

RESEARCH

Open Access

Understanding cooperativity of microRNAs via microRNA association networks

Young-Ji Na^{1*}, Ju Han Kim^{1,2}

From Asia Pacific Bioinformatics Network (APBioNet) Twelfth International Conference on Bioinformatics (InCoB2013)

Taicang, China. 20-22 September 2013

Abstract

Background: MicroRNAs (miRNAs) are key components in post-transcriptional gene regulation in multicellular organisms. As they control cooperatively a large number of their target genes, they affect the complexity of gene regulation. One of the challenges to understand miRNA-mediated regulation is to identify co-regulating miRNAs that simultaneously regulate their target genes in a network perspective.

Results: We created miRNA association network by using miRNAs sharing target genes based on sequence complementarity and co-expression patterns of miRNA-target pairs. The degree of association between miRNAs can be assessed by the level of concordance between targets of miRNAs. Cooperatively regulating miRNAs have been identified by network topology-based approach. Cooperativity of miRNAs is evaluated by their shared transcription factors and functional coherence of target genes. Pathway enrichment analysis of target genes in the cooperatively regulating miRNAs revealed the mutually exclusive functional landscape of miRNA cooperativity. In addition, we found that one miRNA in the miRNA association network could be involved in many cooperatively regulating miRNAs in a condition-specific and combinatorial manner. Sequence and structural similarity analysis within miRNA association network showed that pre-miRNA secondary structure may be involved in the expression of mature miRNA's function.

Conclusions: On the system level, we identified cooperatively regulating miRNAs in the miRNA association network. We showed that the secondary structures of pre-miRNAs in cooperatively regulating miRNAs are highly similar. This study demonstrates the potential importance of the secondary structures of pre-miRNAs in both cooperativity and specificity of target genes.

Background

MicroRNAs (miRNAs) play crucial regulatory roles in repressing mRNA translation or mediating mRNA degradation by targeting mRNAs in a sequence-specific manner [1]. MiRNA-mediated regulation has been found to encompass a wide spectrum of biological processes ranging from growth and development to oncogenesis [2-5]. In general, one miRNA can target more than one gene (*i.e.*, multiplicity), and one gene can be controlled by more than one miRNA (*i.e.*, cooperativity) [6]. Cooperative

binding of one or several distinct miRNAs to a single target gene has been shown to be important for the functionality of miRNA-mediated gene regulation [6,7]. For example, genetic evidences in previous research suggest that the *lin-28* gene is cooperatively regulated by the *lin-4* miRNA and another unidentified miRNA [8]. Krek *et al.* [9] also presented that miR-375, miR-124 and let-7b jointly regulate *Mtpn*, providing evidence to support coordinate control of miRNAs in mammals. Studying the cooperativity of miRNAs is a substantial step for understanding the contribution of co-regulating miRNAs towards a complex interplay between miRNAs.

Recently several studies have attempted to develop methods to understand miRNA cooperativity. Boross *et al.*

* Correspondence: yjna01@snu.ac.kr

¹Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, Seoul 110-799, Korea
Full list of author information is available at the end of the article

computed the correlations between the gene silencing scores of individual miRNAs [10]. Xu *et al.* also developed a computational method to identify significant synergistic miRNA pairs via functional co-regulating miRNAs that they jointly regulate [11]. Most of these studies however did not actually considered co-expression profiles of mRNAs and miRNAs. Considering that most miRNAs exert their functions through interactions with other miRNAs, an understanding of a miRNA network context using both co-expression pattern and the sequence complementarity between miRNAs and mRNAs is essential to discover the cooperative regulation of miRNAs.

In this paper, we propose a computational method to construct a miRNA association network (MRAN) by integrating multiple genomic data sources including miRNA-mRNA binding information and miRNA-mRNA co-expression profiles. While sequence complementarity-based miRNA-mRNA target relationships serve as a static set, dynamic expression profiles of miRNAs and mRNAs are used to identify those subsets that are concurrently active. We evaluate the miRNAs modules determined to be significantly cooperative by assessing the functional coherence of target genes co-regulated by the miRNA modules. Then, we apply graph theoretical methods to characterize MRAN and analyse whether miRNAs belonging to the network are associated with each other in a condition-dependent manner or not. Finally we report that co-regulating miRNAs tend to have structural similarities.

Results

Construction and characterization of miRNA association networks

Table 1 lists the four datasets used in the present study to create MRANs following the steps illustrated in Figure 1. Table 2 shows the distribution of miRNAs and mRNAs showing inverse expression pattern in each condition. Figure 2(A) shows the global MRAN created by superimposing four conditions specific MRANs (see Figure 1) consists of 241 miRNAs and 559 connections. In MRAN, a node corresponds to each miRNA that has the significant inverse expression pattern with its targets under each experimental condition (Pearson's correlation

coefficient $r < 0$), and edges represent target overlap score p value < 0.05 . MiRNAs were clustered into distinct groups.

Molecular COmplex DETection algorithm (MCODE) [12] revealed 12 Cooperative MiRNA Modules (CMM) (Table 3). The numbers of miRNAs and connections of the CMMs range from 2 to 19 and 2 to 96, respectively (Table 3). CMM 2 and 4 consist of let-7 family members and miR-103/107 family, respectively. Table 3 exhibits miRNA family members included in CMMs and Figure 3 shows their condition-specific cooperativity. The first known human miRNA *let-7* and its family members are highly conserved across species in sequence and function. Misregulation of *let-7* leads to a less differentiated cellular state and the development of cell-based diseases such as cancer. The role of miR-103 and miR-107 in regulation of CDK5R1 expression and in cellular migration and neural development is well documented [13].

Transcription factors are thought to regulate the transcription of miRNAs in a pol II dependent manner similar to that of protein-coding genes; that is, by binding to conventional transcription factor binding site sequences (TFBS) located in or near promoter regions that lie upstream of the miRNAs [14,15]. We examined the promoter regions of miRNAs for the presence of regulatory motifs. We determined over-represented motifs and found 285 significantly over-represented TFBS pairs (1 in prostate cancer, 241 in radiation, 6 in Actinomycin-D treatment and 37 in circadian rhythm) among the 6790 pairs (*i.e.*, 766 in prostate cancer, 5105 in radiation, 744 in Actinomycin-D treatment, and 175 in circadian rhythm) using hypergeometric distribution (Bonferroni corrected p -value $< 1.0e-05$) (Figure 2(B)).

