Lance Middleton
BIOC218 Final
06/02/2013

A critical review of MuTaME scoring for ceRNA predictions.

Introduction:

Since their discovery, microRNAs (miRs) have been implicated in many human diseases [5,12,18,19,57]. For review, miRs are evolutionarily conserved regulators [26] of mRNA translation and stability [16,40,41,58]. These short (~21 nucleotide) RNAs are transcribed by RNA polymerase II [22] as longer precursors (pre-miRs), which form characteristic hairpin structures (primary miRs, or pri-miRs) by base pair complementarity [24]. These pri-miRNAs are cleaved within the nucleus by nuclear RNAse III Drosha [23] to yield individual hairpins available for export to the cytosol by exportin-5 [59]. In the cytosol another RNAse III, Dicer, removes the end of the hairpin structure [17] to make mature miRs. These mature miRs are then integrated into RNA-Induced Silencing Complexes (RISCs) [14], which target mRNAs for silencing by complementary base pairing.

Individual miRs regulate hundreds to thousands of target mRNAs [25,27]. An interesting corollary of this shared regulation emerged recently when it was shown that several transcripts help regulate tumor suppressor PTEN expression levels by competing for a shared pool of regulatory miRs [52]. Specifically, these authors demonstrated that transcripts termed competing endogenous RNAs (ceRNAs) featured the same miR binding sites, called miR Response Elements (MREs), in their 3' un-translated regions (3'UTRs) as PTEN and thus, could titrate regulatory miRs away from the PTEN transcript, allowing the translation of PTEN protein. Their work was the first proof of principle that protein-coding transcripts have a secondary layer of functionality as trans-regulators of the expression of similar mRNAs. This idea has broad implications in cancer where the expression of genes can be greatly inflated or repressed to help neoplastic cells leverage or ablate biological pathways for survival [47]. In this review, we will explore the "mutually targeted MRE enrichment" (MuTaME) prediction approach put forth by Tay et al in order to predict likely ceRNAs. We will examine how MuTaME predictions work and how this approach might be improved with currently available information.

Experimental Prelude:

Prior to computing MuTaME scores, Tay et al had a set of 10 miRs validated to target PTEN. For scientists interested in identifying ceRNAs, a few experimental steps must be taken before MuTaME scores can be generated. These include miR profiling and validation of miR/mRNA physical associations. Following miR extraction from tissue, there are three widely-used approaches for miR expression profiling. These include microarray profiling, real-time PCR, and next generation sequencing. Here we explore these popular means of pre-computational validation. For brevity, we will only touch on secondary validation approaches that establish the physical associations between miRs and mRNAs.

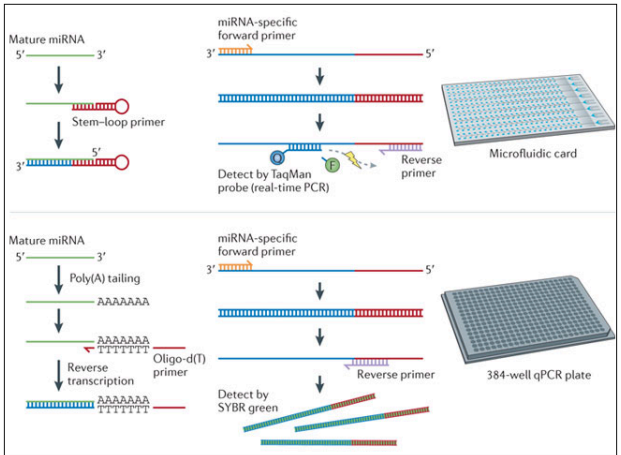

Figure 1. qRT-PCR. Adapted from "MicroRNA profiling: approaches and considerations" by Pritchard C.C. et al, 2012, Nat Revs Genet 13, 358-369.

miR Profiling:

