding p-value is displayed, relaying the significance of the mirexpr supplementation effect across all methods. For all datasets except Rat the FPs showed a significant decrease.
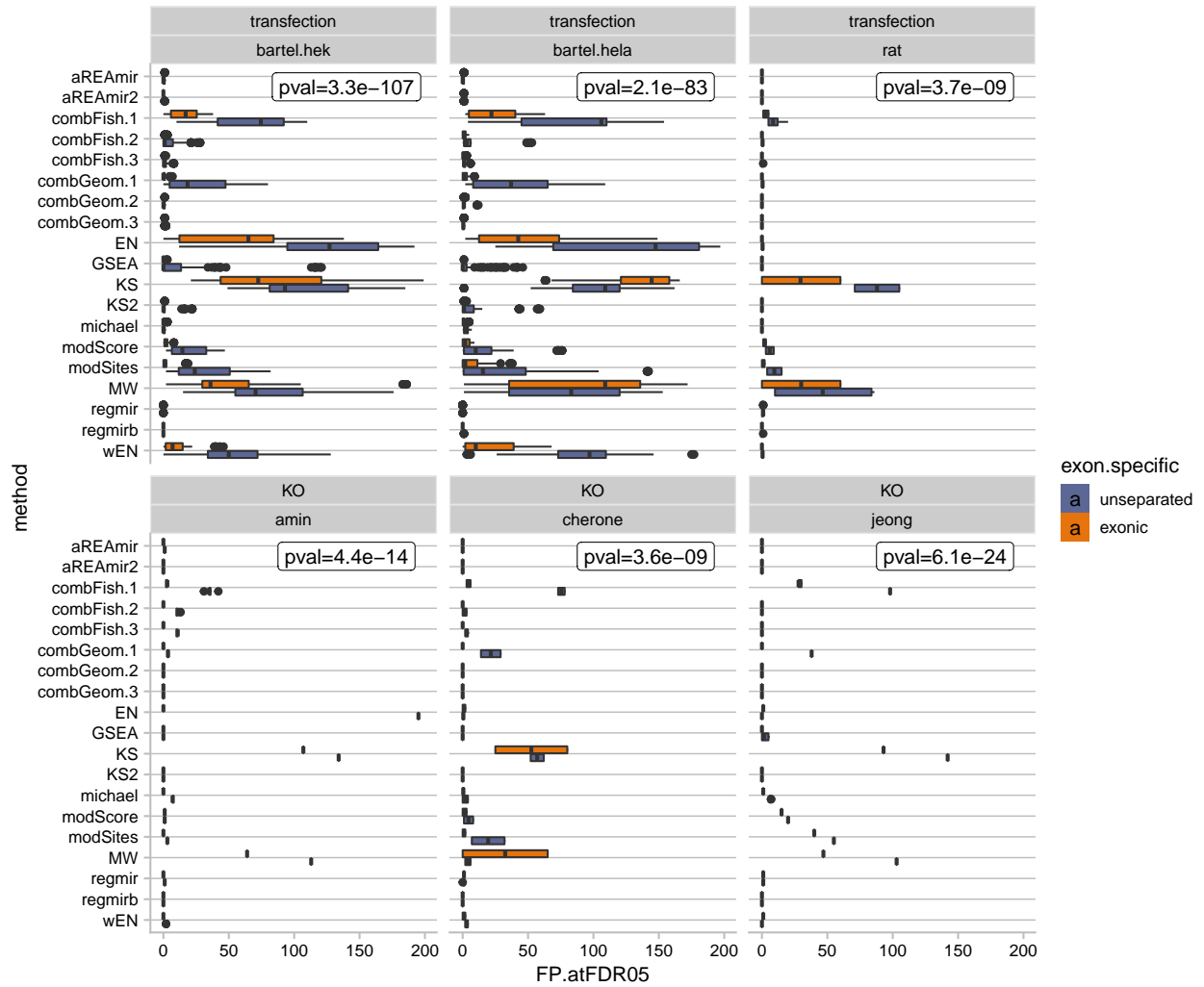
Figure 15: Comparison of FPs between enrichMiR analyses with and without exon-specific DEA inputs for all methods. The FP values (x-axis) are split according to their dataset membership and experimental design. For each dataset the corresponding p-value is displayed, relaying the significance of the mirexpr supplementation effect across all methods. For all datasets except Rat the FP numbers showed a significant decrease.

(a)

(b)

| symbol	FDR	logCPM
<factor>	<numeric>	<numeric>
SMARCA5	2.46e-12	1.62
PTK2	2.93e-10	-0.399
HPS5	2.64e-08	-0.673
SCAF8	8.3e-07	-0.916
BRWD3	1.53e-06	-1.43
CHD7	1.69e-06	-1.62
SNHG14	6.77e-06	-2.52
BABAM1	7.56e-06	-1.83
ASPH	7.95e-06	-2.15
MRPS35	1.23e-05	-1.4

(c)

symbol	FDR	logCPM
<factor>	<numeric>	<numeric>
RTN4	6.16e-05	1.43
AKT3	0.00302	-2.68
SETD3	0.00355	-0.236
HIPK3	0.00356	1.66
MAP2	0.00671	-1.72
GTDC1	0.00742	-2.38
RSRC1	0.0337	-3.09
RAB6A	0.0353	-1.32
PSD3	0.0405	-2.44
ANK2	0.0475	-2.9

(d)