Mutually exclusive functional landscape of miRNA cooperativity

The biology of miRNA function will be dictated by the mRNA transcripts targeted by specific miRNAs [16]. Functional enrichment analysis of miRNA's target transcripts hence can be used as a proxy for evaluation of the functional coherence of CMMs. We performed functional enrichment analysis by using all KEGG and

Table 1 Co-transcriptomic expression datasets

Experiment	Sample	mRNA expression profile	miRNA expression profile	Reference
Circadian rhythm	Mouse liver	Illumina Mouse-6 Expression BeadChip	Ambion mirVana™ miR Bioarray V2	Na et al., 2009 [49]
ActinomycinD treatment	Mouse brain neuroblast N2a cells	Affymetrix Mouse Gene 1.0 ST Array	Ambion mirVana™ miR Bioarray V2	Unpublished data
Prostate cancer	Human prostate adenocarcinomas	Affymetrix Human Gene 1.0 ST Array	OSU-CCC MicroRNA Microarray	Prueitt et al., 2008 [37]
Radiation treatment	Human lung cancer H460 and H1229 cells	Affymetrix Human Gene 1.0 ST Array	Ambion mirVana™ miR Bioarray V2	Lee et al., 2008 [36]

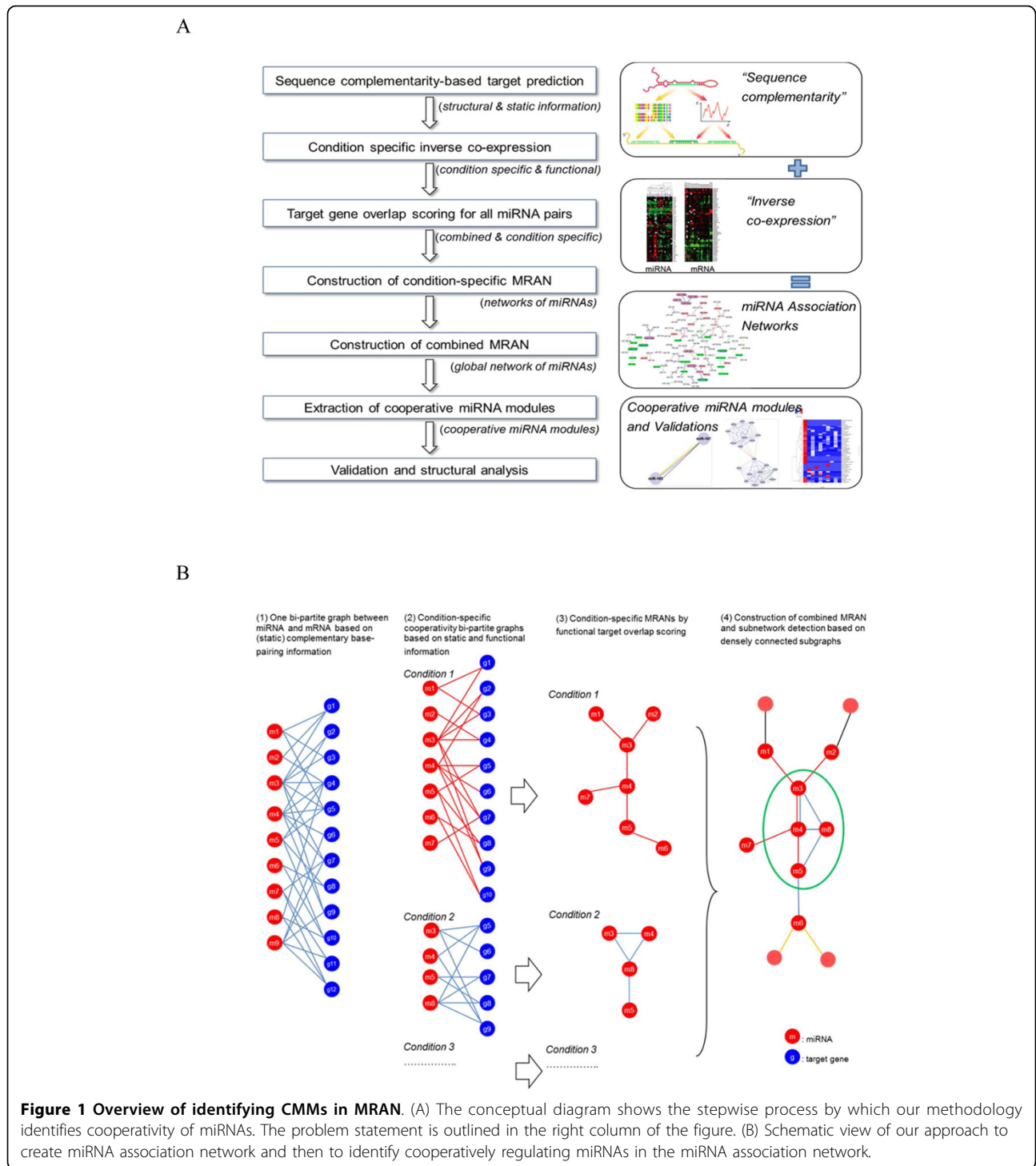
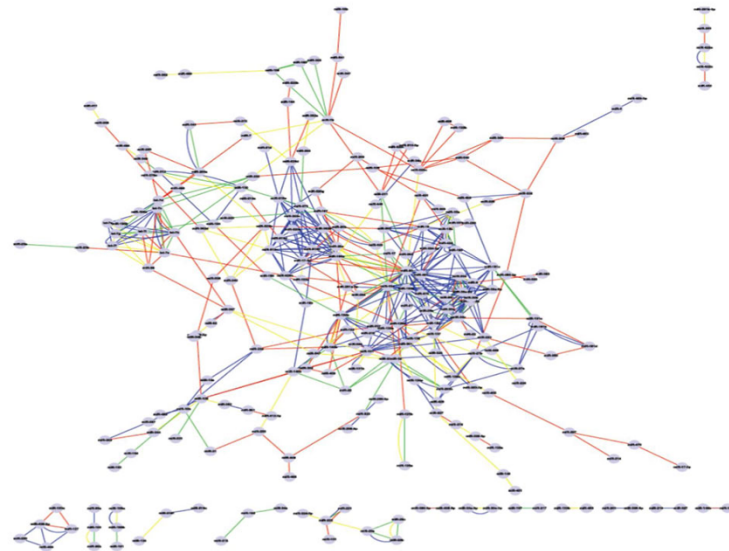


Table 2 Distribution of miRNAs and mRNAs with inverse expression patterns

Category	Circadian rhythm	ActinomycinD	Prostate cancer	Radiation
No. of miRNAs	266	275	157	302
No. of genes	638	2446	1268	4354
No. of miRNA-gene pairs	1104	5095	2673	20522

