BIOINFORMATICS DISCOVERY NOTE Vol. 28 no. 4 2012, pages 453–456
doi:10.1093/bioinformatics/btr693

Genome analysis Advance Access publication December 16, 2011

The role of miRNAs in complex formation and control

Wilson Wen Bin Goh[1], Hirotaka Oikawa[2], Judy Chia Ghee Sng[2,3], Marek Sergot[1] and Limsoon Wong[4,*]

[1]Department of Computing, Imperial College London, London, UK, [2]Singapore Institute for Clinical Sciences, A*STAR, [3]Department of Physiology and [4]Department of Computer Science, National University of Singapore, Singapore

Associate Editor: Ivo Hofacker

ABSTRACT

Summary: microRibonucleic acid (miRNAs) are small regulatory molecules that act by mRNA degradation or via translational repression. Although many miRNAs are ubiquitously expressed, a small subset have differential expression patterns that may give rise to tissue-specific complexes.

Motivation: This work studies gene targeting patterns amongst miRNAs with differential expression profiles, and links this to control and regulation of protein complexes.

Results: We find that, when a pair of miRNAs are not expressed in the same tissues, there is a higher tendency for them to target the direct partners of the same hub proteins. At the same time, they also avoid targeting the same set of hub-spokes. Moreover, the complexes corresponding to these hub-spokes tend to be specific and nonoverlapping. This suggests that the effect of miRNAs on the formation of complexes is specific.

Contact: wongls@comp.nus.edu.sg

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on July 7, 2011; revised on November 28, 2011; accepted on December 13, 2011

1 INTRODUCTION

microRibonucleic acid (miRNAs) are small non-coding regulatory molecules that act by downregulating gene expression at the post-transcriptional and pre-translational stage. They have an extensive array of functions and are implicated in diverse processes such as growth, development and differentiation. They also play important roles in diseases such as cancer (Meltzer, 2005) and heart disease (van Rooij and Olson, 2007).

miRNAs are known to play roles in regulating the protein interaction network. (Liang and Li, 2007) showed that miRNAs tend to target hubs in a network. Moreover, this regulation is more important for intermodular hubs than intramodular ones. (Hsu *et al.*, 2008) further built upon this and discovered that while proteins directly regulated by miRNAs might not form a network module themselves, the miRNA-target genes and their interacting neighbors jointly showed significantly higher modularity. This is an early indication that miRNAs target select components of complexes. However, it is not until recently when Sass *et al.* (2011) confirmed that individual miRNAs or co-expressed miRNAs frequently target several components of protein complexes, indicating a coordinate post-transcriptional regulation of protein complexes by miRNAs.

In the human protein interaction network, several large hubs are often observed (Barabasi and Albert, 1999), usually corresponding to critical gene or biological process regulators (Jeong *et al.*, 2001). The partners (hub-spokes) are functional units controlled by the hub, and can be considered as part of a functional unit such as a complex. Given that different tissues utilize their gene repertoire differently, it is inconceivable that interactions between hubs and all adjoining hub-spokes exist simultaneously since there are limitations in physical binding area, and so on. One possible mode of control is via expressional suppression, thereby giving rise to different or subforms of complexes. Since Sass *et al.* (2011) showed that coexpressing miRNAs have a propensity to target complexes, we extend this by suggesting that anti-coexpressing miRNAs can effect the formation of different protein complexes by suppressing gene expression. That is, miRNAs with widely different expression profiles can control mutually exclusive or tissue-specific biological processes by impacting on the formation of different complexes.

2 METHODS

2.1 Data sources

2.1.1 DIANA DIANA-microT (ver 3.0) calculates parameters for each miRNA based on binding and conservation levels (Maragkakis *et al.*, 2009). The total predicted score of a miRNA-target pair is the weighted sum of conserved and unconserved miRNA recognition element (MRE) of a gene. A signal-to-noise ratio (SNR) and a precision score specific for each interaction that can be used as a confidence estimate of the correctness of each predicted miRNA-target relationship is provided in the dataset. The Human dataset is downloaded from http://diana.cslab.ece.ntua.gr.

2.1.2 Mammalian microRNA expression atlas Human miRNA expression is based on the dataset of Landgraf *et al.* (2007). This dataset is extensive and includes expression information for 516 miRNAs across 79 normal tissues, cancer tissues and cell lines. Forty normal tissues are extracted for evaluation.