1. qRT-PCR – Quantitative reverse transcription is a highly quantitative miR profiling method in common use. Similar to older reverse transcription methods used to convert mRNA into cDNA, this process uses miRs as the reverse transcription template to make cDNAs (figure 1). However, unlike methods that generate cDNAs from longer mRNAs, this process utilizes unique reverse transcription primers that recognize miRs and add short, additional sequences to the cDNA product which are targeted for the priming of the polymerase chain reaction (PCR). These RT primers typically come as either poly-a-extended [48] or stem-loop primers [4]. After cDNA synthesis, PCR proceeds normally although the RT primer used will influence the fluorescent chemistry used in the quantification of cDNA. Poly-a-extension primers render cDNAs amenable to double stranded DNA detection with SYBR green reagents

whereas stem-loop primers generate cDNAs that can be detected by hydrolysable fluorescent probes. These reactions are carried out at medium throughput using microfluidic cards or qPCR plates that come pre-seeded with RT primers and are amenable to automated technologies. As with conventional RT-PCR, cDNA production can be quantified by standard curves or by comparative CT methods that calculate the difference in threshold cycles for target and control samples [28]. In this way, data are normalized to the expression miRs that are constitutively expressed [39]. However, data can also be normalized by the global mean expression of all miRs [34]. After processing, the investigator is able to determine which miRs are being expressed in a particular tissue and in what relative quantities.

2. miR Microarrays – Hybridization-based microarray technologies are among the oldest and most common means of profiling miR expression [20,53]. Similar to mRNA microarrays, this approach requires labeling miRs with detectable fluorescent markers that can then be hybridized with complementary, oligonucleotide arrays (figure 2). Several labeling chemistries are available [50,56,] and differ based on the commercial platform in use. From fluorescent intensity measurements across the (sample) x (miR) matrix, relative changes between normal control samples and experimental samples can be determined once the data are normalized. However, it should be noted that there are a plethora of normalization approaches used for miR microarray data. Methods currently in practice include scaling [37], invariant-based regression [11], median-centering [51], quartile normalization [21], and variance-stabilizing normalization [6]. Notably, each of these normalization techniques relies on a unique set of assumptions about the origin of the variability in the fluorescent measurements (i.e., what is true biological variation between normal controls and experimental replicates and what is random, “noisy” variation) [42]. After processing, the investigator is able to determine which miRs are being differentially expressed between two biological materials in a relatively cheap, high-throughput fashion.

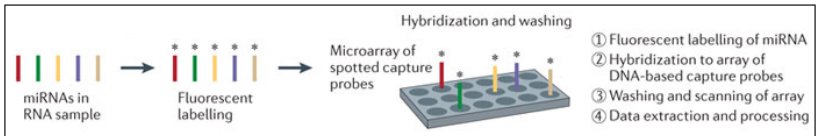

Figure 2. miR Microarray. Adapted from "MicroRNA profiling: approaches and considerations" by Pritchard C.C. et al, 2012, Nat Revs Genet 13, 358-369.

3. miR-seq – miRs are amenable to next-generation sequencing technologies just like the rest of the transcriptome [38]. As such, a plethora of sequencing platforms and chemistries are available for use though, for our purposes, we will mainly only consider two. These technologies utilize a common workflow that features template preparation, sequencing, and assembly/alignment [Reviewed in 43] (figure 3). miRs are converted to cDNA by reverse transcription as with RT-PCR and sequencing adapters are ligated to the ends of each cDNA that serve as universal priming sites. These individual cDNA templates with adapters are then spatially separated from one another either by 1.) emulsion in an oil-based liquid where single templates are attached to single beads by complementary binding, or 2.) distribution across a surface array bound to complementary sequences. Sequestered templates then undergo standard PCR amplification via priming of the adapters, the products of which are amenable to massively parallel sequencing. Those fragments amplified on a solid-phase platform (ex: Illumina) are then sequenced by the synthesis of a complementary DNA strand. This happens in a process known as cyclic reversible termination where a DNA polymerase integrates a single, fluorescently-labeled nucleotide to a strand complementary to the PCR product [35]. The fluorescently labeled nucleotide is imaged and then its fluorescence is quenched. Repeated rounds of individual fluorescent nucleotide addition, measurement, and quenching occur until the entire sequence of the complementary strand is complete. Those fragments amplified in an emulsion reaction (ex: 454) are also sequenced by the synthesis of complementary DNA strands in a process termed pyrosequencing [46]. In this case, the integration of a new nucleotide to the growing complementary strand releases a pyrophosphate, which becomes enzymatically converted to light for detection. All sequences detected by either technology can then be aligned against reference miR sequences for identification and quantification. The abundance of individual miRs can then be determined by normalizing the number of sequencing reads for a unique miR to the total number of miR sequencing reads, expressed as