A



B

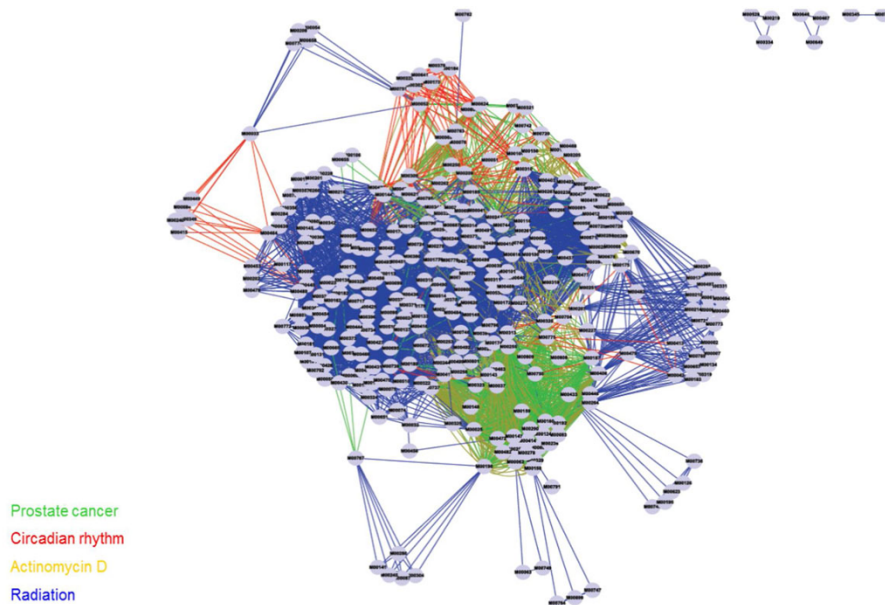


Figure 2 Global MRAN and co-occurrence graph of regulators on MRAN. (A) Global MRAN. Nodes of the network represent miRNAs, and edges represent condition-specific cooperativity between a miRNA pair. (B) Co-occurrence graph of regulators on MRAN. In the co-occurrence graph of regulators on MRAN, nodes represent transcription factors (TFs), and edges statistical significance of co-occurrences. The colors of edges indicate different conditions (red: circadian rhythm, blue: radiation treatment, yellow: Actinomycin D treatment, green: prostate adenocarcinomas).

BioCarta pathways for each CMM. Heat maps in Figure 4(A) and 4(B) exhibit $-\log_{10} p$ -values obtained by hypergeometric tests between (horizontal axis) CMMs' target transcripts and (vertical axis) KEGG and BioCarta pathways. All pathways mapped to at least one CMM and all modules

mapped to at least one pathway appear in the Heat maps. Those modules and pathways that have no mapping at all are omitted from each heat map. Ten (*i.e.*, CMMs 1~9 and 12) and nine (*i.e.*, CMMs 1~9) CMMs have at least one KEGG and BioCarta pathway mappings, respectively.

Table 3 Cooperative miRNA modules extracted from global MRAN

Subnetwork	No. of miRNAs	No. of links	No. of targets	miRNA family list
1	19	96	1536	miR-30a-5p, miR-30b, miR-30c, miR-30d, miR-30e, miR-30e-5p miR-106a, miR-20a, miR-519d, miR-17-5p, miR-93 miR-302b, miR-302c, miR-373 miR-519b, miR-519c miR-181c miR-19a miR-9
2	8	32	361	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7g, let-7i
3	6	20	403	miR-15a, miR-15b, miR-16, miR-195, miR-424 miR-503
4	2	4	100	miR-107, miR-103
5	4	7	552	miR-130a, miR-130b, miR-301 miR-106b
6	3	5	214	miR-132, miR-212 miR-194
7	3	5	349	miR-29a, miR-29b, miR-29c
8	4	6	105	miR-185 miR-198 miR-326 miR-7b
9	5	7	653	miR-26a, miR-26b miR-181a, miR-181d miR-101
10	3	3	10	miR-302d, miR-520b miR-349
11	2	2	33	miR-422a, miR-422b
12	3	3	132	miR-10a, miR-10b miR-339

CMM 1 has many significantly enriched pathways both in KEGG and BioCarta and other CMMs show small numbers of differently enriched pathways. One noteworthy pattern in both heat maps is the finding that functionally enriched pathways (in red) show little overlap between CMMs (Figure 4). It seems evident from the heat maps that CMMs extracted by mining the global MRAN exhibit functional landscape of mutually exclusive CMMs.

Properties of the condition-dependent miRNA association network

One miRNA may be involved in many CMMs in a condition-specific and combinatorial manner. Figure 3 illustrates CMMs 1 to 4 from Table 3. Expression profiles of miR-93 and miR-106a show very high correlation coefficient (>0.8) as well as high target overlap score (*i.e.*, shared target genes and inverse co-expression patterns with their shared target transcript) both in prostate cancer and circadian rhythm. So do the expression profiles of miR-93 and miR-17-5p in circadian rhythm. It seems that the function of a miRNA in the context of cooperativity can better be defined by its interactions with other miRNAs (or ‘the company it keeps’) rather than by its individual characteristics. Microarray

analysis by Volinia *et al.* for determining miRNA signatures in prostate cancer includes both miR-93 and miR-106a, which have well-characterized cancer associations [17]. Several lines of evidences suggest that miR-93 may have different partners in different conditions. miR-93 and miR-130b affect the proliferation and survival of HTLV-1-infected/transformed cells [18]. miR-93 and miR-98 are expressed at higher levels in small-cell than in non-small-cell lung cancer cell lines and immortalized human bronchial epithelial cells (HBEC) [19]. MiRNAs-93, 92, 21, 126 and 29a were significantly over-expressed in the serum from ovarian cancer patients compared to controls [20].

The three miRNA families in CMM 3 are in fact regarded as a larger family of miR-15/107 group. These miRNAs are involved in cell division, metabolism, stress response, and angiogenesis in vertebrate species and have been implicated in human cancers, cardiovascular disease and neurodegenerative disease, including Alzheimer’s disease [21]. Membership in this group is defined based on sequence similarity near the mature miRNAs’ 5’ end: all include the sequence AGCAGC. While all vertebrates studied to date express miR-15a, miR-15b, miR-16, miR-103, and miR-107, mammals alone are known to express miR-195, miR-424, miR-497, miR-503, and miR-646.

