miRNA–mRNA network involved in folliculogenesis interactome: systems biology approach

Abolfazl Bahrami¹, Seyed Reza Miraie-Ashtiani¹, Mostafa Sadeghi¹ and Ali Najafi²

¹Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran and ²Molecular Biology Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran

Correspondence should be addressed to A Najafi or S R Miraie-Ashtiani; Email: Najafi74@bsmu.ac.ir or ashtiani@ut.ac.ir

Abstract

At later phases of folliculogenesis, the mammalian ovarian follicle contains layers of granulosa cells surrounding an antral cavity. To better understand the molecular basis of follicular growth and granulosa cell maturation, we study transcriptome profiling of granulosa cells from small (<5 mm) and large (>10 mm) bovine follicles using simultaneous method of Affymetrix microarrays (24,128 probe sets) and RNA-Seq data sets. This study proposes a computational method to discover the functional miRNA–mRNA regulatory modules, that is, groups of miRNAs and their target mRNAs that are believed to take part cooperatively in post-transcriptional gene regulation under specific conditions. The reconstructed network was named Integrated miRNA–mRNA Bipartite Network. 277 genes and 6 key modules were disclosed through clustering for mRNA master list. The 66 genes are among the genes that belong to at least two modules. All these genes, being involved in at least one of the phenomena, namely cell survival, proliferation, metastasis and apoptosis, have an overexpression pattern (P < 0.01). For miRNA master list, a total of 172 sequences were differentially expressed (P < 0.01) between dominant (large) and each of subordinate (small) follicles. Within the follicle, these miRNAs were predominantly expressed in mural granulosa cells. Finally, predicted and validated targets of these miRNAs enriched in dominant (large) follicles were identified, which are mapped to signaling pathways involved in follicular cell proliferation, steroidogenesis, PI3K/AKT/mTOR and Ras/Raf/MEK/ERK. The identification of miRNAs and their target mRNAs and the construction of their regulatory networks may give new insights into biological procedures.

Reproduction (2017) 154 51–65

Introduction

The mammalian ovary produces oocytes for fertilization and the hormones progesterone and estradiol. Oocytes mature in ovarian follicles surrounded by pre-granulosa cells at the primordial follicle stage and by granulosa cells, which commence replicating at the primary follicle stage. Once activated, the follicle grows by development of the oocyte and replication of the granulosa cells from about 20 to 50 million cells in the cow (Rodgers et al. 1999). During growth, a fluid-filled antrum also develops in the middle of the follicle (Rodgers & Irving-Rodgers 2010) and bovine follicles need to develop to over 10–12 mm in diameter, finally by antrum expansion, to be capable of ovulation. However, instead of one primordial follicle growing to the necessary size and then ovulating, many follicles start growth during the course of the cycle. The majority of these growing follicles become atretic leaving in cows only one or two follicles to ovulate. The process of follicle growth during a cycle is not random either as two or three waves of follicles emerge from a pool of follicles of approximately 4–5 mm in diameter during each estrous cycle (Ginther et al. 1996). During these waves, follicles continue to enlarge over several days until one follicle that is growing faster and is larger than the others gains dominance (Ginther et al. 2001). As the larger dominant follicle continues to expand, the smaller follicles in the wave undergo atresia. The process of atresia in large follicles to have expanded an antral cavity is characterized by death of the mural granulosa cells with the attendance of pyknotic nuclei followed by loss of these layers into the antrum (Irving-Rodgers et al. 2011). The cell death processes can involve necrosis, apoptosis, comification and autophagy, and any of the major cell types of the follicle can be involved, depending upon the stage of follicular development when atresia occurs (Rodgers & Irving-Rodgers 2010).