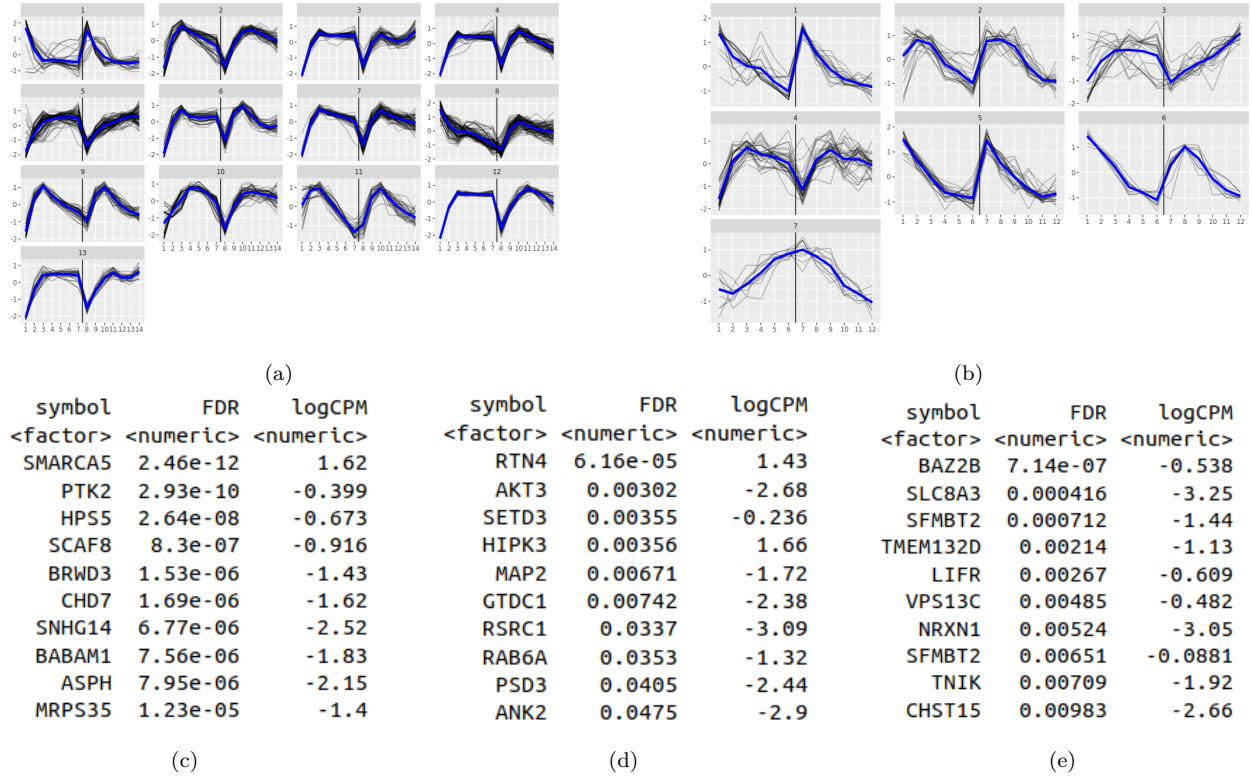| symbol	FDR	logCPM
<factor>	<numeric>	<numeric>
BAZ2B	7.14e-07	-0.538
SLC8A3	0.000416	-3.25
SFMBT2	0.000712	-1.44
TMEM132D	0.00214	-1.13
LIFR	0.00267	-0.609
VPS13C	0.00485	-0.482
NRXN1	0.00524	-3.05
SFMBT2	0.00651	-0.0881
TNIK	0.00709	-1.92
CHST15	0.00983	-2.66

(e)

Figure 16: Combined expression patterns of circular and linear counterparts across neuronal maturation stages of Ngn2 programmed hiPSC cells. (A) Expression patterns of genes differentially expressed (logFC > 1 and FDR < 0.05) between the stem cell stage (stage 0) and differentiated stages (stages 9-40). 13 clusters were generated each separated into linear (left) and circular (right) variants. (B) Expression patterns of genes differentially expressed (logFC > 1 and FDR < 0.05) between early neuronal stage (stage 9) and later neuronal stages (stages 15-40). 7 clusters were generated. (C-E) Lists of genes ordered according to FDR. List (C) is composed of genes found in cluster 8 of in (A). Lists (D) and (E) derive from clusters 7 and 3 of the clustering shown in (B) respectively.

Table 6: List of test significances and DEA results derived from the latest recorded transitions of our Ngn2 neurodifferentiation dataset. 'miRNA.logFC.33v27' represents the log FC of the earlier transition with corresponding FDR values under 'miRNA.FDR.33v27'. The same information is shown for the later transition. The last column shows the mean log CPM over all 3 involved stages (stages 27, 33 and 40). Here we see that miR-153 levels increase over both transitions while the miR-181 levels first rise then decrease.

	family	michael.FDR	miRNA.logFC.33v27	miRNA.FDR.33v27	miRNA.logFC.40v33	miRNA.FDR.40v33	miRNA.mean.logCPM
hsa-miR-210-3p	UGUGCGU	0.0381	-0.0546	0.982	-0.190	0.750	5.56
hsa-miR-181a-5p	ACAUUCA	0.0182	0.2130	0.928	-0.484	0.454	16.20
hsa-miR-181b-5p	ACAUUCA	0.0182	0.0818	0.964	-0.355	0.453	12.50
hsa-miR-181c-5p	ACAUUCA	0.0182	0.4550	0.805	-0.490	0.478	11.20
hsa-miR-181d-5p	ACAUUCA	0.0182	0.3770	0.693	-0.107	0.826	9.80
hsa-miR-153-3p	UGCAUAG	0.0182	0.5520	0.649	0.294	0.524	8.86
hsa-let-7a-5p	GAGGUAG	0.0182	1.3300	0.649	0.695	0.300	9.45
hsa-let-7e-5p	GAGGUAG	0.0182	1.2200	0.649	0.333	0.556	9.68
hsa-let-7f-5p	GAGGUAG	0.0182	0.7330	0.693	0.190	0.774	8.29
hsa-let-7g-5p	GAGGUAG	0.0182	0.7060	0.757	0.506	0.455	5.17
hsa-let-7i-5p	GAGGUAG	0.0182	0.3930	0.898	-0.156	0.849	3.88
hsa-miR-98-5p	GAGGUAG	0.0182	0.9820	0.676	0.219	0.769	4.24

References

Agarwal, Vikram et al. (Aug. 2015). "Predicting effective microRNA target sites in mammalian mRNAs". In: *eLife* 4. Ed. by Elisa Izaurralde, e05005. ISSN: 2050-084X. DOI: 10.7554/eLife.05005[16]. URL: https://doi.org/10.7554/eLife.05005 (visited on 05/25/2019).

Amin, Neal D. et al. (Dec. 2015). "Loss of motoneuron-specific microRNA-218 causes systemic neuromuscular failure". In: *Science (New York, N.Y.)* 350.6267, pp. 1525–1529. ISSN: 0036-8075. DOI: 10.1126/science.aad2509[17]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4913787/ (visited on 06/03/2020).