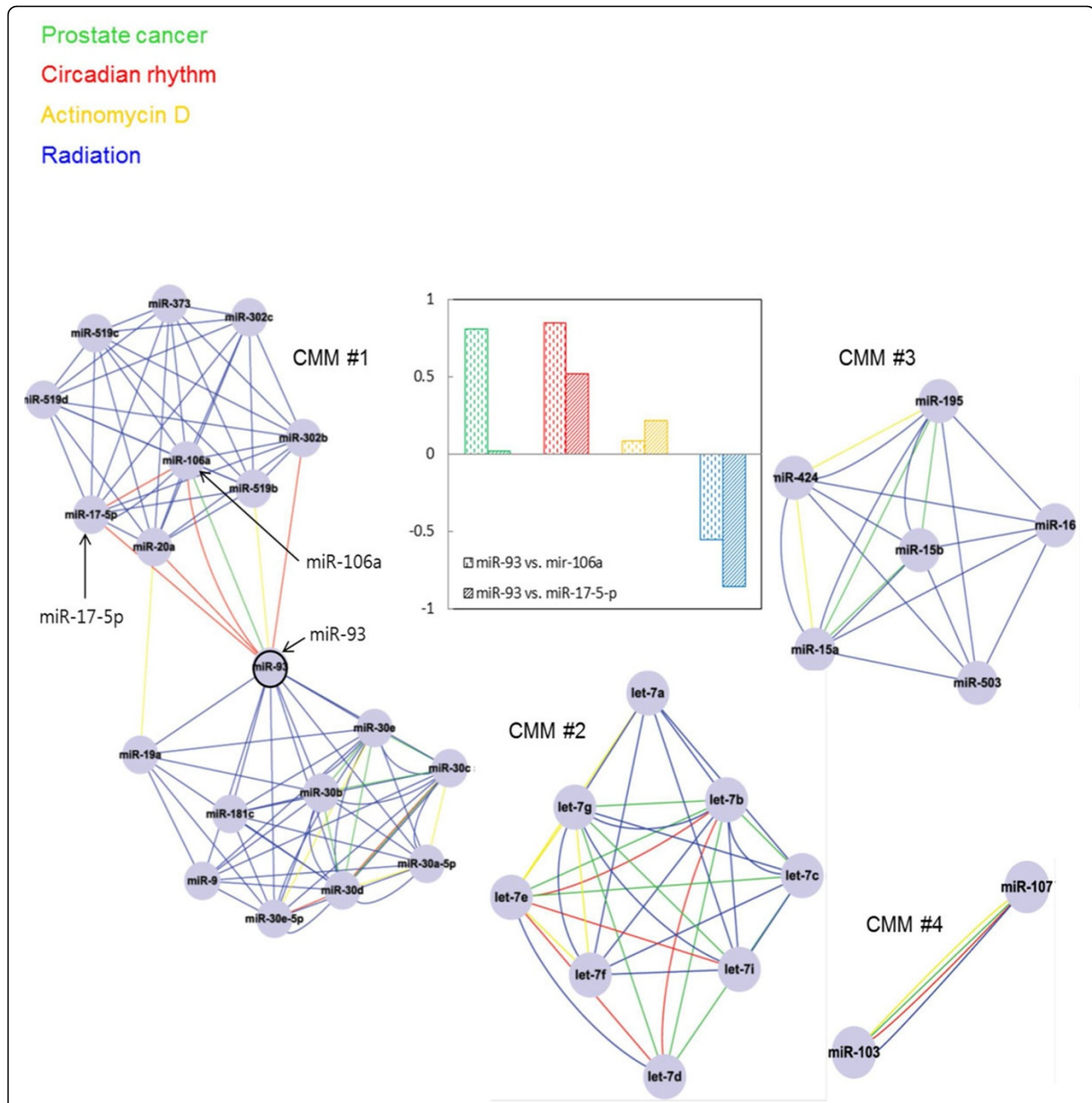


Figure 3 One miRNA participating in multiple condition specificity. The four representative CMMs identified with MCODE. The layout of this network was produced with Cytoscape [48]. Expression profiles of miR-93 and its partner miRNAs according to different conditions show correlation patterns.

Sequence and structural similarities within miRNA association network

Recognition of only 6 ~ 7 nt base pair in the seed sequence of the miRNA are enough to induce functional inhibition of the target gene [22]. Although critical points for target recognition are the short sites that match the seed region of the miRNA, the possible role of the secondary structure of miRNA cannot be overlooked in post-transcriptional

regulation of miRNA expression. If we assume that the distributions of the similarities obtained within CMMs, between CMMs, and among random miRNAs is a normal distribution, the appropriate test is Wilcoxon signed rank test. Table 4 shows the results from statistical comparison of the distributions of the similarities. As expected, seed sequences within CMMs were significantly more similar than those between CMMs. In contrast, mature miRNA

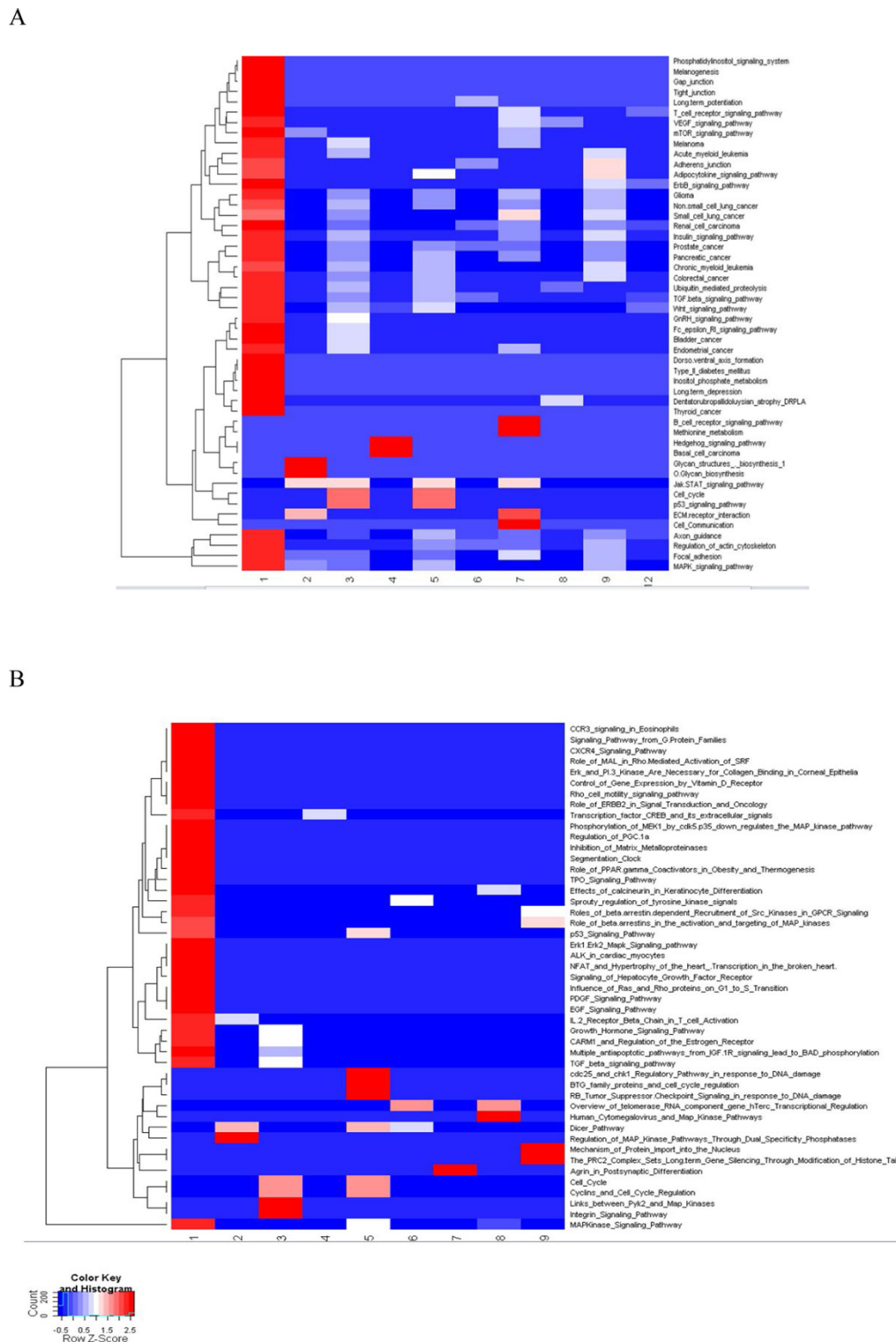


Figure 4 Heatmap of over-represented biological pathways of target genes of CMM. Color gradient represents statistical significance as the $-\log_{10}$ (p-value) in the hypergeometric test for enrichment analysis using (A) KEGG pathway and (B) BioCarta pathway. Red indicates significant associations while blue indicates insignificant associations.