MicroRNAs (miRNAs) are a category of single-stranded, non-coding RNAs with ~20–23 nucleotides in length. miRNAs target coding mRNAs through complementary base-pairing for repressing translation, cleavage and fathering protein degradation (Bartel 2004, He & Hannon 2004). Expanding number of evidence suggests that miRNAs play important roles in cell proliferation, differentiation, growth, mobility and apoptosis (Ambros 2004, Du & Zamore 2007, Bushati & Cohen 2007, Najafi et al. 2016). Recent studies have
reported differentially regulated miRNAs in different traits. Thus, detecting miRNAs and their target mRNAs, and further building their regulatory networks may give new insights into biological procedures. Previous works mostly focus on the genomewide discovery of miRNAs and prediction of putative target mRNAs (Krek et al. 2005). These methods have identified a large number of miRNAs and their target mRNA or genes. For example, miR base has deposited numerous miRNAs and their target genes from several species (Griffiths-Jones et al. 2008) up to date. Thus, the detection of the functional regulation networks that associate miRNAs and their target mRNAs or genes with conditions and integrated miRNA–mRNA network at different stages of traits is crucial for understanding the regulatory mechanism in complex cellular systems. Yoon and De Micheli (Yoon & De Micheli 2005) suggested a prediction method for miRNA regulatory modules at the sequence level. Their method is based on a seeing that the binding strength of miRNAs and target mRNAs are modest and similar when multiple binding sites exist on a target. Huang and coworkers (Huang et al. 2006) and Joung and coworkers (Joung et al. 2007) integrate both expression profiles and sequence information of miRNAs and mRNAs to detect the relevant miRNA–mRNA pairs. The integrated approaches potentially decrease false discovery rate and facilitate the interventional experiments to validate the targets of miRNAs. Huang and coworkers (Huang et al. 2006) proposed Bayesian networks to model the regulatory mechanism of miRNAs. Joung and coworkers (Joung et al. 2007) utilized a bi-clustering approach to identify MRMs (miRNA–mRNA regulatory modules). The binding strength between miRNAs and mRNAs based on the complementary base-pairing as well as their expression profiles across different conditions is taken into account. They showed that using diverse resources including expression profiles and sequence information of miRNAs and mRNAs can attain better prediction. On the other hand, systems biology is a system-level approach to surve a complete understanding of complex biological systems outside the molecular-level scale (Najafi et al. 2014). In lieu of analyzing individual aspects or components of the organism, such as metabolism or a cell nucleus, systems biologists focus on all the aspects and the interactions among them, all as part of one system. These interactions are finally responsible for an organism’s functions and form.

Unlike previous works as discussed previously, this work detects miRNAs and mRNAs regulatory modules within differentiated conditions. This study suggests a computational method to identify and discover the integrated miRNA–mRNA bipartite networks (IMBNs), that is, groups of miRNAs and their target genes or mRNAs, which are believed to participate in post-transcriptional gene regulation under specific conditions such as folliculogenesis. The concept of IMBNs is an extension of MRMs introduced by Yoon and coworkers (Yoon & De Micheli 2005) but distinct from their work as it associates MRMs with conditions. In our study, we intend to identify the group of miRNA–mRNA target pairs that associate with certain condition or trait, such as folliculogenesis. Associated miRNA–mRNA modules with conditions or traits in the model may help detect the functional groups that are involved in conditions or traits directly, thus decreasing the false discovery rate of target prediction as well as to build integrated miRNA–mRNA bipartite networks. Accordingly, we have prepared a master list of folliculogenesis-related miRNAs and target genes. Protein–protein interaction networks (PPI), gene regulatory networks (GRN) and IMBNs were then generated and topological analysis was performed on the merged network. This work is the first reconstruction of an integrated miRNA–mRNA bipartite networks and detection of functional modules in the bovine folliculogenesis pathways.

Materials and methods
Static network reconstruction plays a central role in systems biology. Therefore, the reconstruction process is highly related to our knowledge base in characterizing folliculogenesis. However, owing to the large amount of data, we need appropriate computational tools for gathering, constructing and analyzing generated networks and pathways (Fig. 1).

Data collection
We have had two types of high-throughput data sets in our study; first, data set relates to microarray data and the second to RNA-Seq data. Hence we have analyzed the data sets separately. In all cases, the HGNC (HUGO Gene Nomenclature Committee) was used to identify gene and protein symbols.