Arvey, Aaron et al. (Apr. 2010). "Target mRNA abundance dilutes microRNA and siRNA activity". In: *Molecular Systems Biology* 6, p. 363. ISSN: 1744-4292. DOI: 10.1038/msb.2010.24[18]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2872614/ (visited on 06/15/2020).

Backes, Christina et al. (July 2016). "miEAA: microRNA enrichment analysis and annotation". en. In: *Nucleic Acids Research* 44.W1. Publisher: Oxford Academic, W110–W116. ISSN: 0305-1048. DOI: 10.1093/nar/gkw345[19]. URL: https://academic.oup.com/nar/article/44/W1/W110/2499340 (visited on 06/16/2020).

Baek, Daehyun et al. (Sept. 2008). "The impact of microRNAs on protein output". en. In: *Nature* 455.7209. Number: 7209 Publisher: Nature Publishing Group, pp. 64–71. ISSN: 1476-4687. DOI: 10.1038/nature07242[20]. URL: https://www.nature.com/articles/nature07242 (visited on 06/15/2020).

Bail, Sophie et al. (Jan. 2010). "Differential regulation of microRNA stability". en. In: *RNA* 16.5. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab, pp. 1032–1039. ISSN: 1355-3382, 1469-9001. DOI: 10.1261/rna.1851510[21]. URL: http://rnajournal.cshlp.org/content/16/5/1032 (visited on 06/18/2020).

Bartel, David P. (Mar. 2018). "Metazoan MicroRNAs". English. In: *Cell* 173.1, pp. 20–51. ISSN: 0092-8674, 1097-4172. DOI: 10.1016/j.cell.2018.03.006[22]. URL: https://www.cell.com/cell/abstract/S0092-8674(18)30286-1 (visited on 05/25/2019).

Beveridge, Natalie J. et al. (Apr. 2008). "Dysregulation of miRNA 181b in the temporal cortex in schizophrenia". en. In: *Human Molecular Genetics* 17.8. Publisher: Oxford Academic, pp. 1156–1168. ISSN: 0964-6906.

---

[16]https://doi.org/10.7554/eLife.05005
[17]https://doi.org/10.1126/science.aad2509
[18]https://doi.org/10.1038/msb.2010.24
[19]https://doi.org/10.1093/nar/gkw345
[20]https://doi.org/10.1038/nature07242
[21]https://doi.org/10.1261/rna.1851510
[22]https://doi.org/10.1016/j.cell.2018.03.006

DOI: 10.1093/hmg/ddn005[23]. URL: https://academic.oup.com/hmg/article/17/8/1156/651613 (visited on 06/17/2020).

Biasini, Adriano et al. (Jan. 2020). "Translation is required for miRNA-dependent decay of endogenous transcripts". en. In: *bioRxiv*. Publisher: Cold Spring Harbor Laboratory Section: New Results, p. 2020.01.21.913483. DOI: 10.1101/2020.01.21.913483[24]. URL: https://www.biorxiv.org/content/10.1101/2020.01.21.913483v1 (visited on 06/19/2020).

Carrella, Sabrina, Sara Barbato, et al. (Dec. 2015). "TGF- Controls miR-181/ERK Regulatory Network during Retinal Axon Specification and Growth". In: *PLoS ONE* 10.12. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0144129[25]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4671616/ (visited on 06/17/2020).

Carrella, Sabrina, Ylenia D'Agostino, et al. (Nov. 2015). "miR-181a/b control the assembly of visual circuitry by regulating retinal axon specification and growth". In: *Developmental Neurobiology* 75.11, pp. 1252–1267. ISSN: 1932-8451. DOI: 10.1002/dneu.22282[26]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5033011/ (visited on 06/17/2020).

Chandrasekar, Vijay and Jean-Luc Dreyer (Nov. 2009). "microRNAs miR-124, let-7d and miR-181a regulate Cocaine-induced Plasticity". en. In: *Molecular and Cellular Neuroscience* 42.4, pp. 350–362. ISSN: 1044-7431. DOI: 10.1016/j.mcn.2009.08.009[27]. URL: http://www.sciencedirect.com/science/article/pii/S1044743109001900 (visited on 06/17/2020).

Cheng, Jun, Franziska Metge, and Christoph Dieterich (Apr. 2016). "Specific identification and quantification of circular RNAs from sequencing data". en. In: *Bioinformatics* 32.7, pp. 1094–1096. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btv656[28]. URL: https://academic.oup.com/bioinformatics/article/32/7/1094/1743588 (visited on 07/15/2019).

Cherone, Jennifer M., Vjola Jorgji, and Christopher B. Burge (Nov. 2019). "Cotargeting among microRNAs in the brain". In: *Genome Research* 29.11, pp. 1791–1804. ISSN: 1088-9051. DOI: 10.1101/gr.249201.119[29]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6836737/ (visited on 06/03/2020).

Chi, Sung Wook et al. (July 2009). "Ago HITS-CLIP decodes miRNA-mRNA interaction maps". In: *Nature* 460.7254, pp. 479–486. ISSN: 0028-0836. DOI: 10.1038/nature08170[30]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2733940/ (visited on 06/14/2020).

Dobin, Alexander et al. (Jan. 2013). "STAR: ultrafast universal RNA-seq aligner". In: *Bioinformatics* 29.1, pp. 15–21. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/bts635[31]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3530905/ (visited on 06/19/2020).

Doench, John G. and Phillip A. Sharp (Mar. 2004). "Specificity of microRNA target selection in translational repression". In: *Genes & Development* 18.5, pp. 504–511. ISSN: 0890-9369. DOI: 10.1101/gad.1184404[32]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC374233/ (visited on 06/15/2020).

Dongen, Stijn van, Cei Abreu-Goodger, and Anton J. Enright (Dec. 2008). "Detecting microRNA binding and siRNA off-target effects from expression data". en. In: *Nature Methods* 5.12. Number: 12 Publisher: Nature Publishing Group, pp. 1023–1025. ISSN: 1548-7105. DOI: 10.1038/nmeth.1267[33]. URL: https://www.nature.com/articles/nmeth.1267 (visited on 06/16/2020).