Table 4 Sequence and structure similarities of cooperative miRNA modules

		Comparison between CMMs		CMMs vs. random miRNAs	
		<i>H. sapiens</i>	<i>M. musculus</i>	<i>H. sapiens</i>	<i>M. musculus</i>
Mature form	seed sequence	1.87E-01 *	2.59E-01 *	5.50E-02 *	1.18E-01 *
	mature sequence	3.67E-01	8.77E-01	8.97E-01	9.96E-01
Precursor form	Sequence	6.21E-01	4.59E-01	3.52E-01	1.35E-01
	secondary structure	3.19E-08 *	1.06E-03 *	< 2.2e-16*	< 2.2e-16*

* $p < 0.005$

sequence did not achieve statistical significance. Neither did precursor miRNA sequence.

Very interestingly, however, secondary structure of precursor miRNA showed statistically significant differences such that precursor structure comparison showed even much smaller p -values than seed sequence comparison, which is believed to be the primary regulatory element of miRNA. Although CMMs tend to include miRNA families as shown in Table 3, structural similarity of pre-miRNAs within CMM cannot be explained by miRNA families because many of them originate from different precursors. Moreover, a CMM tends to contain more than one miRNA families and structural similarity of pre-miRNA is even bigger than seed sequence similarity (Table 3). It is suggested that pre-miRNA secondary structure may be involved in the expression of mature miRNA's function.

Evaluation

To evaluate the efficiency of our pipeline, we compared target gene enrichment scores of our miRNA modules with those of miRNA modules from previous study using enrichment scores provided by DAVID Gene Functional Classification [23].

The enrichment scores are intended to order the relative importance of the gene groups. A higher score indicates that the group members are involved in more important (enriched) roles. The enrichment score of each group is measured by the geometric mean of the EASE Scores (modified Fisher Exact) associated with the enriched annotation terms that belong to this gene group.

We used Wilcoxon Rank Sum test to compare target gene enrichment scores between this study and previous study. A statistically significant difference in the target gene enrichment score distribution was observed (p -value = $1.495e-06$). In addition, in order to compare the distribution of enrichment scores of target genes between this study and previous study, we made boxplot and density plot (Figure 5(A,B)). These plots show that the enrichment scores of targets genes in the miRNA modules obtained from this study are statistically significantly higher than those from previous study. In conclusion, we can demonstrate that our pipeline for finding miRNA modules is more efficient than previous method in terms of enrichment of target gene function.

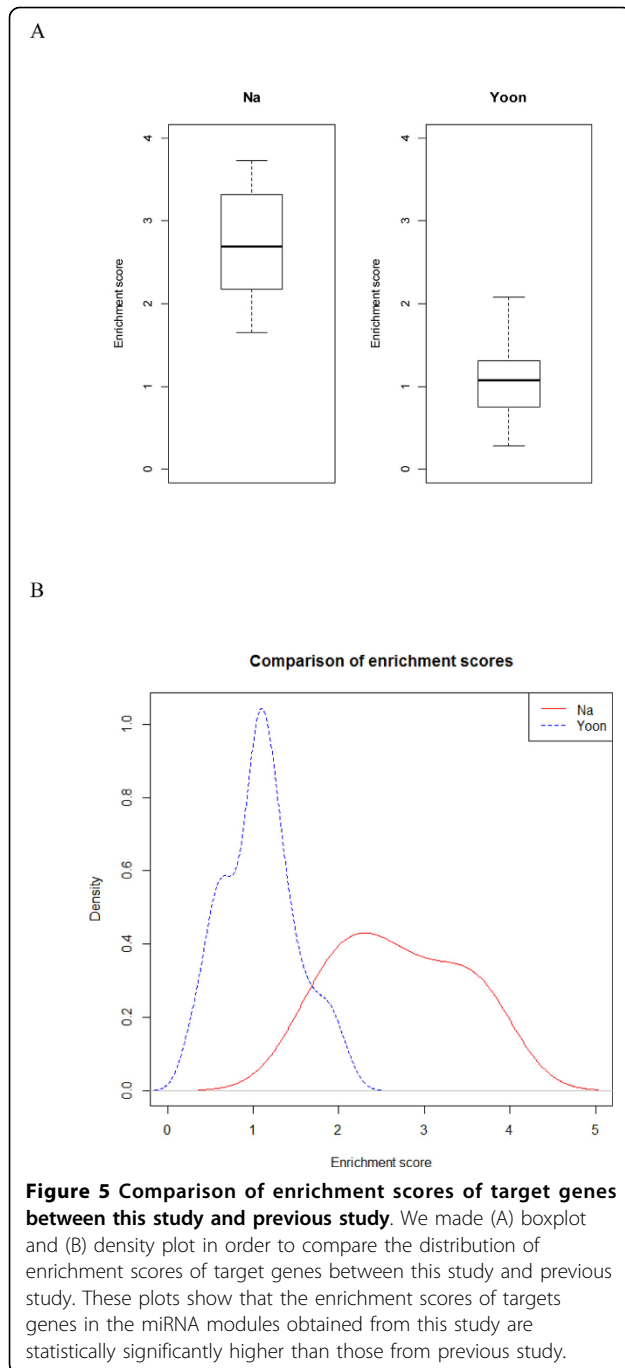
Discussion

MiRNAs can bind to one gene [24-26] and the target sites may overlap to some degree [27]. In many studies, individual effect of a miRNA may appear to be small but when they cooperate, the effect can be of significant proportions [28]. Sætrom *et al.* (2007) observed that lin-41 down-regulation in *C. elegans* requires cooperativity between three miRNA sites [29]. Mavrakis *et al.* (2011) showed that a small set of miRNAs are responsible for the cooperative suppression of several tumor-suppressor genes in T-cell acute lymphoblastic leukemia (T-ALL) [30]. Cooperativity therefore provides the mechanistic basis for reading out combinatorial expression patterns of miRNAs.

Network-based approach of the present study aims to investigate cooperative translational control of miRNAs. We constructed MRAN and showed that CMMs share their transcription factors in the association network. Transcription factors may bind directly to the pri-miR and/or pre-miR to regulate their processing [31]. TFBS within pre-miRs might serve specifically to regulate transcription of the primary miRNA gene transcript (pri-miR) itself as well as transcription of nearby downstream genes. It may be a crucial factor that co-regulating miRNAs act cooperatively to target genes.

It is demonstrated that individual miRNA may interact with different miRNA partners in a condition-specific manner. Some miRNAs may act as 'global facilitator miRNAs' that assist their target-specific partners in their functions. They may inhibit their target genes depending on conditions and partner miRNAs, enabling a post-transcriptional response that integrates multiple environmental signals and pathways.