2.1.3 Human Protein Reference Database Human Protein Reference Database (HPRD) release 9 at http://www.hprd.org is used as the reference network (Goel *et al.*, 2011). It is cleaned using iterative AdjustCD (Liu *et al.*, 2009). No new edges are added. Two iterations are used. Threshold is set at a score of 0.3 where the functional homogeneity is 0.6 (4800 edges). Hubs (Supplementary Table S1) are defined as nodes having 50 or more connections (based on the long tail of the degree distribution curve) and are further filtered using Gene Ontology for involvement in biological regulation process (GO:0050789).

*To whom correspondence should be addressed.

2.1.4 CORUM CORUM September 2009 release at http://mips.helmholtz-muenchen.de provides a resource of manually annotated protein complexes from mammalian organisms (Ruepp *et al.*, 2010). A total of 1200 human complexes are used.

2.2 Analytical pipeline

2.2.1 Coexpression calculation using Hamming distance, D A miRNA expression profile is regarded as a binary vector across a series of 40 normal tissues (1 for expressed, 0 for not expressed). To relate the degree of similarity of two miRNAs in terms of expression profile, the Hamming distance (D) and its P-value can be calculated based on their intersection as described by Wu *et al.* (2003); see Supplementary Materials. To find pairs that are strongly anti-coexpressed, we set a P-value threshold ≤ 0.05 and with Hamming distance ≥ 35 (maximum is 40).

2.2.2 Spoke and complex intersection analytical pipeline miRNA pairs identified as anti-coexpressed are then checked for their hub-spoke targeting overlap using precision scores from DIANA. Targeted spokes are then checked for correspondence with known complexes from CORUM. The overlap between targeted spokes and their corresponding complexes is calculated and the average overlap reported.

3 RESULTS AND DISCUSSION

3.1 Anti-coexpressed miRNA pairs are few

Based on the expression of 516 miRNAs across 40 normal human tissues, most miRNAs are coexpressing (Fig. 1 C). This indicates that many miRNAs tend to be found in the same tissues. Anti-coexpressed miRNAs are rarer. In order to obtain a reasonably sized set of anti-coexpressed miRNAs, we define an anti-coexpressed pair as any pair with Hamming distance ≥ 35 (maximum 40) and which meets a P-value threshold of ≤ 0.05.

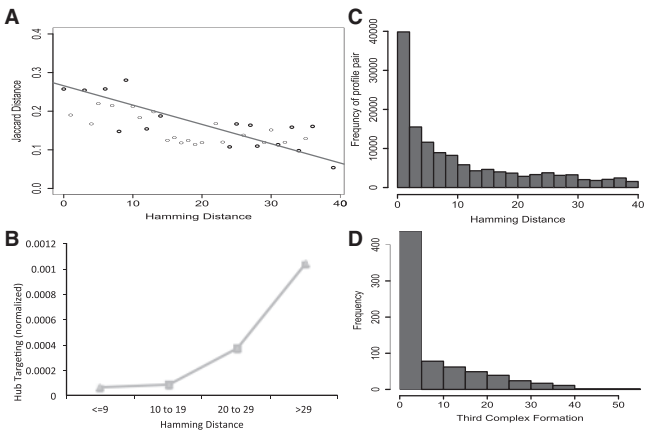

Fig. 1. **(A)** Hamming distance against Jaccard distance shows that on average, anti-coexpressed miRNAs tend to avoid targeting the same spokes. Adjusted $R^2 = 0.3$, $P = 0.00024$. **(B)** miRNA profile pairs (Hamming distance) are plotted against the normalized average number of hubs (hubs affected/total number of unique miRNA–target relationships) in HPRD whose spokes are targeted. At higher Hamming distances, the tendency to control hub-spokes increases dramatically. **(C)** The frequency distribution of Hamming distance based on all miRNA pairs shows that most miRNA pairs are coexpressed. **(D)** In considering all hub-spokes disrupted by anti-coexpressing complex pairs, it is found that disrupted complexes tend not to possess common components that can be used to make an alternative complex (Third Complex).

3.2 Anti-coexpressed miRNAs tend to target hub-spokes very strongly

From the cleaned network, we retrieve the first-degree neighbors of each hub (spokes). For each pair of miRNAs with Hamming distance D, we check the set of miRNA-targeted hub-spokes across all hubs. We define the probability each miRNA targets a hub-spoke based on the precision score provided in DIANA. We perform simulation 1000 times and take an average of the overlap of targeted hub-spokes between the miRNA pair.

We divide D into four bins: ≤ 9, 10–19, 20–29 and >29. For each bin, to quantify hub-spoke targeting propensity, the number of affected hubs is divided over the total number of unique miRNA-target pairs. This provides confidence that the hub-spoke targeting is not spurious as a result of many predicted targets. This is also important because we realize from Figure 1C that there are many co-expressed miRNAs, and the likelihood of targeting a hub-spoke is therefore quite high. However, in Figure 1 B, it is observed that anti-coexpressing miRNAs tend to target hub-spokes with significantly higher propensity.