Doxakis, Epaminondas (Apr. 2010). "Post-transcriptional Regulation of -Synuclein Expression by mir-7 and mir-153". In: *The Journal of Biological Chemistry* 285.17, pp. 12726–12734. ISSN: 0021-9258. DOI:

---

[23]https://doi.org/10.1093/hmg/ddn005
[24]https://doi.org/10.1101/2020.01.21.913483
[25]https://doi.org/10.1371/journal.pone.0144129
[26]https://doi.org/10.1002/dneu.22282
[27]https://doi.org/10.1016/j.mcn.2009.08.009
[28]https://doi.org/10.1093/bioinformatics/btv656
[29]https://doi.org/10.1101/gr.249201.119
[30]https://doi.org/10.1038/nature08170
[31]https://doi.org/10.1093/bioinformatics/bts635
[32]https://doi.org/10.1101/gad.1184404
[33]https://doi.org/10.1038/nmeth.1267

10.1074/jbc.M109.086827[34]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2857101/ (visited on 06/11/2020).

Dwivedi, Y. et al. (Nov. 2015). "Chronic corticosterone-mediated dysregulation of microRNA network in prefrontal cortex of rats: relevance to depression pathophysiology". en. In: *Translational Psychiatry* 5.11. Number: 11 Publisher: Nature Publishing Group, e682–e682. ISSN: 2158-3188. DOI: 10.1038/tp.2015.175[35]. URL: https://www.nature.com/articles/tp2015175 (visited on 06/17/2020).

Ebert, Margaret S, Joel R Neilson, and Phillip A Sharp (Sept. 2007). "MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells". In: *Nature methods* 4.9. ISSN: 1548-7091. DOI: 10.1038/nmeth1079[36]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3857099/ (visited on 06/15/2020).

Fan, Yannan et al. (July 2016). "miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis". en. In: *Nucleic Acids Research* 44.W1. Publisher: Oxford Academic, W135–W141. ISSN: 0305-1048. DOI: 10.1093/nar/gkw288[37]. URL: https://academic.oup.com/nar/article/44/W1/W135/2499313 (visited on 06/16/2020).

Gaidatzis, Dimos et al. (July 2015). "Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation". en. In: *Nature Biotechnology* 33.7. Number: 7 Publisher: Nature Publishing Group, pp. 722–729. ISSN: 1546-1696. DOI: 10.1038/nbt.3269[38]. URL: https://www.nature.com/articles/nbt.3269 (visited on 06/13/2020).

Gray, Laura J. et al. (July 2010). "Region and diagnosis-specific changes in synaptic proteins in schizophrenia and bipolar I disorder". en. In: *Psychiatry Research* 178.2, pp. 374–380. ISSN: 0165-1781. DOI: 10.1016/j.psychres.2008.07.012[39]. URL: http://www.sciencedirect.com/science/article/pii/S0165178108002321 (visited on 06/11/2020).

Grimson, Andrew et al. (July 2007). "MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing". en. In: *Molecular Cell* 27.1, pp. 91–105. ISSN: 1097-2765. DOI: 10.1016/j.molcel.2007.06.017[40]. URL: http://www.sciencedirect.com/science/article/pii/S1097276507004078 (visited on 06/18/2020).

Grosswendt, Stefanie et al. (June 2014). "Unambiguous Identification of miRNA:target site Interactions by Different Types of Ligation Reactions". In: *Molecular cell* 54.6, pp. 1042–1054. ISSN: 1097-2765. DOI: 10.1016/j.molcel.2014.03.049[41]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4181535/ (visited on 06/14/2020).

Gu, Jin and Zhenyu Xuan (Nov. 2013). "Inferring the Perturbed microRNA Regulatory Networks in Cancer Using Hierarchical Gene Co-Expression Signatures". In: *PLoS ONE* 8.11. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0081032[42]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3835731/ (visited on 06/16/2020).

Guo, Yanwen et al. (Feb. 2015). "Characterization of the mammalian miRNA turnover landscape". en. In: *Nucleic Acids Research* 43.4. Publisher: Oxford Academic, pp. 2326–2341. ISSN: 0305-1048. DOI: 10.1093/nar/gkv057[43]. URL: https://academic.oup.com/nar/article/43/4/2326/2411404 (visited on 06/18/2020).

Hansen, Thomas B et al. (Nov. 2011). "miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA". In: *The EMBO Journal* 30.21, pp. 4414–4422. ISSN: 0261-4189. DOI: 10.1038/emboj.2011.359[44]. URL: https://www.embopress.org/doi/10.1038/emboj.2011.359 (visited on 10/21/2019).

Hansen, Thomas B. et al. (Mar. 2013). "Natural RNA circles function as efficient microRNA sponges". en. In: *Nature* 495.7441. Number: 7441 Publisher: Nature Publishing Group, pp. 384–388. ISSN: 1476-4687. DOI: 10.1038/nature11993[45]. URL: https://www.nature.com/articles/nature11993 (visited on 06/16/2020).

[34]https://doi.org/10.1074/jbc.M109.086827
[35]https://doi.org/10.1038/tp.2015.175
[36]https://doi.org/10.1038/nmeth1079
[37]https://doi.org/10.1093/nar/gkw288
[38]https://doi.org/10.1038/nbt.3269
[39]https://doi.org/10.1016/j.psychres.2008.07.012
[40]https://doi.org/10.1016/j.molcel.2007.06.017
[41]https://doi.org/10.1016/j.molcel.2014.03.049
[42]https://doi.org/10.1371/journal.pone.0081032
[43]https://doi.org/10.1093/nar/gkv057
[44]https://doi.org/10.1038/emboj.2011.359
[45]https://doi.org/10.1038/nature11993

He, Miao et al. (Jan. 2012). "Cell-type based analysis of microRNA profiles in the mouse brain". In: *Neuron* 73.1, pp. 35–48. ISSN: 0896-6273. DOI: 10.1016/j.neuron.2011.11.010[46]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3270494/ (visited on 06/14/2020).