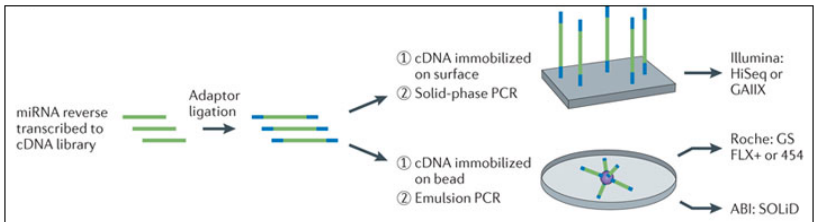

Figure 3. miR-seq. Adapted from "MicroRNA profiling: approaches and considerations" by Pritchard C.C. et al, 2012, Nat Revs Genet 13, 358-369.

a percentage of total sequencing reads [43]. After processing, the investigator is able to determine the relative abundance of each miR. Notably, some miRs may not align to known sequences, thus miR-seq enables investigators to identify new miRs.

Once an investigator has established the expression of a set of miRs they will need to validate the association of these miRs with a particular mRNA transcript of interest. Often this is conducted by Argonaute II immunoprecipitation [Karginov, 2007], which retrieves mRNAs bound to RNA-induced silencing complexes containing miRs of interest. Similar methods have been proposed where investigators overexpress and precipitate dominant negative RISC components via a so-called RISCtrap technique that captures miR/RISC/mRNA complexes of low abundance [2]. Others have proposed more classical cloning-based strategies whereby miR targeting can be confirmed by positive and negative selection for selectable markers [10]. Others still have confirmed miR targeting of mRNA transcripts indirectly by measuring the degradation of radiolabeled proteins after miR overexpression and translational repression [55]. Lastly, the most common method of validation, where a single miR/mRNA interaction is under investigation, is bioluminescence based. In this technique, forced expression of a construct bearing the 3'UTR of the mRNA of interest fused in frame with the coding sequence of firefly luciferase allows the investigator to determine if a miR of interest antagonizes the production of functional luciferase protein [7].

The rna22 target prediction algorithm:

Once an investigator has a gene of interest and a set of miRs validated to regulate its transcript, she can determine which other genes these miRs are likely to target. Because few miR:mRNA interactions have been validated, this step in the MuTaME approach requires the use of a miR target prediction algorithm. Tay et al choose rna22 (http://cbcsrv.watson.ibm.com/rna22.html) [36] to generate miR-binding predictions for the entire protein-coding transcriptome because it was shown by other authors to yield fewer false positive predictions than other algorithms [13,45]. Rna22 is unique among miR target prediction algorithms in that it relies on pattern-recognition of short sequences within known miRs (i.e., it doesn't rely on known miR:mRNA interactions) to determine likely miR targets. Here we will review the features of the rna22 target prediction algorithm.

First, rna22 utilizes the most current RFAM database release to compile known miRs. In doing so, rna22 filters out identical and near-identical miR sequences using the BLASTN algorithm [1]. Rna22 then attempts to discover patterns within the remaining, unique miRs. To discover sequence patterns of variable lengths, rna22 applies the Teiresias algorithm [44] to the unique miRs. During the scanning phase of Teiresias, given a set S = $\{s_1, s_2, \ldots, s_n\}$ miR sequences and parameters L,W, and K, Rna22 uses Teiresias to find all maximal <L,W> patterns that appear K times. These variables are described below:

1. L – This value is the minimum size of discovered patterns. Rna22 sets L= 4.
2. W – This value is the length of the event stream (length of sequence) wherein a pattern of size L will be discovered (L $\leq$ W). Rna22 sets W = 12.
3. K – This value is the number of instances a pattern must be observed before it can be reported. Rna22 sets K = 2.