3.3 Anti-coexpressed miRNAs avoid targeting the same set of hub-spokes

Although anti-coexpressing miRNAs tend to target hub-spokes, they do not target the same spokes or complexes. We measure overlap by Jaccard Distance (J score) for a given miRNA pair with a Hamming distance D. J score is calculated as the intersection of targeted hub-spokes between a given miRNA pair divided by their respective union. Here, J decreases linearly with increasing D. Aside from a single outlier, a non-binned plot shows a well-defined linear anti-correlation ($P = 0.00024$) (Fig. 1A). This suggests that anti-coexpressed miRNAs also avoid each other's targets. This antagonism may give rise to differential functionality, e.g. different complexes being formed. Moreover, this control appears quite specific: only a small subset of CORUM complexes (54/1200) are controlled by these anti-coexpressed pairs.

3.4 Reproducibility in other prediction algorithms

To show that the above observations are reproducible, we perform similar analysis using PicTar (Krek *et al.*, 2005), PITA (Kertesz *et al.*, 2007) and TargetScan 5.1 (TS) (Friedman *et al.*, 2009). The propensity to target hub-spokes by anti-coexpressed miRNAs in TS, PicTar and PITA is similar (Supplementary Fig. S5). All show that anti-coexpressed miRNAs have a significantly higher propensity to target the spokes of hubs.

In all datasets, D and J are consistently inversely related (Supplementary Fig. 6). This strongly supports the notion that anti-coexpressed miRNAs avoid each other's targets and may be important in the formation of different protein complexes or acquiring tissue-specific functionalities.

3.5 Case study on miR-29b Triad

To better understand the functional control between anti-coexpressed miRNAs, we construct an anti-coexpression network (Supplementary Fig. S3); where the nodes are miRNAs and edges are anti-coexpression relationships. The anti-coexpression triad consisting of miR-29b and miR-140-5p, both anti-coexpressed against miR-499-5p (Supplementary Fig. S4), is selected for further

study. miRs-140-5p and 29b are both ubiquitously expressed (38/40 tissues), whereas miR-499-5p is found only in heart. miR-29b is found in all tissues except heart and uterus, whereas miR-140-5p is found in all except heart and B-cell. Due to these expression patterns, we hypothesize that the removal of targets of miRs-29b and 140-5p in heart tissue should affect the acquisition of functions specific to heart.

Using human tissue-specific networks generated by (Emig and Albrecht, 2011), we compare heart and liver networks and check the affected nodes and edges if miRs-29b and 140-5-p are allowed to be expressed in heart using the target predictions in DIANA. We then test these affected nodes for GO term function enrichment. If miRs-29b and 140-5p were allowed to be expressed in heart, their high confidence targets (precision ≥ 0.9) would affect blood related, extra-cellular matrix, chemical stimulation (miR-29b) and nervous system development (miR-140-5p) functions in addition to other more generic functions (Supplementary Table S2). For pathway enrichment, in miR-29b, we find 'myometrial relaxation and contraction pathway' and 'calcium regulation in cardiac cell' ($P \leq 0.01$). These suggest that removal of miRs-29b and miR-140-5p target genes affects heart-related function and indicate that tissue-specific miRNAs do regulate tissue-specific functions.

3.6 How miRNAs may control BAF complex in VPA-treated mice

To further show that miRNAs disrupt complexes specifically, we examine the disrupted complexes where there is experimental evidence for both miRNA upregulation and mRNA downregulation. However, this dataset comes from mouse.

Valproic Acid (VPA) is used medically as an anticonvulsant and a mood stabilizer, but has effects on learning and memory (Bredy and Barad, 2008). It functions primarily as an epigenetic regulator via its histone deacetylase inhibitor activity. The treatment of mice over a 2-day period with VPA also indicates readjustment of miRNA levels in the brain. There are 136 overexpressed miRNAs and 1000 downregulated genes; for details, see Supplementary Material.

The effects of VPA treatment on the brain is well studied. Its role in epigenetic remodeling has a tangible impact on neuronal differentiation (Kondo, 2006). VPA prompts the differentiation of hippocampal neural progenitor cells into neurons, but also prevents differentiation into oligodendrocytes and astrocytes (Hsieh *et al.*, 2004). In a screen of several psychoactive drugs in rats, it has been shown that only VPA exerted clear definable increase in acetylation (Perisic *et al.*, 2010).