Gene expression microarray and RNA-Seq data
Datasets for gene expression microarray of granulosa cells were obtained from GEO database. For including the most number of genes possible and in order to maximize the number of databases, the Affymetrix GeneChip Bovine Genome Array (Bos taurus) were downloaded from either NCBI GEO (Barrett et al. 2013) or EBI Array Express (Kolesnikov et al. 2015) and used to build the main master lists and various datasets were applied to denote multiple perturbed states in folliculogenesis. In total, 209 samples were chosen. 83 samples from seven datasets i.e. GSE56145 (Platforms: GPL13226; Agilent-028298 Embryovine Bovine 45K microarray) (Douville & Sirard, 2014), GSE39589 (Platforms: GPL2112; Affymetrix Bovine Genome Array) (Hatzirodos et al. 2014), GSE50854 (Platforms: GPL13226; Agilent-028298 Embryovine Bovine 45K microarray) (Golini et al. 2014), GSE63904 (Platforms: GPL13226;
Agilent-028298 Embryogene Bovine 45K microarray (Girard et al., 2015), GSE63919 (Platforms: GPL13226; Agilent-028298 Embryogene Bovine 45K microarray) (Girard et al., 2015), GSE2598 (Platforms: GPL13226; Agilent-028298 Embryogene Bovine 45K microarray) (Mihm et al., 2008), GSE69247 (Platforms: GPL13226; Agilent-028298 Embryogene Bovine 45K microarray) (Gilbert et al., 2011) and GSE11495 (Platforms: GPL2112; Affymetrix Bovine Genome Array) (Skinner et al., 2008) datasets as well as 43 samples were chosen from two datasets i.e. GSE55890 (Platforms: GPL18422; Exiqon miRCURY LNA microRNA array 6th generation product no. 34013 mirBase v18) (Zielak-Steciwko et al., 2014) and GSE54692 (Platforms: GPL11434; miRCURY LNA microRNA Array, 6th generation – hsa, mmu and mo) (Sontakke et al., 2014) datasets of dominant and subordinate follicles for obtaining mRNA and miRNA main master list for microarray data, respectively, and 83 samples from three dataset i.e. GSE34317 (Platforms: GPL11153; Illumina Genome Analyzer II) (Walsh et al., 2012), GSE55987 (Platforms: GPL15749; Illumina HiSeq 2000) (Salilew-Wondim et al., 2014) and GSE56002 (Platforms: GPL15749; Illumina HiSeq 2000) (Gebremedhyn et al., 2015) datasets of dominant and subordinate follicles for obtaining mRNA and miRNA main master list for RNA-Seq data, respectively, were selected.

**Gene expression microarray analysis**

Arrays from individual experiments were preprocessed with R lumi package (Du et al., 2008). Lumi R package is designed to pre-process the microarray data. It includes data input, quality control, variance stabilization, normalization and gene annotation part. After data pre-processing, differential gene expression was evaluated using R limma package ($P<0.01$). Volcanic plot was constructed using R language. Data validation was performed utilizing Array Mining web-sever (www.arraymining.net) with different statistical methods such as eBayse and ENSEMBLE. Finally, master list 1 was built for both mRNA and miRNA data sets.
RNA-Seq analysis

Quality assessment of raw sequencing data and adapter trimming

Quality evaluation of the raw sequence data was performed using FastQC, a freely available sequence analysis tool (http://www.bioinformatics.babraham.ac.uk/publications.html) and FASTQ Groomer (Cock et al. 2010). For this, the raw fastq data were imported into the program and the basic statistics, the sequence quality, quality score and sequence content were evaluated per base basis. The raw sequence data that satisfies the basic requirements of quality parameters were used for downstream analysis. The adapters, PCR primers and non-informative sequences were removed from downstream analysis using the Trimmomatic software (Bolger et al. 2014).

Sequence alignment and identification of known miRNAs and mRNAs

Sequence alignment and detection of known and novel microRNAs and mRNAs were performed using TopHat2 software (Kim et al. 2013). Prior to performing sequence alignment, the BosTau8 assembly of the bovine genome (UMD 3.1.1) was retrieved from UCSC Genome Browser (https://genome.ucsc.edu/) and indexed using Bowtie 2, a freely available tool for aligning sequencing reads to long reference sequences (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml). The sequence reads were then aligned to the indexed bovine reference genome.

Assemble and analyze transcripts

After mapping the reads, the next step was to assemble the reads into complete transcripts that can be analyzed for differential expression and phenomena such as splicing events and transcriptional start sites. For this, we have used Cufflinks and Cuffmerge software (Trapnell et al. 2010).

Analyzing differentially expressed transcripts

Following detection and identification of transcripts, the expression level of those transcripts was compared between the granulosa samples of subordinate follicle (SF) and dominant follicle (DF) to understand whether accumulation or degradation of transcripts is associated with follicular selection, recruitment and dominance.