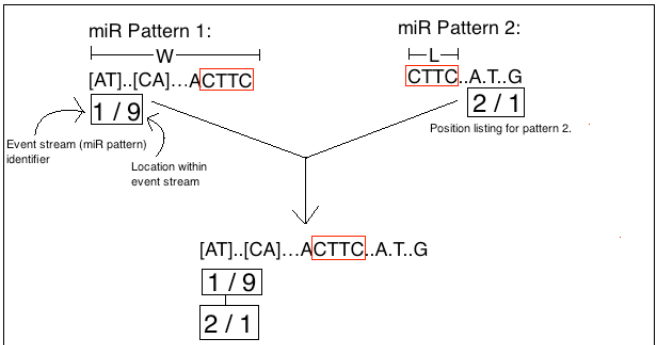

Figure 4. Convolution phase of Teiresias. Adapted from "The emergence of pattern discovery techniques in computational biology." by Rigoutsos et al. 2000. Metab Eng. 2, 159-177.

Next, Teiresias enters a convolution phase where it recursively combines patterns of minimum length (L) into increasingly longer patterns (figure 4). The K value, or "support" of the patterns decreases as longer patterns are constructed. To accomplish this task, Teiresias considers patterns two at a time for common "prefixes" and "suffixes" (i.e., common sequences at either the beginning or end of a pattern, respectively). When the suffix of

one pattern is found to be the same as the prefix of another pattern, Tieresias will consider, for each pattern, where in an event stream (i.e., miR sequence) the pattern occurs. The two patterns are then combined into a new pattern and the new pattern is given a position listing (location description, says where the pattern occurs and in what event stream) made from the two smaller, contributing patterns. Position lists of the contributing patterns are updated and the contributing patterns are returned for repeated rounds of convolution until they can no longer form larger patterns at a minimum of K frequency [36, supplement]. The rna22 algorithm then takes all conserved patterns that it identifies and applies a second-order Markov chain to estimate each pattern's statistical significance. Specifically, it estimates the frequency of nucleotide triplets that occur anywhere across 23 positions of sequence (the length of a miR). Based on the frequency of each triplet, Bayes' theorem can be applied to determine the probability that any pattern might be generated from a random database of patterns. Those patterns with an estimated log probability $\geq$ -38 are discarded [36].

With a number of miR patterns identified (233,554 at the time of the original publication [36]), rna22 can begin to search for sequences of reverse complementary to these patterns in the protein-coding transcriptome. According to the logic of rna22, if a protein-coding sequence contains one or more patterns complementary to those identified by the Teiresias algorithm in the 3'UTR, it is likely to be a miR target. This likelihood increases for a given protein-coding sequence as the number of identified patterns in the 3'UTR increases. Rna22 looks across the transcriptome for 3'UTRs that have reverse complementarity to the miR-patterns and also considers the probability (mentioned above) that those patterns were identified relative to chance. In this way, rna22 identifies "target islands," or regions of 3'UTRs where miR complementary patterns cluster. Those target islands that contain $\geq$ 30 reverse complement patterns are considered putative miR binding sites.

Once all target islands are identified for each transcript, rna22 determines the miR likely to bind each target. To accomplish this task, rna22 considers the target island sequence, a hypothetical hairpin linking sequence "GCGGGGACGC" [49], and a miR sequence. With these three sequences concatenated, rna22 applies the Vienna RNA package [15] as though the three components are to fold into a heteroduplex like a pri-miR. Vienna predicts the structure and folding energy of each hypothetical heteroduplex. In order for Vienna to make these predictions, rna22 users specify the following:

1. The minimum required number of base-pairs between the miR and the target sequence (M).

2. The maximum allowed number of unpaired bases between the miR's seed sequence and the target sequence (G).

3. The maximum allowed free energy (kcal/mol) of the miR/target interaction in the heteroduplex (E). This value serves as a cutoff by which rna22 further defines likely miR/MRE binding events.