VPA treatment shows elevation of 136 miRNAs (t-test $P \leq 0.01$). We investigate how VPA mediates its effects via miRNA activity. To determine which downregulated genes are miRNA targets, we combine scores from TS (Friedman *et al.*, 2009), PT (Krek *et al.*, 2005) and DIANA (Maragkakis *et al.*, 2009). Those downregulated genes with a combined score ≥ 0.9 are deemed high confidence (236 genes). GO term analysis using GO Term Finder reveals that a good number of these genes are involved in chromatin modification ($n = 14$, $P \leq$ 2.12e-05), nervous system development ($n = 23$, $P \leq$ 0.0008) and cell differentiation ($n = 32$, $P \leq 0.00129$). This implies that the miRNA targets play a role in epigenetic regulation and neuronal maturation.

To investigate further, the 236 high confidence miRNA targets are compared against CORUM (Ruepp *et al.*, 2010). A total of 53 unique complexes are disrupted. Of these, 15 contain HDAC 1 and/or 2. Eight possess Swi/Snf components involved in neuronal differentiation and neurogenesis. Five belong to the Polycomb family of epigenetic regulating complexes, which play a role in cell fate transition and neuronal differentiation (Bracken *et al.*, 2006). And eight complexes are SMAD regulators, which act as transcription factors. In particular, SMAD can induce proliferation and differentiation of hippocampal neurons via brain-derived neurotrophic factor (BDNF) and transforming growth factor beta (Lu *et al.*, 2005).

Several disrupted complexes are involved in neuronal function. For instance, SNARE, BHC and HPRC1L. Another, the BRG1/brm-associated factor (BAF) family of complexes, act as a switch in neuronal differentiation, and exists in two forms; neural progenitor (npBAF) and matured (nBAF). As neural progenitors exit mitosis and differentiate into neurons, npBAF complexes, which contain ACTL6a and PHF10, are exchanged for ACTL6b and DPF1/DPF3 subunits in nBAF. The npBAF complex is essential for self-renewal/proliferative capacity of multipotent neural stem cells. The nBAF complex along with CREST plays a role regulating the activity of genes essential for dendrite growth.

Yoo *et al.* (2009) reported that the transition from npBAF to nBAF is mediated by repression of ACTL6a by miR-9* and miR-124. Our miRNA chip analysis in the 2-day period also indicates that miR-9* and miR-124 are strongly upregulated due to VPA, with corresponding decrease in ACTL6a levels. This agrees with the article. However, it appears that VPA response is not so simple. First, there are a multitude of 137 significantly upregulated miRNAs of which miR-9* and miR-124 are only two. Second, the downregulation of ACTL6a is not very significant.

Therefore, we check the expression level of PHF10, another component of npBAF. We find that its expression level increases significantly (t-test, $P = 0.01$). However, PHF10 is not a strong miRNA target and is predicted as a target of upregulated miR-96 by only DIANA (precision = 0.46).

To investigate further, we check the nBAF-specific components expression as well. We find that both ACTL6b and DPF1 levels, specific components of npBAF, are decreased. In particular, DPF1 decrease is significant for both probes (t-test, P = 0.01 and 0.02). DPF1 is also a predicted target for a multitude of miRNAs and supported by TargetScan, DIANA and PicTar with a combined precision score of 0.81 (for only upregulated miRNAs). This suggests that DPF1 is likely to be regulated by the upregulated miRNAs. The expression levels and targeting miRNAs for DPF1 and PHF10 are found in Supplementary Table S3.

The rise in precursor BAF components may correspond to VPA's neurological effects. Additionally, it appears that this effect is mediated by miRNA activity. However, the role of miRNAs in regulating the BAF switch appears to be complex. First, both ACTL6a and ACTL6b, components of npBAF and nBAF, respectively, are downregulated. However, neither of these show significant downregulation relative to controls. Furthermore, it was reported by (Yoo *et al.*, 2009) that ACTL6a expression tends to repress ACTL6b. In the case of ACTL6a, it appears that the upregulation of known targeters miR-9* and miR-124 as a result of VPA treatment do not have a strong effect in its suppression, suggesting activation of other compensatory processes due to VPA.

On the other hand, PHF10 appears to evade miRNA regulation whereas DPF1 appears to be targeted by a good number of upregulated miRNAs. This may result in more npBAF complexes being formed, with the possible effect of reactivation of neural plasticity and learning in adult mice.