MiRNA biosynthesis can no longer be viewed as one general pathway universal to all miRNAs. Many steps can be performed in multiple ways, omitted or replaced, and are affected by different mechanisms for individual miRNAs. Most importantly, these specific differences in miRNA processing suggest multiple opportunities for post-transcriptional regulation of miRNA expression [32]. We found that the secondary structures of pre-miRNAs in CMMs are highly similar. On the contrary, pre-miRNAs in CMMs are not similar at the sequence level. The secondary structures of pre-miRNAs have been reported to



serve as a signature motif that is recognized by the nuclear export factor Exportin-5 in the biogenesis of miRNA [33]. Structure-function study raises the possibility of the involvement of pre-miRNA secondary structure into specific functions of targets via different miRNA biogenesis pathways. It has been further reported that the processing of pre-miRNAs is specifically regulated [34]. These differential processes potentially may have caused selection of distinctive structural features that are involved in discriminating

regulatory interactions. Although the seed sequences of miRNAs in CMMs are mostly similar (or identical), there is no one-to-one mapping between CMMs and seed sequences. Instead, CMMs seem to respond to a small group of similar seeds that may be shifted from the usual position (residues 2-8) to, e.g. the positions 1-7 or 4-10.

Our results highlight the potential importance of the secondary structures of pre-miRNAs in both cooperativity and specificity of target genes. Vermeulen *et al.* identified the features of miRNA structure that affects Dicer specificity and efficiency showing that various attributes of the 3' end structure which play a primary role in determining the position of Dicer cleavage [35]. In addition, the functional regulatory networks of miRNAs can provide insights into the intricacies of miRNA processing. The systematic study of identifying co-regulating miRNAs showed that specifically regulated processing of pre-miRNAs may have caused selection of distinctive structural features that are involved in discriminating regulatory interactions. It lends further credibility to the hypothesis that structural subclasses could be associated with processing differences of the precursors.

Conclusions

Cooperative signal integration on target genes of miRNAs is key features of the control of translation by miRNAs. Here, we constructed MRAN using CMMs to investigate miRNA cooperativity based on integration of multiple genomic data sources. The functional regulatory networks of miRNAs can provide insight into the intricacies of miRNA processing. Pre-miRNA secondary structure is suggested to be involved in mature miRNA function. In conclusion, the molecular dissection of miRNA modulation will help to unravel their functional comprehension and escalate one level towards their molecular decoding.

Methods

Overview of finding co-regulating miRNAs

Steps for identifying co-regulating miRNAs are presented in Figure 1. Co-regulating miRNAs correspond to a group of miRNAs silencing together target genes. Sequence complementarity-based computational prediction can consider static target relation only but not cooperative binding of miRNAs that may occur in a condition-specific manner. One needs to systematically investigate expression patterns between miRNAs and mRNAs across different conditions. We used miRNA-mRNA co-expression profiles to select condition-specific modules of cooperative miRNAs at the level of functional expression. Using miRNAs which share common targets, we constructed MRAN and extracted clusters of tightly co-regulating miRNAs from MRAN based on the network analytic approach.

Datasets

We first used miRNA-mRNA co-transcriptomic datasets obtained from the same samples by ourselves in our previous studies (Table 1). Circadian rhythm dataset [12] consists of triplicated mRNA (Illumina Mouse-6 Expression BeadChip) and duplicated miRNA (Ambion mirVana™ miRNA Bioarray V2) expression profiles from 12 time points with four-hour interval across two complete circadian cycles, *i.e.*, 48 hours, resulting 36 and 24 hybridizations for mRNA and miRNA, respectively. One can download the dataset at Gene Expression Omnibus (GSE11516). Radiation dataset [36] consists of triplicated mRNA (Affymetrix Human Gene 1.0 ST Array) and duplicated miRNA (Ambion mirVana™ miRNA Bioarray V2) expression profiles from 6 time points, resulting 18 and 12 hybridizations for mRNA and miRNA, respectively. Two lung cancer cell lines, H460 and H1299 were irradiated at 2 Gy, and harvested after 0, 2, 4, 8, 12 and 24 hours to examine the expressions. Actinomycin D dataset (unpublished data) is created to explore the bio-molecular mechanism of miRNA decay. It consists of triplicated mRNA (Affymetrix Mouse Gene 1.0 ST Array) and duplicated miRNA (Ambion mirVana™ miRNA Bioarray V2) expression profiles from 7 time points, resulting 21 and 14 hybridizations for mRNA and miRNA, respectively. N2a mouse neuroblastoma cells were treated with Actinomycin D to block transcription at 0, 1, 2, 4, 8, 12 and 24 h time points. We also included external dataset from 57 prostate adenocarcinomas (GSE7055) [37] (Table 1).

Normalizations of the miRNA and mRNA expression profiles were performed separately. For datasets using Affymetrix GeneChips, we used RMA (Robust Multi-Array) normalization method [38]. For datasets using Ambion miRNA chips and OSU-CCC chips, we used vsn transformation after background subtraction [39] and applied quantile normalization method [40]. For datasets using Illumina chips, we used chip-wise method using the rank invariant algorithm [41].

Target overlap score and MRAN construction

MRAN is defined as a combined network of cooperative miRNAs sharing condition-specific target genes. A node in MRAN corresponds to a miRNA and an edge corresponds to a condition-specific cooperativity between a miRNA pair. Condition specific cooperativity is defined as condition specific target sharing and determined by (static) significant sequence complementarity by computational target prediction algorithms provided by TargetScan and (functional) significant inverse expression pattern at a specific experimental condition (by Pearson's correlation coefficient). We used the Pearson's correlation coefficient (PCC) for measuring similarity/dissimilarity between expression patterns of miRNAs and genes. The PCC of a pair of genes commonly returns a real value in [+1, -1].

$PCC(x, y) > 0$ represents that x and y are positively correlated with the degree of correlation. On the other hand, $PCC(x, y) < 0$ represents that x and y are negatively correlated with a value $|PCC(x, y)|$. A positive value of the PCC indicates that miRNAs and genes are co-expressed and a negative value of the PCC indicates that opposite expression pattern exists between them. Considering the conservation of miRNA-target relationship, we retrieved and applied 50,339 and 50,349 human and mouse miRNA-target pairs for the 162 and 7,927 conserved miRNAs and mRNAs by using TargetScan target prediction database (release 4.2) [42].

Target overlap score between a pair of miRNAs is defined as Jaccard similarity coefficient, representing the fraction of shared condition specific targets. We define the target overlap score t_{ij} between miRNAs i and j as follows:

$$t_{ij} = \begin{cases} \frac{|Targets(i) \cap Targets(j)|}{|Targets(i) \cup Targets(j)|}, & i \neq j \\ 1, & i = j \end{cases}$$

where $Targets(i)$ represents to the set of targets of miRNAs i . We measured the statistical significance of the Jaccard similarity coefficient by using the exact randomization tests [43]. The $n \times n$ condition specific target overlap matrix $T = [t_{ij}]$ is transformed into an $n \times n$ adjacency matrix $A = [a_{ij}]$, which encodes the connection strengths between pairs of nodes based on the probabilities related to the Jaccard similarity coefficient. Since the networks considered here are undirected, A is a symmetric matrix with non-negative entries.