In this way, rna22 identifies likely miR binding targets by looking for recurrent sequence patterns in known miRs, determining the statistical significance of these patterns, and calculating the folding energy associated with the formation of a heteroduplex between individual miRs and target islands of potential miR binding sites.

ceRNA criteria:

Tay et al took a list of 10 miRs that were known from the literature to regulate PTEN and they generated a list of other, likely targets for these miRs using rna22. First they discard all predicted targets that have fewer than 7/10 miRs in common with PTEN. Next, they filter these likely targets by a few other criteria to bolster confidence that some of the targets compete with PTEN for miR binding. Gene X is predicted to be a ceRNA of PTEN if:

1.) **Genes X and PTEN have many miRs in common.** This criteria is expressed as (the number of PTEN miRs that are predicted to target gene X's 3'UTR) / (the total number of miRs that target the PTEN 3'UTR). The more miRs the two transcripts are predicted to share, the more this ratio approximates a value of 1. As the number of predicted, shared miRs between two transcripts increases, so does the likelihood that at

least some of the miRs actually regulate both transcripts. Furthermore, if two transcripts are to exhibit a cross-talk phenomenon, then sharing more miRs means that the two transcripts have more ways they can regulate one another.

2.) **Shared miRs are predicted to target gene X over a short region.** For a given miR, miR-n, gene X has a high ratio of (MREs for miR-n)/(distance in bases between left-most and right-most MREs of miR-n). This consideration arises because, longer 3'UTRs cause target prediction algorithms to return more spurious target predictions [36]. This is approximated by the formula: $E=K*L1*L2*\exp(-\lambda S)$, where:

- λ = A constant determined by the scoring system.
- K = A constant determined by the database used.
- L1 = The length of the 3'UTR.
- L2 = The length of the MRE.
- E = The mean of a poisson distribution of the number of matches between L1 and L2.

As the length of 3'UTR under consideration (L1) increases, the higher the mean of the distribution becomes. Thus, target predictions across the entire length of a 3'UTR can be noisy. This criterion prioritizes targets with shorter islands because they are less likely to contain spurious predictions.

3.) **There is an even distribution of miR targets within target islands in gene X.** This criteria is expressed as (distance between leftmost and rightmost MRE for miR-n)^2 / (sum of squared distances between miR-n MREs). This criterion does two things. First, it penalizes targets where most *but not all* MREs aggregate in a small neighborhood. In this way, longer target islands are penalized because they may introduce spurious predictions. Second, it rewards instances that maximize the minimum distance between the same MREs. The MREs that are targeted by miR-n must be spaced out enough that miR-n+RISC can sit on the mRNA target without steric interference by another miR-n+RISC complex. By maximizing the minimum distance between MREs, we increase confidence that miR binding to all these MREs is possible simultaneously.

4.) **Fewest possible miRs give rise to a large number of MREs in gene X.** This can be expressed as [(All MREs in transcript X for all miRs predicted to target X's 3'UTR) - (# of miRs predicted to target X's 3'UTR + 1)] / (MREs in X's 3'UTR for all considered miRs). This criteria rewards transcripts for having a higher number of predicted MREs for each targeting miR. As the number of predicted MREs in target X for a particular miR increases, so does the likelihood that the miR actually regulates the target in at least one place.

Once 1-4 above are calculated for all the miRs validated to regulate a gene of interest, the values can be multiplied out for every protein-coding transcript in the transcriptome to generate MuTaME scores. The transcriptome can then be rank-ordered based on the likelihood each gene functions as a ceRNA of your gene of interest. To review, the MuTaMe scoring approach takes validated regulatory miRs of a gene of interest and uses rna22 to generate lists of other genes likely regulated by those same miRs. This list is then rank-ordered by applying the criteria above which prioritize potential ceRNAs based on the number of miRs they share with a gene of interest, attributes of ceRNA target islands, and how likely miRs are to regulate ceRNA 3'UTRs.