Differentially expressed transcripts were analyzed from read counts of the samples using Cuffdiff software (Trapnell et al. 2010). For this, normalization of the data was performed in such a way that the raw read count of each transcript was multiplied by the sample size factor that was calculated as ratios of observed count. For each transcript in each granulosa sample, the observed count is the ratio of raw count for each transcript to the geometric mean across the samples. At the end of analysis, transcripts with log2 fold change difference >2.0, with P value < 0.01 and false discovery rate (FDR) < 10% were considered as differentially expressed transcripts. For visualizing and interpreting RNA-Seq results, Integrative Genomics Viewer (IGV) tool was used (James et al. 2011).

Finally, master list 2 was built for both mRNA and miRNA RNA-Seq data sets.

Literature mining

Extensive literature surveys were then used to identify and confirm factors (gene, protein and metabolite) involved in folliculogenesis. Literature mining was performed using keywords including bovine folliculogenesis in PubMed and Google Scholar databases without time limitation. Folliculogenesis studies were selected for further detailed review. Finally, after removing redundant reports, iHOP (Fernandez et al. 2007) is a Web-based tool and Agilent literature search is a plugin in Cytoscape (Smoot et al. 2011) that was used to automate analysis of abstracts in search for gene names. Finally, data were manually curated.

Gene annotation

Gene set annotation enrichment analysis was also performed using DAVID (http://david.abcc.ncifcrf.gov) and g:Profiler (http://biit.cs.ut.ee/gprofiler/) web tools that provides a set of functional annotation tools for the genes categorized into gene ontology (GO) terms.

Target gene prediction and validation of differentially expressed miRNAs

The functional annotation of differentially expressed miRNAs was analyzed based on the functional annotation of their potential target genes. The predicted and validated target genes were identified using miRWalk 2.0 (Dweep et al. 2015), a comprehensive atlas of microRNA-target interactions tool which integrates 12 miRNA target prediction tools. The target genes that were predicted and validated by at least five prediction tools were selected and submitted to The Database for Annotation, Visualization and Integrated Discovery (DAVID) and the potential Kyoto Encyclopedia of Genes and Genomes (KEGG), Protein AAnalysis Throught Evolutionary Relationships (PANTHER) and Reactome pathways databases significantly enriched by target genes of each miRNAs were identified.

Integrated miRNA–mRNA Bipartite Network reconstruction

Integrated miRNA–mRNA bipartite network was reconstructed based on the molecular interactions and the master gene list documented in related papers and online interaction databases. The Pathway Resource List (http://pathguide.org) is a meta-database that provides an overview of more than 300 web-accessible biological pathway and network databases (Bader et al. 2006). Protein–protein interaction (PPI) data were abstracted from Database of Interacting Proteins (DIP) (Xenarios et al. 2002), Biomolecular Interaction Network Database (BIND) (Bader et al. 2003), Mammalian Protein–Protein Interactions Database (MIPS) (Pagel et al. 2005) and Biological General Repository for Interaction Datasets (BioGRID) (Chat-
Bovine folliculogenesis interactome

Aryamontri et al. 2013). Furthermore, pathway data were obtained from the analysis of related experiments and search in pathway database such as Gene MANIA (Mostafavi et al. 2008) and Search Tool for the Retrieval of Interacting Genes/Proteins (STRING). String uses eight major sources of interaction/association data (neighborhood, co-occurrence, fusion, experimental, co-expression, database and text mining) to define the interaction between proteins using a probabilistic confidence score (Szklarczyk et al. 2015). Each miRNA and target gene was entered into the database manually, and resulting interactions were added to the networks using Cytoscape software. Also, Cytoscape plugins were used for extracting, combining, visualizing and analyzing unified interactive data. Molecular species in generated networks are represented as nodes and the interactions between these nodes as edges.

Some of the achieved interactions were checked by pathway databases including PANTHER, Reactome and KEGG. Also, generated networks were compiled in Simple Interaction Format (SIF; see section on supplementary data given at the end of this article) bowed to Cytoscape for topological analysis. After merging networks, the statistical and topological significance of the network was calculated using the Network Analyzer plugin in Cytoscape. The degree (connectivity) of nodes, the shortest path lengths between any two nodes, the mean path length (the average of the shortest path lengths) and the network diameter in comparison with random networks (Barabasi-Albert and Erdos-Renyi models; generated by random networks) were analyzed.