4 CONCLUSION

We have shown how miRNA expression patterns could impact the formation of complexes. First, anti-coexpressed miRNAs tend to regulate hub-spokes significantly. Second, anti-coexpressed miRNAs affect different spokes corresponding to different complexes. The activity of anti-coexpressed miRNAs is also specific in that targeted complexes tend not to have components common to other complexes, thereby limiting collateral or non-specific effects.

miRs-29b and 140-5p are ubiquitously expressed but not in heart tissue and are anti-coexpressed to miR-499-5p. Furthermore, the targets of miRs-29b and 140-5p are enriched for heart-related functions. This reinforces the notion that anti-coexpressed miRNAs are involved in the acquisition of tissue-specific functions.

In VPA-treated mice, it appears that miRNAs play a role in the complex formation. The increase in precursor BAF complexes due to miRNAs targeting specific components of the mature form may have a significant effect in the psychoactive effects of VPA, e.g. by allowing reversion to a higher level of brain plasticity.

Funding: Wellcome Trust Scholarship (83701/Z/07/Z for W.W.B.G.); Singapore National Research Foundation grant (NRF-G-CRP-2007-04-082(d) for L.W.) in part.

Conflict of Interest: none declared.

REFERENCES

Barabasi,A.L. and Albert,R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.

Bracken,A.P. *et al.* (2006) Genome-wide mapping of polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.*, **20**, 1123–1136.

Bredy,T.W. and Barad,M. (2008) The histone deacetylase inhibitor valproic acid enhances acquisition, extinction, and reconsolidation of conditioned fear. *Learn. Mem.*, **15**, 39–45.

Emig,D. and Albrecht,M. (2011) Tissue-specific proteins and functional implications. *J. Proteome Res.*, **10**, 1893–1903.

Friedman,R.C. *et al.* (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, **19**, 92–105.

Goel,R. *et al.* (2011) Human protein reference database and human proteinpedia as discovery resources for molecular biotechnology. *Mol. Biotechnol.*, **48**, 87–95.

Hsieh,J. *et al.* (2004) Histone deacetylase inhibition-mediated neuronal differentiation of multipotent adult neural progenitor cells. *Proc. Natl Acad. Sci. USA*, **101**, 16659–16664.

Hsu,C.-W. *et al.* (2008) Characterization of microRNA-regulated protein-protein interaction network. *Proteomics*, **8**, 1975–1979.

Jeong,H. *et al.* (2001) Lethality and centrality in protein networks. *Nature*, **411**, 41–42.

Kertesz,M. *et al.* (2007) The role of site accessibility in microRNA target recognition. *Nat. Genet.*, **39**, 1278–1284.

Kondo,T. (2006) Epigenetic alchemy for cell fate conversion. *Curr. Opin. Genet. Dev.*, **16**, 502–507.

Krek,A. *et al.* (2005) Combinatorial microRNA target predictions. *Nat. Genet.*, **37**, 495–500.

Landgraf,P. *et al.* (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*, **129**, 1401–1414.

Liang,H. and Li,W.-H. (2007) MicroRNA regulation of human protein protein interaction network. *RNA*, **13**, 1402–1408.

Liu,G. *et al.* (2009) Complex discovery from weighted PPI networks. *Bioinformatics*, **25**, 1891–1897.

Lu,J. *et al.* (2005) Smad pathway mediation of BDNF and TGF beta 2 regulation of proliferation and differentiation of hippocampal granule neurons. *Development*, **132**, 3231–3242.

Maragkakis,M. *et al.* (2009) Diana-microt web server: elucidating microRNA functions through target prediction. *Nucleic Acids Res.*, **37**, W273–W276.

Meltzer,P.S. (2005) Cancer genomics: small RNAs with big impacts. *Nature*, **435**, 745–746.

Perisic,T. *et al.* (2010) Valproate and amitriptyline exert common and divergent influences on global and gene promoter-specific chromatin modifications in rat primary astrocytes. *Neuropsychopharmacology*, **35**, 792–805.

Ruepp,A. *et al.* (2010) Corum: the comprehensive resource of mammalian protein complexes–2009. *Nucleic Acids Res.*, **38**, D497–D501.

Sass,S. *et al.* (2011) MicroRNAs coordinately regulate protein complexes. *BMC Syst. Biol.*, **5**, 136.

van Rooij,E. and Olson,E.N. (2007) MicroRNAs put their signatures on the heart. *Physiol. Genomics*, **31**, 365–366.

Wu,J. *et al.* (2003) Identification of functional links between genes using phylogenetic profiles. *Bioinformatics*, **19**, 1524–1530.

Yoo,A.S. *et al.* (2009) MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature*, **460**, 642–646.