Future directions:

By implementing rna22 and the selection criteria proposed, the MuTaME approach attempts to increase the sensitivity of shared miR regulation predictions. This increased sensitivity likely comes at the expense of specificity as many of the genes that score low may still have features that increase the likelihood of shared miR regulation. Here, we will explore a few considerations that could improve both the sensitivity and specificity of the approach using currently available information about miR targeting of mRNAs, mRNA secondary structure, and mRNA expression levels.

Currently, MuTaME scoring predictions are based on shared regulation of 3'UTRs. However, there are even more predicted MREs found in 5'UTRs and coding sequences [36]. While the 3'UTR has conventionally

been thought to be the primary location of miR regulation, there are increasingly found to be miR regulatory elements within coding sequences [9] and 5'UTRs [30]. By considering target islands within these regions of transcripts while still applying he sensitivity-enhancing criteria above, more transcripts will likely make MuTaME cutting scores and those with already high scores may receive yet higher scores by having high performing target islands (or, even average performing target islands) in multiple locations. Thus, extending the search for miR regulation to the entire transcript has the potential to increase the specificity of MuTaME score predictions. The sensitivity of these predictions may or may not be improved since considering multiple transcript regions will likely re-introduce some spurious predictions.

Next, we consider the secondary structure of candidate ceRNA target islands. It has been thought for some time that, because of regions of base-pairing complementarity, mRNAs display secondary structure [33,54]. Of relevance to ceRNA predictions is the observation that mRNAs with secondary structures permitting short stretches (4 nucleotides) of accessible nucleotides for miR binding are potentially better regulated by these miRs [29]. It has also been shown through experimentation that mRNA secondary structure influences miR regulation [3]. This nuanced interplay is increasingly under investigation [31,32]. There currently exist several RNA secondary structure prediction algorithms [8,60,61]. An investigator could potentially analyze the target islands computed by rna22 through any number of secondary structure prediction algorithms to estimate the local secondary structure of that island. The degree of secondary structure will influence likelihood that a miR:mRNA heteroduplex forms. This added consideration would likely discredit some predicted targets as ceRNAs, by virtue of inaccessible MREs. Accordingly, it should help increase the specificity of predictions.

Finally, the relative stoichiometry of potential ceRNA transcripts is likely a meaningful consideration in predicting miR titration events. For two candidate ceRNAs of gene X that have the same number of predicted, accessible MREs, the one whose expression is *most likely* to affect the expression of gene X is the one that is more highly expressed. This is because the transcript count serves as a coefficient for all of the miR-binding events that occur for a given target gene. As a result, MuTaME score predictions could include a coefficient as expressed by (total transcripts of candidate ceRNA – total transcripts of gene X) / (total transcripts of gene X). This way MuTaME scores can be weighted by the expression of candidate ceRNAs relative to a given gene of interest. Notably this consideration adds an additional experimental step to the process (either PCR or RNA-seq based quantification). However, it should enrich a list of candidate ceRNAs for those whose expression is most likely to be affected by the gene of interest, thus making downstream validation studies easier.

Conclusion:

Because mRNAs can be regulated by the same miRs, it is likely that a vast biological ceRNA network exists. This possibility adds a new wrinkle to the central dogma of biology where mRNAs are considered to be mere carriers of protein-coding information. Through this lens, transcripts are viewed to compete with one another for post-transcriptional regulation and ultimately endure mutual preservation or destruction. The first foray into predicting this competition by Tay et al. was successful and it yielded proof-of-concept that such a competition exists. However, more work remains to be done in refining ceRNA predictions. Here we have reviewed the experimental backing and computational workings that lead to ceRNA predictions and why those steps have already proven to be successful. In addition, we have proposed a few novel, readily implementable considerations that stand to improve the sensitivity and specificity of these predictions.

Lance Middleton
BIOC218 Final
06/02/2013

Works Cited:

1. Altschul S.F. et al. Basic local alignment search tool. J. Mol. Biol. 1990;215(3):403-410.

2. Cambronne X.A. et al. Capturing microRNA targets using an RNA-induced silencing complex (RISC)-trap approach. Proc. Natl. Acad. Sci. USA. 2012;109(50):20473-20478.