**Clustering of Integrated miRNA–mRNA Bipartite Network**

Topological characteristics of the integrated miRNA–mRNA bipartite network were examined by Cytoscape 3.4 and for clustering; ClusterONE (Nepusz et al. 2012) and MCODE (Bader & Hogue 2003) were used. ClusterONE, a Cytoscape plugin for clustering, as the clustering method in this section, was used. This algorithm is swift and can be run in a command-line mode, which does not need to load the large genomewide network in Cytoscape. ClusterONE is designed to discover densely connected sub-graphs of a network by maximizing edges within a cluster and minimizing edges between different clusters. It allows the overlapping of sub-graphs, which are necessary in co-expression networks, as a gene may participate in more than one functional module. MCODE is another method for clustering that was used. MCODE is a clustering algorithm, which can be operated for undirected or directed graphs.

**Results**

**Candidate miRNA and gene list**

Following the workflow described in the Methods section for data retrieval, 277 genes (263 genes from microarray and RNA-Seq data sets and 14 genes from literature) and 162 miRNAs were selected according to databases, literature mining, related microarray and RNA-Seq data.

The genes and miRNA that are known to be involved in folliculogenesis are given in Supplementary Tables 1 and 2. These genes are annotated and described based on molecular process and functionality in the gene ontology (GO) and other annotation databases. The volcanic plot for microarray was shown in the Fig. 2. As well as for validation of alignment related to transcripts, result of reads relevant to one of the most important gene (CYP17A1) was show in Fig. 3.

**MicroRNAs were differentially expressed in DF compared to SF**

In order to study the global microRNA (miRNA) expression change in dominant (large) follicles compared with subordinate (small) follicles, we first obtained samples from dominant (large) follicles (n = 23) and subordinate (small) follicles (n = 20) for microarray data sets and dominant (large) follicles (n = 9) and subordinate (small) follicles (n = 9) for RNA-Seq data. All reads were mapped with annotated miRNAs in miRBase database, and approximately 35% of the clean reads were mapped to miRNA in the database, while nearly another 35% of the clean reads were mapped to ribosome RNA (rRNA) or small nucleolar RNAs (snoRNA) when mapped to other small RNA datasets. After normalizing and analyzing all the mapped reads in both groups, we identified the existence of 753 miRNAs altogether. After applying a stringent filtering criteria that compared DF with SF (False discover rate, FDR < 0.01, fold change > 2 and P < 0.01), we identified 89 upregulated and 72 downregulated miRNAs (Supplementary Table 1).
Global transcriptome analysis revealed DF altered gene expressions

To fully decipher the DF miRNA regulatory network, we first analyzed the global transcriptome in both DF and SF. Taking advantage of hierarchical cluster, we found the global gene expression is greatly altered (Supplementary Table 2), with 233 genes significantly upregulated (fold change >2, FDR<0.01) and 44 genes significantly downregulated (Fold change <2, FDR<0.01). However, these terms were not found in the GO analysis in genes downregulated in OPLL. In the KEGG pathway analysis of altered genes, we also found that PI3K/AKT/mTOR signaling pathways are highly enriched in deregulated genes. Taken these data together, we found that DF expresses more cell survival-related genes than SF cells.

Integrated miRNA–mRNA Bipartite Network

We compiled a gene and miRNA list (nodes) involved in folliculogenesis based on PPI resources, literature mining, GRNs and miRNAs–gene targets. Briefly, PPI (miRNAs–Genes) networks are commonly represented in an undirected graph format, with nodes corresponding to proteins (miRNAs and genes) and edges corresponding to physical protein–protein interactions (miRNAs targeted genes). Gene expression data and co-expressed genes have become a fruitful resource in representing the molecular state of cells. GRNs are commonly shown in a directed graph. GRNs consist of nodes, indicating the expression profile of a particular gene, and edges indicating significant associations between expression profiles (for example Pearson correlation). Integrated miRNA–mRNA bipartite network obtained from the master list are given in Cytoscape.sif format in File S4. Based on the current knowledge of interactions in databases, we could identify interactions for 531 nodes and 959 edges in the master list. Also, the related files of networks are presented in Supplementary Table 4. The folliculogenesis network was created by merging networks. The complex network is composed of several functional modules or sub-networks contributing to various biological processes in folliculogenesis pathway. Six major functional modules are illustrated in Figs. 4, 5, 6, 7, 8 and 9.