Hicks, Steven D. and Frank A. Middleton (2016). "A Comparative Review of microRNA Expression Patterns in Autism Spectrum Disorder". English. In: *Frontiers in Psychiatry* 7. Publisher: Frontiers. ISSN: 1664-0640. DOI: 10.3389/fpsyt.2016.00176[47]. URL: https://www.frontiersin.org/articles/10.3389/fpsyt.2016.00176/full (visited on 06/12/2020).

Hu, Zhonghua et al. (Aug. 2019). "Temporal dynamics of miRNAs in human DLPFC and its association with miRNA dysregulation in schizophrenia". In: *Translational Psychiatry* 9. ISSN: 2158-3188. DOI: 10.1038/s41398-019-0538-y[48]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6702224/ (visited on 06/12/2020).

Jeong, Geon et al. (Dec. 2016). "Knockout of miR-221 and miR-222 reveals common and specific targets for paralogous miRNAs". In: *RNA Biology* 14.2, pp. 197–205. ISSN: 1547-6286. DOI: 10.1080/15476286.2016.1269994[49]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5324760/ (visited on 06/03/2020).

Kingston, Elena R. and David P. Bartel (Apr. 2019). "Global analyses of the dynamics of mammalian microRNA metabolism". en. In: *bioRxiv*. Publisher: Cold Spring Harbor Laboratory Section: New Results, p. 607150. DOI: 10.1101/607150[50]. URL: https://www.biorxiv.org/content/10.1101/607150v1 (visited on 06/18/2020).

Krol, Jacek et al. (May 2010). "Characterizing Light-Regulated Retinal MicroRNAs Reveals Rapid Turnover as a Common Property of Neuronal MicroRNAs". en. In: *Cell* 141.4, pp. 618–631. ISSN: 0092-8674. DOI: 10.1016/j.cell.2010.03.039[51]. URL: http://www.sciencedirect.com/science/article/pii/S0092867410003570 (visited on 06/15/2020).

Kudla, Grzegorz et al. (June 2011). "Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast". In: *Proceedings of the National Academy of Sciences of the United States of America* 108.24, pp. 10010–10015. ISSN: 0027-8424. DOI: 10.1073/pnas.1017386108[52]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3116431/ (visited on 06/14/2020).

Lee, Byunghan et al. (Aug. 2016). "deepTarget: End-to-end Learning Framework for microRNA Target Prediction using Deep Recurrent Neural Networks". en. In: *arXiv:1603.09123 [cs, q-bio]*. arXiv: 1603.09123. URL: http://arxiv.org/abs/1603.09123 (visited on 06/16/2020).

Lee, Rosalind C., Rhonda L. Feinbaum, and Victor Ambros (Dec. 1993). "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14". en. In: *Cell* 75.5, pp. 843–854. ISSN: 0092-8674. DOI: 10.1016/0092-8674(93)90529-Y[53]. URL: http://www.sciencedirect.com/science/article/pii/009286749390529Y (visited on 06/15/2020).

Lee, Ko-Woon et al. (Feb. 2006). "Cocaine-induced dendritic spine formation in D1 and D2 dopamine receptor-containing medium spiny neurons in nucleus accumbens". en. In: *Proceedings of the National Academy of Sciences* 103.9. Publisher: National Academy of Sciences Section: Biological Sciences, pp. 3399–3404. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0511244103[54]. URL: https://www.pnas.org/content/103/9/3399 (visited on 06/17/2020).

Lewis, Benjamin P., Christopher B. Burge, and David P. Bartel (Jan. 2005). "Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets". en. In: *Cell* 120.1, pp. 15–20. ISSN: 0092-8674. DOI: 10.1016/j.cell.2004.12.035[55]. URL: http://www.sciencedirect.com/science/article/pii/S0092867404012607 (visited on 06/15/2020).

---

[46] https://doi.org/10.1016/j.neuron.2011.11.010
[47] https://doi.org/10.3389/fpsyt.2016.00176
[48] https://doi.org/10.1038/s41398-019-0538-y
[49] https://doi.org/10.1080/15476286.2016.1269994
[50] https://doi.org/10.1101/607150
[51] https://doi.org/10.1016/j.cell.2010.03.039
[52] https://doi.org/10.1073/pnas.1017386108
[53] https://doi.org/10.1016/0092-8674(93)90529-Y
[54] https://doi.org/10.1073/pnas.0511244103
[55] https://doi.org/10.1016/j.cell.2004.12.035

Lewis, Benjamin P., I-hung Shih, et al. (Dec. 2003). "Prediction of Mammalian MicroRNA Targets". en. In: *Cell* 115.7, pp. 787–798. ISSN: 0092-8674. DOI: 10.1016/S0092-8674(03)01018-3[56]. URL: http://www.sciencedirect.com/science/article/pii/S0092867403010183 (visited on 06/16/2020).

Liu, Wenjie and Joseph Irudayaraj (Dec. 2019). "Understanding the dynamics and structure of epigenetic states with single-molecule fluorescence microscopy". en. In: *Current Opinion in Biomedical Engineering*. Molecular & CellularEngineering: single molecule technology Neural Engineering: High Resolution Cell Imaging 12, pp. 18–24. ISSN: 2468-4511. DOI: 10.1016/j.cobme.2019.08.010[57]. URL: http://www.sciencedirect.com/science/article/pii/S2468451119300303 (visited on 06/16/2020).

Liu, Yang et al. (Sept. 2013). "microRNA-181a is involved in insulin-like growth factor-1-mediated regulation of the transcription factor CREB1". eng. In: *Journal of Neurochemistry* 126.6, pp. 771–780. ISSN: 1471-4159. DOI: 10.1111/jnc.12370[58].

Llave, Cesar et al. (Sept. 2002). "Cleavage of Scarecrow-like mRNA Targets Directed by a Class of Arabidopsis miRNA". en. In: *Science* 297.5589. Publisher: American Association for the Advancement of Science Section: Report, pp. 2053–2056. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1076311[59]. URL: https://science.sciencemag.org/content/297/5589/2053 (visited on 06/18/2020).

Ma, Jin-Biao et al. (Mar. 2005). "Structural basis for 5 -end-specific recognition of guide RNA by the A. fulgidus Piwi protein". en. In: *Nature* 434.7033. Number: 7033 Publisher: Nature Publishing Group, pp. 666–670. ISSN: 1476-4687. DOI: 10.1038/nature03514[60]. URL: https://www.nature.com/articles/nature03514 (visited on 06/18/2020).