Figure 1 outlines the steps of the present study. Figure 1 (B) conceptualizes (1) the miRNA-mRNA target relationship as a directed bi-partite graph at the sequence level. (2) Many condition-specific directed bi-partite graphs are emerging by combining inversely co-expression patterns between miRNA and mRNA pairs. (3) By means of applying target overlap scoring, many condition-specific MRANs are obtained as undirected weighted graph of miRNAs. (4) Many condition-specific MRANs are merged into a combined multiple graph to form global MRAN (see Figure 2(A)). Now one can extract cooperative miRNA modules (CMMs) by applying subnetwork detection algorithms and then characterize and evaluate them in terms of biological relevance.

Extracting cooperative miRNA modules

In the present study, a subnetwork detection algorithm, called MCODE [12], was applied to detect coherent groups in the global MRAN. The MCODE is a graph theoretic clustering algorithm specifically designed to find complexes by identifying densely connected subgraphs in networks. MCODE algorithm consists of three stages: vertex weighting, complex prediction and an optional

post-processing step. The weighting of nodes is based on the core clustering coefficient. Once the weights are computed, the algorithm traverses the weighted graph in a greedy fashion to isolate densely connected regions. The post-processing step filters or adds nodes based on connectivity criteria. MCODE has a parameter that specifies the size of clusters returned. All MCODE parameters are applied with default values. Subnetworks are filtered if they do not contain at least a 2-core (graph of minimum degree 2). This approach allows us to assign one miRNA to multiple clusters, considering a biological principle that miRNAs can have multiple functions.

Regulators of cooperative miRNAs

To investigate the modular nature of global MRAN, we used predicted binding sites for all Position Specific Scoring Matrices (PSSMs) from TRANSFAC version 8.3, as they are defined by the UCSC hg17 genome assembly, in the tfbsConsSites (<http://genome.ucsc.edu/>) and tfbsConsFactors. All RefSeq genes genomic locations were taken from hg17. In mouse, regulators of miRNAs were identified in the human genome in the regions orthologous to the mouse. For example, hsa-miR-101 in human and mmu-miR-101 in mouse are orthologous. To assess the statistical significance of the rate of co-occurrence of motif pairs, we used cumulative hypergeometric distribution to calculate the probability of obtaining the rate of co-occurrence, C , equal to or higher than the observed rate of co-occurrence, c' , by chance:

$$P(C \geq c') = \sum_{i=c'}^{\min(m_1, m_2)} \frac{\binom{m_1}{i} \binom{N - m_1}{m_2 - i}}{\binom{N}{m_2}}$$

where m_1 and m_2 denote the numbers of miRNAs having each of the two motifs, N denotes the total number of miRNAs in the genome, and i the summation index.

In the co-occurrence graph of regulators on MRAN, nodes represent transcription factors (TFs), and edges statistical significance of co-occurrences. Hypergeometric p -values were used as weights (see Figure 2(B)).

Functional coherence analysis of target genes

Plausible characteristics of an extracted subnetwork as a CMM might be functional coherence, meaningfulness as well as distinctness from other modules. We performed functional enrichment analysis by using KEGG (Kyoto Encyclopedia of Genes and Genomes) [44] and BioCarta (<http://www.biocarta.com>) pathways to test their functional coherence and meaningfulness. KEGG pathways mainly include cellular processes related to metabolism and biosynthesis. Those on BioCarta cover a wider variety of cellular processes including a large number of signal

transduction and immune signalling pathways as well as metabolic and biosynthetic pathways. Pathways with p -value < 0.05 as revealed by the hypergeometric test were considered statistically significant in the present study. To visualize the distinctness of the cooperative modules in terms of functional enrichment, we created heat maps of the significances (see Figure 4).

Sequence/structure analysis of the co-regulating miRNAs

Structural similarity of the members of CMMs was evaluated. T-COFFEE (version 5.05) [45] was used to calculate pairwise sequence similarity between miRNAs. It uses information from a pre-compiled library of different pairwise alignments including both local and global alignments. RNAdistance program of the Vienna RNA package (Version 1.7.1) [46] was applied to calculate structural distances. Sequence-structure based clustering using the LocARNA-RNAclust pipeline [47] was applied to pre-miRNA sequences. LocARNA uses a complex RNA energy model for simultaneous folding and sequence/structure alignment of the RNAs [47]. The resulting alignment scores can be used to cluster RNAs according to their sequential and structural similarities.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YJN designed the study, performed data analysis and wrote the paper. JHK supervised the research.

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012-0000994). YJN's educational training was supported in part by a grant of the Ministry of Health and Welfare (A112020) and by the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (No. 2012M3A9D1054622).

Declarations

Publication of this article was funded by the Ministry of Education, Science and Technology, Korea; the Ministry of Health and Welfare, Korea and the Ministry of Science, ICT & Future Planning, Korea.

This article has been published as part of *BMC Genomics* Volume 14 Supplement 5, 2013: Twelfth International Conference on Bioinformatics (InCoB2013): Computational biology. The full contents of the supplement are available online at <http://www.biomedcentral.com/bmcgenomics/supplements/14/S5>.

Authors' details

¹Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, Seoul 110-799, Korea. ²Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul 110-799, Korea.

Published: 16 October 2013

References

1. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**(2):281-297.
2. Hammond SM: **MicroRNAs as oncogenes.** *Curr Opin Genet Dev* 2006, **16**(1):4-9.