3. Cao S. & Chen S.J. Predicting kissing interactions in microRNA-target complex and assessment of microRNA activity. Nucl. Acids. Res. 2012;40(10):4681-4690.

4. Chen C. et al. Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res. 2005;33(20):e179.

5. Croce C.M. Causes and consequences of microRNA dysfunction in cancer. Nat. Rev. Genet. 2009;10(10)704-714.

6. Davison T.S. et al. Analyzing micro-RNA expression using microarrays. Methods Enzymol. 2006;411:14–34.

7. Doench J.G. et al. siRNAs can function as miRNAs. Genes & Dev. 2003;17:438-442.

8. Ding Y. & Lawrence C.E. A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Res.* 2003;31(24):7280–301.

9. Forman J.J. et al. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. Proc. Natl. Acad. Sci. USA. 2008;105(39):14879-14884.

10. Gaken J. et al. A functional assay for microRNA target identification and validation. Nucl. Acids Res. 2012;40(10):e75.

11. Garzon R. et al. Distinctive microRNA signature of acute myeloid leukemia bearing cytoplasmic mutated nucleophosmin. Proc. Natl. Acad. Sci. 2008;105:3945–3950.

12. Goettsch C. et al. MicroRNA in Cardiovascular Calcification: Focus on Targets and Extracellular Vesicle Delivery Mechanisms. Circ. Res. 2013;112;1073-1084.

13. Hammell M. mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. Nat. Methods. 2008;5(9):813-819.

14. Hammond S.M. et al. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature. 2000;404:293–295.

15. Hofacker I.L. et al. Fast folding and comparison of RNA secondary structures. Monatshefte f Chemie. 1994;125:167–188.

16. Humphreys D.T. et al. MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc. Natl. Acad. Sci. USA. 2005;*102:16961–16966.

17. Hutvágner G. et al. A cellular function for the RNA-interference enzyme dicer in the maturation of the let-7 small temporal RNA. Science. 2001;293:834–838.

18. Jiang Q. et al. miR2Disease: A manually curated database for microRNA deregulation in human disease. Nucleic Acids Res. 2009;37:D98–D104.

19. Kantharidis P. et al. Diabetes Complications: The MicroRNA Perspective. Diabetes. 2011;60(7):1832-1837.

20. Krichevsky A.M. et al. A microRNA array reveals extensive regulation of microRNAs during brain development. RNA. 2003;9:1274–1281.

21. Laurent L.C. et al. Comprehensive microRNA profiling reveals a unique human embryonic stem cell signature dominated by a single seed sequence. Stem cells 2008;26:1506–1516.

22. Lee Y. et al. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004;23:4051–60.

23. Lee Y. et al. The nuclear RNase III drosha initiates microRNA processing. Nature. 2003;425:415–419.

24. Lee Y. et al. MicroRNA maturation: stepwise processing and subcellular localization. EMBO J. 2002;21:4663–4670.

25. Lewis B.P. et al. Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. Cell. 2005;120(1):15-20.

26. Lim L.P. et al. The microRNAs of Caenorhabditis elegans. Genes Dev. 2003;17(8):991-1008.

27. Lim L.P. et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005;433:769-773.

28. Livak K.J. and Schmittgen T.D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta}C_T$ Method. Methods. 2001;25(4)402-408.

29. Long D. et al. Potent effect of target structure on microRNA function. Nat. Struct. Mol. Biol. 2007;14:287-294.

30. Lytle J.R. et al. Target mRNAs are repressed as efficiently by microRNA binding sites in the 5'UTR as in the 3'UTR. Proc. Natl. Acad. Sci. USA. 2007;104(23):9667-9672.

31. Marin R.M. & Vanicek J. Efficient use of accessibility in microRNA target prediction. Nucl. Acids Res. 2011;39(1)19-29.