McCall, Matthew N. et al. (Jan. 2017). "Toward the human cellular microRNAome". en. In: *Genome Research* 27.10. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab, pp. 1769–1781. ISSN: 1088-9051, 1549-5469. DOI: 10.1101/gr.222067.117[61]. URL: http://genome.cshlp.org/content/27/10/1769 (visited on 06/19/2020).

McGeary, Sean E., Kathy S. Lin, Charlie Y. Shi, Namita Bisaria, et al. (Sept. 2018). "The biochemical basis of microRNA targeting efficacy". en. In: *bioRxiv*, p. 414763. DOI: 10.1101/414763[62]. URL: https://www.biorxiv.org/content/10.1101/414763v1 (visited on 08/09/2019).

McGeary, Sean E., Kathy S. Lin, Charlie Y. Shi, Thy M. Pham, et al. (Dec. 2019). "The biochemical basis of microRNA targeting efficacy". en. In: *Science* 366.6472. Publisher: American Association for the Advancement of Science Section: Research Article. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aav1741[63]. URL: https://science.sciencemag.org/content/366/6472/eaav1741 (visited on 06/03/2020).

Millar, Anthony A. and Peter M. Waterhouse (July 2005). "Plant and animal microRNAs: similarities and differences". en. In: *Functional & Integrative Genomics* 5.3, pp. 129–135. ISSN: 1438-7948. DOI: 10.1007/s10142-005-0145-2[64]. URL: https://doi.org/10.1007/s10142-005-0145-2 (visited on 06/18/2020).

Mukherji, Shankar et al. (Sept. 2011). "MicroRNAs can generate thresholds in target gene expression". en. In: *Nature Genetics* 43.9. Number: 9 Publisher: Nature Publishing Group, pp. 854–859. ISSN: 1546-1718. DOI: 10.1038/ng.905[65]. URL: https://www.nature.com/articles/ng.905 (visited on 06/15/2020).

Nestler, Eric J. (Feb. 2001). "Molecular basis of long-term plasticity underlying addiction". en. In: *Nature Reviews Neuroscience* 2.2. Number: 2 Publisher: Nature Publishing Group, pp. 119–128. ISSN: 1471-0048. DOI: 10.1038/35053570[66]. URL: https://www.nature.com/articles/35053570 (visited on 06/17/2020).

Nicolas, Francisco E. (2011). "Experimental Validation of MicroRNA Targets Using a Luciferase Reporter System". en. In: *MicroRNAs in Development: Methods and Protocols*. Ed. by Tamas Dalmay. Methods in

---

[56]https://doi.org/10.1016/S0092-8674(03)01018-3
[57]https://doi.org/10.1016/j.cobme.2019.08.010
[58]https://doi.org/10.1111/jnc.12370
[59]https://doi.org/10.1126/science.1076311
[60]https://doi.org/10.1038/nature03514
[61]https://doi.org/10.1101/gr.222067.117
[62]https://doi.org/10.1101/414763
[63]https://doi.org/10.1126/science.aav1741
[64]https://doi.org/10.1007/s10142-005-0145-2
[65]https://doi.org/10.1038/ng.905
[66]https://doi.org/10.1038/35053570

Molecular Biology. Totowa, NJ: Humana Press, pp. 139–152. ISBN: 978-1-61779-083-6. DOI: 10.1007/978-1-61779-083-6_11[67]. URL: https://doi.org/10.1007/978-1-61779-083-6_11 (visited on 06/18/2020).

O'Brien, Jacob et al. (Aug. 2018). "Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation". In: *Frontiers in Endocrinology* 9. ISSN: 1664-2392. DOI: 10.3389/fendo.2018.00402[68]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6085463/ (visited on 06/18/2020).

Oliveira, Arthur C. et al. (2017). "Combining Results from Distinct MicroRNA Target Prediction Tools Enhances the Performance of Analyses". English. In: *Frontiers in Genetics* 8. Publisher: Frontiers. ISSN: 1664-8021. DOI: 10.3389/fgene.2017.00059[69]. URL: https://www.frontiersin.org/articles/10.3389/fgene.2017.00059/full (visited on 06/14/2020).

Ouyang, Yi-Bing et al. (Jan. 2012). "miR-181 regulates GRP78 and influences outcome from cerebral ischemia in vitro and in vivo". In: *Neurobiology of disease* 45.1, pp. 555–563. ISSN: 0969-9961. DOI: 10.1016/j.nbd.2011.09.012[70]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3251314/ (visited on 06/19/2020).

Qiao, Jing et al. (Feb. 2020). "MicroRNA-153 improves the neurogenesis of neural stem cells and enhances the cognitive ability of aged mice through the notch signaling pathway". en. In: *Cell Death & Differentiation* 27.2. Number: 2 Publisher: Nature Publishing Group, pp. 808–825. ISSN: 1476-5403. DOI: 10.1038/s41418-019-0388-4[71]. URL: https://www.nature.com/articles/s41418-019-0388-4 (visited on 06/11/2020).

Rajman, Marek et al. (June 2017). "A microRNA-129-5p/Rbfox crosstalk coordinates homeostatic downscaling of excitatory synapses". In: *The EMBO Journal* 36.12, pp. 1770–1787. ISSN: 0261-4189. DOI: 10.15252/embj.201695748[72]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5470041/ (visited on 06/16/2020).

Reinhart, Brenda J. et al. (Feb. 2000). "The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans". en. In: *Nature* 403.6772. Number: 6772 Publisher: Nature Publishing Group, pp. 901–906. ISSN: 1476-4687. DOI: 10.1038/35002607[73]. URL: https://www.nature.com/articles/35002607 (visited on 06/15/2020).

Robinson, Mark D., Davis J. McCarthy, and Gordon K. Smyth (Jan. 2010). "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data". In: *Bioinformatics* 26.1, pp. 139–140. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btp616[74]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2796818/ (visited on 06/19/2020).

Rybak-Wolf, Agnieszka et al. (June 2015a). "Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed". en. In: *Molecular Cell* 58.5, pp. 870–885. ISSN: 1097-2765. DOI: 10.1016/j.molcel.2015.03.027[75]. URL: http://www.sciencedirect.com/science/article/pii/S109727651500218X (visited on 06/16/2020).