Network topology analysis

Network parameters were performed by using Network Analyzer plugin of Cytoscape for the IMBN. Topological analysis of a network identifies the global qualitative virtues of the system. Network topology is used to cater the significance of a node in communicating with other nodes. Topological parameters such as topological coefficient, the average clustering coefficient, average degree and power law distribution of degrees and
betweenness centrality have been assessed to predict the topology of the networks. The distribution of clustering coefficient is an important characteristic feature of biological scale-free networks. Therefore, characteristic path length, the clustering coefficient, network centralization and network density of the IMBN were compared with simulated randomized model networks (Erdos–Renyi and Barabasi–Albert models; generated by random networks, a plugin in Cytoscape) as shown in Table 1. The general structure of the network is far from random by the comparison with the simulated network. In particular, the clustering coefficient of the IMBN greatly differs from random network. The characteristic path length donates the expected distance between two connected nodes in each network and the network density and network centralization show how densely the network is populated with edges and reflect the inclination of a network to contain hub nodes.

**Clustering of Integrated miRNA–mRNA Bipartite Network**

To analyze the results, we applied clustering on the Integrated miRNA–mRNA Bipartite Network generated by merged networks. Clustering algorithms are used to find important sub-networks or modules. The results were first obtained from clustering with MCODE and clusterONE. ClusterONE and MCODE were run with its default parameter settings. ClusterONE output was 26 clusters or modules including 124 genes and
32 miRNAs. MCODE output was 14 modules including 89 genes and 27 miRNAs (considering overlapping cluster genes and miRNAs). Some of the clusters turned out to be subclusters of other larger clusters, therefore deleting such subclusters the final number of clusters was reduced from 26 to 6, including 277 genes and 27 miRNAs. Many of the genes and miRNAs are repeated in more than one cluster and overall, there were 66 unique genes out of 6 modules. These clusters are available in Supplementary Table 3. Interestingly, 161

Figure 5 79 genes and 10 miRNAs in module 2; each node denotes a gene (rectangle) and miRNA (octagon) in module and edges depict miRNA–mRNA relationships.
miRNAs and 277 genes were commonly detected in all samples regardless of the follicle stage. Re-processing of two independent miRNAs expression series in folliculogenesis resulted in identification of 15 upregulated and 12 downregulated miRNAs (Table 2).

Discussion

miRNAs comprise one of numerous classes of gene regulatory molecules in multi-cellular organisms. They influence the output of many protein-coding genes by binding to and inhibiting mRNAs. Increasing number of evidences suggest that miRNAs play critical roles in traits including productive traits and processes. Computational methods modeling miRNA–mRNA regulatory modules will help understand the complex biological procedures. The method proposed in this study goes further by associating the different phases with miRNA–mRNA regulatory modules. The IMBNs defined in this work consist of miRNAs, target mRNAs

Figure 6 48 genes and 5 miRNAs in module 3; each node denotes a gene (rectangle) and miRNA (octagon) in module and edges depict miRNA–mRNA relationships.
and associated conditions, abnormal and normal in the demonstrated case. By associating the miRNA–mRNA pairs with environments, the method may identify the regulatory networks of ‘miRNA–mRNA bipartite’. It will bring a new chance to the treatment of genetic traits such as folliculogenesis process. On the other hand, one of the most significant roles of the ovary is to undergo a continuous folliculogenesis to produce competent and viable oocytes. Understanding the molecular mechanism of follicular development is necessary to unravel the complex synergies orchestrated during the process of forming the fertilizable ovum. The granulosa cells are required during the stages of oocyte growth and development including ovulation and fertilization.

Although several authors (Bao & Garverick et al. 1998, Vitt et al. 2000, Hayashi et al. 2010) have described the presence of altered gene expression patterns in granulosa cells of different follicular stages, the post-transcription regulation mechanisms that lead the follicle to become dominant or subordinate are still unclear. Here, we investigated the accumulation or degradation of miRNA–mRNA bipartite networks in granulosa cells of subordinate and dominant follicles during folliculogenesis. For this, we have integrated miRNA and mRNA data from granulosa cell samples derived from subordinated and dominant follicles. As the integrating data were generated from numerous biological replicates of each sample, we believe that the analysis has been robust and the findings could be repeatable. Therefore, from our data, we detected several miRNAs and mRNAs in granulosa cells of dominant and subordinate follicles.