3. Esquela-Kerscher A, Slack FJ: **Oncomirs - microRNAs with a role in cancer.** *Nat Rev Cancer* 2006, **6**(4):259-269.
4. Gregory RI, Shiekhattar R: **MicroRNA biogenesis and cancer.** *Cancer Res* 2005, **65**(9):3509-3512.
5. Alvarez-Garcia I, Miska EA: **MicroRNA functions in animal development and human disease.** *Development* 2005, **132**(21):4653-4662.
6. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS: **Human MicroRNA targets.** *PLoS Biol* 2004, **2**(11):e363.
7. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS: **MicroRNA targets in *Drosophila*.** *Genome Biol* 2003, **5**(1):R1.
8. Siggerson K, Tang L, Moss EG: **Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene *lin-28* after translation initiation.** *Dev Biol* 2002, **243**(2):215-225.
9. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, et al: **Combinatorial microRNA target predictions.** *Nat Genet* 2005, **37**(5):495-500.
10. Boross G, Orosz K, Farkas IJ: **Human microRNAs co-silence in well-separated groups and have different predicted essentialities.** *Bioinformatics* 2009, **25**(8):1063-1069.
11. Xu J, Li CX, Li YS, Lv JY, Ma Y, Shao TT, Xu LD, Wang YY, Du L, Zhang YP, et al: **MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features.** *Nucleic Acids Res* 2011, **39**(3):825-836.
12. Bader GD, Hogue CW: **An automated method for finding molecular complexes in large protein interaction networks.** *BMC Bioinformatics* 2003, **4**:2.
13. Moncini S, Salvi A, Zuccotti P, Viero G, Quattrone A, Barlati S, De Petro G, Venturin M, Riva P: **The role of miR-103 and miR-107 in regulation of CDK5R1 expression and in cellular migration.** *PLoS One* 2011, **6**(5):e20038.
14. Turner MJ, Slack FJ: **Transcriptional control of microRNA expression in *C. elegans*: promoting better understanding.** *RNA Biol* 2009, **6**(1):49-53.
15. Xu H, He JH, Xiao ZD, Zhang QQ, Chen YQ, Zhou H, Qu LH: **Liver-enriched transcription factors regulate microRNA-122 that targets CUTL1 during liver development.** *Hepatology* 2010, **52**(4):1431-1442.
16. Zhao Y, Srivastava D: **A developmental view of microRNA function.** *Trends Biochem Sci* 2007, **32**(4):189-197.
17. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, et al: **A microRNA expression signature of human solid tumors defines cancer gene targets.** *Proc Natl Acad Sci USA* 2006, **103**(7):2257-2261.
18. Yeung ML, Yasunaga J, Bannasser Y, Dusetti N, Harris D, Ahmad N, Matsuoka M, Jeang KT: **Roles for microRNAs, miR-93 and miR-130b, and tumor protein 53-induced nuclear protein 1 tumor suppressor in cell growth dysregulation by human T-cell lymphotropic virus 1.** *Cancer Res* 2008, **68**(21):8976-8985.
19. Du L, Schageman JJ, Subauste MC, Saber B, Hammond SM, Prudkin L, Ji L, Roth JA, Minna JD, et al: **miR-93, miR-98, and miR-197 regulate expression of tumor suppressor gene FUS1.** *Mol Cancer Res* 2009, **7**(8):1234-1243.
20. Resnick KE, Alder H, Hagan JP, Richardson DL, Croce CM, Cohn DE: **The detection of differentially expressed microRNAs from the serum of ovarian cancer patients using a novel real-time PCR platform.** *Gynecol Oncol* 2009, **112**(1):55-59.
21. Finnerty JR, Wang WX, Hebert SS, Wilfred BR, Mao G, Nelson PT: **The miR-15/107 group of microRNA genes: evolutionary biology, cellular functions, and roles in human diseases.** *J Mol Biol* 2010, **402**(3):491-509.
22. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB: **Prediction of mammalian microRNA targets.** *Cell* 2003, **115**(7):787-798.
23. Yoon S, De Micheli G: **Prediction of regulatory modules comprising microRNAs and target genes.** *Bioinformatics* 2005, **21**(Suppl 2):ii93-100.
24. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G: **The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*.** *Nature* 2000, **403**(6772):901-906.
25. Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D: **MicroRNAs and other tiny endogenous RNAs in *C. elegans*.** *Curr Biol* 2003, **13**(10):807-818.
26. Vella MC, Choi EY, Lin SY, Reinert K, Slack FJ: **The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the *lin-41* 3'UTR.** *Genes Dev* 2004, **18**(2):132-137.
27. Doench JG, Sharp PA: **Specificity of microRNA target selection in translational repression.** *Genes Dev* 2004, **18**(5):504-511.
28. Chhabra R, Dubey R, Saini N: **Cooperative and individualistic functions of the microRNAs in the miR-23a~27a~24-2 cluster and its implication in human diseases.** *Mol Cancer* 2010, **9**:232.
29. Saetrom P, Heale BS, Snove O Jr, Aagaard L, Alluin J, Rossi JJ: **Distance constraints between microRNA target sites dictate efficacy and cooperativity.** *Nucleic Acids Res* 2007, **35**(7):2333-2342.
30. Mavrakis KJ, Van Der Meulen J, Wolfe AL, Liu X, Mets E, Taghon T, Khan AA, Setti M, Rondou P, Vandenberghe P, et al: **A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL).** *Nat Genet* 2011, **43**(7):673-678.
31. Piriyaopongsa J, Jordan IK, Conley AB, Ronan T, Smalheiser NR: **Transcription factor binding sites are highly enriched within microRNA precursor sequences.** *Biol Direct* 2011, **6**:61.
32. Winter J, Jung S, Keller S, Gregory RI, Diederichs S: **Many roads to maturity: microRNA biogenesis pathways and their regulation.** *Nat Cell Biol* 2009, **11**(3):228-234.
33. Kim VN, Han J, Siomi MC: **Biogenesis of small RNAs in animals.** *Nat Rev Mol Cell Biol* 2009, **10**(2):126-139.
34. Schmittgen TD: **Regulation of microRNA processing in development, differentiation and cancer.** *J Cell Mol Med* 2008, **12**(5B):1811-1819.
35. Vermeulen A, Behlen L, Reynolds A, Wolfson A, Marshall WS, Karpilow J, Khvorova A: **The contributions of dsRNA structure to Dicer specificity and efficiency.** *RNA* 2005, **11**(5):674-682.
36. Lee SY, Choi MK, Lim JM, Wu HG, Kim JH, Park WY: **Differential Expressions of Apoptosis-related Genes in Lung Cancer Cell Lines Determine the Responsiveness to Ionizing Radiation.** *Genomics and Informatics* 2008, **6**(1):36-43.
37. Prueitt RL, Yi M, Hudson RS, Wallace TA, Howe TM, Yfantis HG, Lee DH, Stephens RM, Liu CG, Calin GA, et al: **Expression of microRNAs and protein-coding genes associated with perineural invasion in prostate cancer.** *Prostate* 2008, **68**(11):1152-1164.
38. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Biostatistics* 2003, **4**(2):249-264.
39. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M: **Variance stabilization applied to microarray data calibration and to the quantification of differential expression.** *Bioinformatics* 2002, **18**(Suppl 1):S96-104.
40. Bolstad BM, Irizarry RA, Astrand M, Speed TP: **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.** *Bioinformatics* 2003, **19**(2):185-193.
41. Kuhn K, Baker SC, Chudin E, Lieu MH, Oeser S, Bennett H, Rigault P, Barker D, McDaniel TK, Chee MS: **A novel, high-performance random array platform for quantitative gene expression profiling.** *Genome Res* 2004, **14**(11):2347-2356.
42. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.** *Cell* 2005, **120**(1):15-20.
43. Real R, Vargas JM: **The probabilistic basis of Jaccard's index of similarity.** *Systematic Biology* 1996, **45**:380-385.
44. Kanehisa M: **Pathway databases and higher order function.** *Adv Protein Chem* 2000, **54**:381-408.
45. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment.** *J Mol Biol* 2000, **302**(1):205-217.
46. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P: **Fast Folding and Comparison of Rna Secondary Structures.** *Monatsh Chem* 1994, **125**(2):167-188.
47. Will S, Reiche K, Hofacker IL, Stadler PF, Backofen R: **Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering.** *PLoS Comput Biol* 2007, **3**(4):e65.
48. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13**(11):2498-2504.
49. Na YJ, Sung JH, Lee SC, Lee YJ, Choi YJ, Park WY, Shin HS, Kim JH: **Comprehensive analysis of microRNA-mRNA co-expression in circadian rhythm.** *Exp Mol Med* 2009, **41**(9):638-647.

doi:10.1186/1471-2164-14-S5-S17

Cite this article as: Na and Kim: Understanding cooperativity of microRNAs via microRNA association networks. *BMC Genomics* 2013 **14** (Suppl 5):S17.