– (June 2015b). "Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed". en. In: *Molecular Cell* 58.5, pp. 870–885. ISSN: 1097-2765. DOI: 10.1016/j.molcel.2015.03.027[76]. URL: http://www.sciencedirect.com/science/article/pii/S109727651500218X (visited on 06/16/2020).

Saba, Reuben et al. (Feb. 2012). "Dopamine-Regulated MicroRNA MiR-181a Controls GluA2 Surface Expression in Hippocampal Neurons". en. In: *Molecular and Cellular Biology* 32.3. Publisher: American Society for Microbiology Journals Section: Articles, pp. 619–632. ISSN: 0270-7306, 1098-5549. DOI: 10.1128/MCB.05896-11[77]. URL: https://mcb.asm.org/content/32/3/619 (visited on 06/17/2020).

Schäfer, Moritz and Constance Ciaudo (Jan. 2020). "Prediction of the miRNA interactome – Established methods and upcoming perspectives". en. In: *Computational and Structural Biotechnology Journal* 18,

---

[67]https://doi.org/10.1007/978-1-61779-083-6_11
[68]https://doi.org/10.3389/fendo.2018.00402
[69]https://doi.org/10.3389/fgene.2017.00059
[70]https://doi.org/10.1016/j.nbd.2011.09.012
[71]https://doi.org/10.1038/s41418-019-0388-4
[72]https://doi.org/10.15252/embj.201695748
[73]https://doi.org/10.1038/35002607
[74]https://doi.org/10.1093/bioinformatics/btp616
[75]https://doi.org/10.1016/j.molcel.2015.03.027
[76]https://doi.org/10.1016/j.molcel.2015.03.027
[77]https://doi.org/10.1128/MCB.05896-11

pp. 548–557. ISSN: 2001-0370. DOI: 10.1016/j.csbj.2020.02.019[78]. URL: http://www.sciencedirect.com/science/article/pii/S2001037019304301 (visited on 06/16/2020).

Schirle, Nicole T et al. (2020). "Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets". In: *eLife* 4 (). ISSN: 2050-084X. DOI: 10.7554/eLife.07646[79]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4606517/ (visited on 06/15/2020).

Schnall-Levin, Michael et al. (Sept. 2010). "Conserved microRNA targeting in Drosophila is as widespread in coding regions as in 3 UTRs". In: *Proceedings of the National Academy of Sciences of the United States of America* 107.36, pp. 15751–15756. ISSN: 0027-8424. DOI: 10.1073/pnas.1006172107[80]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2936641/ (visited on 06/18/2020).

Schwarz, Dianne S. et al. (Oct. 2003). "Asymmetry in the Assembly of the RNAi Enzyme Complex". en. In: *Cell* 115.2, pp. 199–208. ISSN: 0092-8674. DOI: 10.1016/S0092-8674(03)00759-1[81]. URL: http://www.sciencedirect.com/science/article/pii/S0092867403007591 (visited on 06/18/2020).

Shao, Yiye and Yinghui Chen (Apr. 2016). "Roles of Circular RNAs in Neurologic Disease". In: *Frontiers in Molecular Neuroscience* 9. ISSN: 1662-5099. DOI: 10.3389/fnmol.2016.00025[82]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4829598/ (visited on 06/16/2020).

Stark, Alexander et al. (Nov. 2007). "Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures". In: *Nature* 450.7167, pp. 219–232. ISSN: 0028-0836. DOI: 10.1038/nature06340[83]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2474711/ (visited on 06/15/2020).

Stark, Kimberly L. et al. (June 2008). "Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model". en. In: *Nature Genetics* 40.6. Number: 6 Publisher: Nature Publishing Group, pp. 751–760. ISSN: 1546-1718. DOI: 10.1038/ng.138[84]. URL: https://www.nature.com/articles/ng.138 (visited on 06/16/2020).

Steinfeld, Israel et al. (Feb. 2013). "miRNA target enrichment analysis reveals directly active miRNAs in health and disease". In: *Nucleic Acids Research* 41.3, e45. ISSN: 0305-1048. DOI: 10.1093/nar/gks1142[85]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3561970/ (visited on 06/16/2020).

Tsuyama, Jun et al. (July 2015). "MicroRNA-153 Regulates the Acquisition of Gliogenic Competence by Neural Stem Cells". In: *Stem Cell Reports* 5.3, pp. 365–377. ISSN: 2213-6711. DOI: 10.1016/j.stemcr.2015.06.006[86]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4618452/ (visited on 06/11/2020).

Ulitsky, Igor, Louise C. Laurent, and Ron Shamir (Aug. 2010). "Towards computational prediction of microRNA function and activity". In: *Nucleic Acids Research* 38.15, e160. ISSN: 0305-1048. DOI: 10.1093/nar/gkq570[87]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2926627/ (visited on 06/16/2020).

Valencia-Sanchez, Marco Antonio et al. (Jan. 2006). "Control of translation and mRNA degradation by miRNAs and siRNAs". en. In: *Genes & Development* 20.5. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab, pp. 515–524. ISSN: 0890-9369, 1549-5477. DOI: 10.1101/gad.1399806[88]. URL: http://genesdev.cshlp.org/content/20/5/515 (visited on 06/18/2020).

Vasudevan, Shobha (May 2012). "Posttranscriptional Upregulation by MicroRNAs". In: *WIREs RNA* 3.3. Publisher: John Wiley & Sons, Ltd, pp. 311–330. ISSN: 1757-7004. DOI: 10.1002/wrna.121[89]. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/wrna.121 (visited on 06/18/2020).