Indeed, the role of these miRNAs and genes during folliculogenesis is not clearly known, but some miRNA family members are believed to be involved in cell proliferation, survival, apoptosis and the other processes by targeting multiple genes. Therefore, the expression of these miRNAs and mRNAs or genes both in SF and DF of the estrous cycle could indicate their basal function/housekeeping cellular roles (Turcatel et al. 2012) during follicular recruitment, selection, dominance or follicular atresia. Herein, we utilized an integrative systems biology approach to ascertain deregulated signaling pathways associated with deregulated miRNA–mRNA interactions in folliculogenesis. Our results have detected 6 key modules in folliculogenesis, which are accessible in detail in Supplementary Table 3. In these modules, we have identified 151 genes and 27 miRNAs, which module 1, 2, 3, 4, 5 and 6 were included 50, 79, 48, 13, 44 and 26 genes, as well as 2, 10, 5, 1, 6 and 3 miRNAs. The merger of the two modules with the most identical genes (named ‘Merged-module’ in this study) is depicted in Fig. 5 and Fig. 6. Signaling pathway and GO enrichment analysis of ‘miRNA-deregulated mRNAs’ delineated important signaling pathways, which are dysregulated by deregulated miRNA and mRNAs. In the module 1, hsa-miR-637 and hsa-miR-498 suppressed COL4A1 gene and both of miRNAs were downregulated and the COL4A1 gene was upregulated. COL4A1 gene encodes a type IV collagen alpha protein. It functions as section of a heterotrimer and interacts with other extracellular matrix components such as proteoglycans, perlecans and laminins (Zenteno et al. 2014). As well as MCL1 gene that relates to cell survival and cell adhesion was upregulated in the dominant follicles because of these follicles need to survival for the ovulating process. On the other hand, this gene encodes an anti-apoptotic protein, which is a member of the Bcl-2 family, so involved in the regulation of apoptosis versus cell survival, and in the maintenance of viability but not of proliferation. Mediates its effects by interactions with a number of other regulators of apoptosis. In the module 2, hsa-miR-216b-5p, hsa-miR-216a-5p, hsa-miR-887-5p, hsa-miR-3689a-3p and hsa-miR-3689b-3psuppressed C3 gene. Complement component C3 plays a key role in the activation of complement system. The encoded pro-protein is proteolytically processed to generate alpha and beta subunits that formation of the mature protein (Lusco et al. 2015), which is then further processed to generate numerous peptide products such as proteins that are relate to folliculogenesis process. Also, hsa-miR-3689a-3p and hsa-miR-887-5p which were upregulated suppressed ITGB3 gene which was downregulated. The ITGB3gene product is the integrin beta chain beta 3. Integrin beta 3 is found along with the alpha IIb chain in platelets. Integrins are known to take part in cell-surface mediated signaling (Bianconi et al. 2015) that in the subordinate follicles, follicles need to cell-surface mediated signaling for cell growth and proliferation so,
this gene was upregulated in the subordinate follicles. 

**PIK3R1** gene was suppressed by hsa-miR-30d-3p in this module. Phosphatidylinositol 3-kinase plays an important key role in the PI3K/Akt signaling pathway (Lin et al. 2015). hsa-miR-30d-3p, hsa-miR-451a and hsa-miR-34b-5p were upregulated suppressed **MYC** gene which was downregulated. The protein encoded by this gene is a nuclear phosphoprotein that plays a role in cell cycle progression, cellular transformation, apoptosis, cell growth and proliferation (Camarda et al. 2016). Because of subordinate follicles critical need to cell growth and proliferation thus **MYC** gene was upregulated in these follicles. In the module 3, has-miR-3666, has-miR548v, bta-miR-144 and bta-miR-222 suppressed **KIT** gene. All of miRNAs were upregulated and **KIT** gene was downregulated. **KIT** gene encodes the homolog of the proto-oncogene c-kit. C-kit was first detected as the cellular homolog of the feline sarcoma viral oncogene v-kit. This protein is a type 3 transmembrane receptor for mast cell growth factor, also known as stem cell factor (Garcia-Montero et al. 2016). As well as **ANGPT2** gene was downregulated. Besides the protein encoded by this gene is an antagonist of angiopoietin 1 (ANGPT1) and endothelial TEK tyrosine kinase. The encoded protein disrupts the vascular remodeling ability of ANGPT1 and may induce endothelial cell apoptosis, therefore this is natural that these genes were downregulated in the dominant follicles. In the module 4, bta-miR-216a suppressed **NR5A2**. The protein encoded by **NR5A2** gene is a DNA-binding zinc finger transcription factor and is a member of the fushi tarazu factor1 sub-family of orphan nuclear receptors. The encoded protein is involved in the expression of genes for cholesterol biosynthesis and may be an important regulator of embryonic development (Montanaro et al. 2015). Also in this module **CYP19A1** gene was upregulated in the dominant follicles because of this gene encodes a member of the cytochrome P450 superfamily of enzymes. This protein localizes to the
endoplasmic reticulum and catalyzes the last steps of estrogen biosynthesis. In the module 5, hsa-miR-3689a-3p, hsa-miR-3689b-3p and has-miR-887-5p suppressed ITGB3 gene, bta-miR-22-3p suppressed CSF1R genes which bta-miR-22-3p was downregulated and CSF1R was upregulated. The protein encoded by CSF1R gene is the receptor for colony stimulating factor 1, a cytokine which controls the differentiation, production and function of macrophages. Ligand binding activates the receptor kinase through a process of transphosphorylation (Cassier et al. 2015). Also, has-miR-4279 suppressed TLR4 genes which encoded a protein which is a member of the Toll-like receptor (TLR) family which plays a fundamental role in activation of