32. Marin R.M. & Vanicek J. Optimal Use of Conservation and Accessibility Filters in MicroRNA Target Prediction. PLoS ONE. 2012;7(2):e32208.

33. Kramer F. & Mills D. Secondary structure formation during RNA-synthesis. Nucleic Acids Res. 1981;9:5109-5124.

34. Mestdagh P. et al. A novel and universal method for microRNA RT-qPCR data normalization. Genome Biol. 2009;**10**:R64.

35. Metzker M.L. et al. Emerging technologies in DNA sequencing. *Genome Res.* 2005;15:1767–1776.

36. Miranda K.C. et al. A Pattern-Based Method for the Identification of MicroRNA Binding Sites and Their Corresponding Heteroduplexes. Cell. 2006;126(6):1203-1217.

37. Miska E.A. et al. Microarray analysis of microRNA expression in the developing mammalian brain. Genome Biol. 2004;5:R68.

38. Morin R.D. et al. Application of massively parallel sequencing to microRNA profiling and discovery of human embryonic stem cells. Genome Res. 2008;18:610-621.

39. Peltier H.J. & Latham G. J. Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. RNA. 2008;14:844–852.

40. Petersen C.P. et al. Short RNAs repress translation after initiation in mammalian cells. Mol. Cell. 2006;21:533–542.

41. Pillai R.S. et al. Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science. 2005;309:1573–1576.

42. Pradervand S. et al. Impact of normalization on miRNA microarray expression profiling. RNA. 2009;15:493-501.

43. Pritchard C.C. et al. MicroRNA profiling: approaches and considerations. Nat. Revs. Genet. 2012;13:358-369.

44. Rigoutsos I. & Floratos A. Combinatorial pattern discovery in biological sequences: The TEIRESIAS algorithm. Bioinformatics. 1998;14(1):55-67.

45. Ritchie W. Predicting microRNA targets and functions: traps for the unwary. Nat. Methods. 2009;6:397-398.

46. Ronaghi M. et al. A sequencing method based on real-time pyrophosphate. Science. 1998;28:363–365.

47. Salmena L. et al. A ceRNA Hypothesis: The Rosetta Stone of a Hidden RNA Language? Cell. 2011;146(3):353-358.

48. Shi R. and Chiang V.L. Facile means for quantifying microRNA expression by real-time PCR. BioTechniques. 2005;39:519-525.

49. Stark A. et al. Identification of *Drosophila* MicroRNA Targets. PLoS Biol. 2003;1(3): e60.

50. Shingara J. et al. An optimized isolation and labeling platform for accurate microRNA expression profiling. RNA. 2005;11:1461–1470.

51. Sun Y. et al. Development of a microarray to detect human and mouse microRNAs and characterization of expression in human organs. Nucleic Acids Res. 2004;32(22):e188.

52. Tay Y. et al. Coding-Independent Regulation of the Tumor Suppressor PTEN by Competing Endogenous mRNAs. Cell. 2011;147(2):344-357.

53. Thomson J.M. et al. A custom microarray platform for analysis of microRNA gene expression. Nat. Methods. 2004;1:47-53.

54. Turner D.H. et al. RNA Structure Prediction. Annu. Rev. Biophys. Biophys. Chem. 1988;17:167-192.

55. Vinther J. et al. Identification of miRNA targets with stable isotope labeling by amino acids in cell culture. Nucl. Acids. Res. 2006;34(16):e107.

56. Wang H. et al. Direct and sensitive miRNA profiling from low-input total RNA. RNA. 2007;13:151–159.

57. Welberg L. Neurodegenerative disorders: Reconnect with microRNA. Nat. Revs. Neurosci. 2010;11:74-75.

58. Wu L. et al. MicroRNAs direct rapid deadenylation of mRNA. Proc. Natl. Acad. Sci. USA. 2006;103:4034–4039.

59. Yi R. et al. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes & Dev. 2003;17: 3011–3016.

60. Zakov S. et al. Rich parameterization improves RNA structure prediction. J. Comput. Biol. 2011;18(11): 1525–1542.

61. Zuker M. & Stiegler P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* 1981;9(1):133–48.