---

[78] https://doi.org/10.1016/j.csbj.2020.02.019

[79] https://doi.org/10.7554/eLife.07646

[80] https://doi.org/10.1073/pnas.1006172107

[81] https://doi.org/10.1016/S0092-8674(03)00759-1

[82] https://doi.org/10.3389/fnmol.2016.00025

[83] https://doi.org/10.1038/nature06340

[84] https://doi.org/10.1038/ng.138

[85] https://doi.org/10.1093/nar/gks1142

[86] https://doi.org/10.1016/j.stemcr.2015.06.006

[87] https://doi.org/10.1093/nar/gkq570

[88] https://doi.org/10.1101/gad.1399806

[89] https://doi.org/10.1002/wrna.121

Venø, Morten T. et al. (Nov. 2015). "Spatio-temporal regulation of circular RNA expression during porcine embryonic brain development". en. In: *Genome Biology* 16.1, p. 245. ISSN: 1474-760X. DOI: 10.1186/s13059-015-0801-3[90]. URL: https://doi.org/10.1186/s13059-015-0801-3 (visited on 06/16/2020).

Vlachos, Ioannis S. et al. (July 2016). "DIANA-mirExTra v2.0: Uncovering microRNAs and transcription factors with crucial roles in NGS expression data". en. In: *Nucleic Acids Research* 44.W1. Publisher: Oxford Academic, W128–W134. ISSN: 0305-1048. DOI: 10.1093/nar/gkw455[91]. URL: https://academic.oup.com/nar/article/44/W1/W128/2499390 (visited on 06/16/2020).

Wei, Chunyao et al. (Feb. 2013). "miR-153 Regulates SNAP-25, Synaptic Transmission, and Neuronal Development". In: *PLoS ONE* 8.2. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0057080[92]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3581580/ (visited on 06/11/2020).

Wingo, Aliza P. et al. (May 2017). "Genome-wide association study of positive emotion identifies a genetic variant and a role for microRNAs". In: *Molecular psychiatry* 22.5, pp. 774–783. ISSN: 1359-4184. DOI: 10.1038/mp.2016.143[93]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5339071/ (visited on 06/17/2020).

Wu, Ligang, Jihua Fan, and Joel G. Belasco (Mar. 2006). "MicroRNAs direct rapid deadenylation of mRNA". In: *Proceedings of the National Academy of Sciences of the United States of America* 103.11, pp. 4034–4039. ISSN: 0027-8424. DOI: 10.1073/pnas.0510928103[94]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1449641/ (visited on 06/18/2020).

Xu, Li-Jun et al. (Feb. 2015). "Post-stroke treatment with miR-181 antagomir reduces injury and improves long-term behavioral recovery in mice after focal cerebral ischemia". en. In: *Experimental Neurology* 264, pp. 1–7. ISSN: 0014-4886. DOI: 10.1016/j.expneurol.2014.11.007[95]. URL: http://www.sciencedirect.com/science/article/pii/S0014488614003677 (visited on 06/19/2020).

You, Yu-Hui et al. (June 2019). "MicroRNA-153 promotes brain-derived neurotrophic factor and hippocampal neuron proliferation to alleviate autism symptoms through inhibition of JAK-STAT pathway by LEPR". In: *Bioscience Reports* 39.6. ISSN: 0144-8463. DOI: 10.1042/BSR20181904[96]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6591574/ (visited on 06/11/2020).

You, Xintian et al. (Apr. 2015). "Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity". en. In: *Nature Neuroscience* 18.4. Number: 4 Publisher: Nature Publishing Group, pp. 603–610. ISSN: 1546-1726. DOI: 10.1038/nn.3975[97]. URL: https://www.nature.com/articles/nn.3975 (visited on 06/16/2020).

Young, Matthew D et al. (2010). "Gene ontology analysis for RNA-seq: accounting for selection bias". In: *Genome Biology* 11.2, R14. ISSN: 1465-6906. DOI: 10.1186/gb-2010-11-2-r14[98]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2872874/ (visited on 06/19/2020).

Zeisel, Amit et al. (Sept. 2011). "Coupled pre-mRNA and mRNA dynamics unveil operational strategies underlying transcriptional responses to stimuli". In: *Molecular Systems Biology* 7, p. 529. ISSN: 1744-4292. DOI: 10.1038/msb.2011.62[99]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3202801/ (visited on 06/13/2020).

Zhang, Kai et al. (Nov. 2018). "A novel class of microRNA-recognition elements that function only within open reading frames". eng. In: *Nature Structural & Molecular Biology* 25.11, pp. 1019–1027. ISSN: 1545-9985. DOI: 10.1038/s41594-018-0136-3[100].

Zhang, Xiao-Ou et al. (Jan. 2016). "Diverse alternative back-splicing and alternative splicing landscape of circular RNAs". en. In: *Genome Research* 26.9. Company: Cold Spring Harbor Laboratory Press Distributor:

---

[90]https://doi.org/10.1186/s13059-015-0801-3
[91]https://doi.org/10.1093/nar/gkw455
[92]https://doi.org/10.1371/journal.pone.0057080
[93]https://doi.org/10.1038/mp.2016.143
[94]https://doi.org/10.1073/pnas.0510928103
[95]https://doi.org/10.1016/j.expneurol.2014.11.007
[96]https://doi.org/10.1042/BSR20181904
[97]https://doi.org/10.1038/nn.3975
[98]https://doi.org/10.1186/gb-2010-11-2-r14
[99]https://doi.org/10.1038/msb.2011.62
[100]https://doi.org/10.1038/s41594-018-0136-3

Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab, pp. 1277–1287. ISSN: 1088-9051, 1549-5469. DOI: 10.1101/gr.202895.115[101]. URL: http://genome.cshlp.org/content/26/9/1277 (visited on 06/19/2020).

Zhao, Yan et al. (Sept. 2016). "Decreased Expression of Plasma MicroRNA in Patients with Methamphetamine (MA) Use Disorder". en. In: *Journal of Neuroimmune Pharmacology* 11.3, pp. 542–548. ISSN: 1557-1904. DOI: 10.1007/s11481-016-9671-z[102]. URL: https://doi.org/10.1007/s11481-016-9671-z (visited on 06/17/2020).

Zhou, Juhua et al. (Apr. 2014). "Dysregulation in microRNA Expression Is Associated with Alterations in Immune Functions in Combat Veterans with Post-Traumatic Stress Disorder". In: *PLoS ONE* 9.4. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0094075[103]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3997344/ (visited on 06/17/2020).

---

[101] https://doi.org/10.1101/gr.202895.115
[102] https://doi.org/10.1007/s11481-016-9671-z
[103] https://doi.org/10.1371/journal.pone.0094075