*Figure 9* 26 genes and 3 miRNAs in module 6; each node denotes a gene (rectangle) and miRNA (octagon) in module and edges depict miRNA–mRNA relationships.

**Table 1** Basic network parameters of the two generated networks, compared with simulated randomized model networks.

<table>
<thead>
<tr>
<th></th>
<th>Integrated miRNA–mRNA Bipartite Network</th>
<th>Simulated Barabasi–Albert model, (scale free)</th>
<th>Simulated Erdos–Renyi model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>531</td>
<td>531</td>
<td>531</td>
</tr>
<tr>
<td>Clustering coefficient</td>
<td>0.108</td>
<td>0.006</td>
<td>0.004</td>
</tr>
<tr>
<td>Characteristic path length</td>
<td>4.192</td>
<td>4.947</td>
<td>4.432</td>
</tr>
<tr>
<td>Network density</td>
<td>0.006</td>
<td>0.007</td>
<td>0.007</td>
</tr>
</tbody>
</table>
innate immunity. They recognize pathogen-associated molecular patterns and mediate the production of cytokines necessary for the development of effective immunity (Bagheri et al. 2016). Both of these genes were upregulated in dominant follicles because of mentioned reasons. In the module 6, hsa-miR-215-5P, bta-miR-20a and bta-miR-216b suppressed LRP8 gene which encode a member of the low-density lipoprotein receptor family. Low-density lipoprotein receptors are cell-surface proteins which play roles in signal transduction. The encoded protein plays a critical role in functions as a receptor for the cholesterol transport protein apolipoprotein E (Shen et al. 2013). On the other hand, Elf4ebp1 gene which relates to PI3K/akt/mTORc1 signaling pathway and protein synthesis was upregulated in the dominant follicles. This gene encodes one member of a family of translation repressor proteins. The protein directly interacts with eukaryotic translation initiation factor 4E (Elf4E), which is a limiting component of the multi-subunit complex that recruits 40S ribosomal subunits to the 5' end of mRNAs. Interaction of this protein with Elf4E inhibits complex assembly and represses translation. Since dominant follicles don’t need to additional growth, protein synthesis via over-expression of this gene was inhibited.

We suggest a computational method for Integrated miRNA–mRNA Bipartite Network using predicted and validated miRNA targets as well as expression profiles of miRNAs and mRNAs. It will give new insight into the biological procedures at the molecular level. The spatio-temporal expression of miRNAs and genes in granulosa cells during folliculogenesis supports the potential role of miRNAs in post-transcriptional regulation of genes involved in follicular development, mainly ovulation of a preovulatory dominant follicle and regression of an ovulatory subordinate follicles. The information we provided here for first time may be helpful in deciphering the molecular mechanism of follicular ovulation and atresia. Further in vitro and in vivo experiment is required to fully understand the specific functional role of classes or module of miRNAs and mRNAs during various stages of the follicular development and the follicular phase of the estrous cycle in particular.

Supplementary data
This is linked to the online version of the paper at http://dx.doi.org/10.1530/REP-17-0049.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This work was financially supported by the University of Tehran, Iran.
References


Reproduction (2017) 154 51–65

www.reproduction-online.org
Bovine folliculogenesis interactome


Received 25 January 2017
First decision 16 March 2017
Revised manuscript received 6 April 2017
Accepted 24 April 2017

www.reproduction-online.org