

# Extensive ceRNA–ceRNA interaction networks mediated by miRNAs regulate development in multiple rhesus tissues

Juan Xu<sup>1,†</sup>, Lin Feng<sup>2,†</sup>, Zujing Han<sup>3,†</sup>, Yongsheng Li<sup>1,†</sup>, Aiwei Wu<sup>1</sup>, Tingting Shao<sup>1</sup>, Na Ding<sup>1</sup>, Lili Li<sup>1</sup>, Wei Deng<sup>4</sup>, Xuebing Di<sup>2</sup>, Jian Wang<sup>3</sup>, Lianfeng Zhang<sup>4,\*</sup>, Xia Li<sup>1,\*</sup>, Kaitai Zhang<sup>2,\*</sup> and Shujun Cheng<sup>1,2,\*</sup>

<sup>1</sup>College of Bioinformatics Science and Technology, Harbin Medical University, 194 Xuefu Road, Harbin 150081, China, <sup>2</sup>State Key Laboratory of Molecular Oncology, Department of Aetiology and Carcinogenesis, Cancer Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100021, China, <sup>3</sup>BGI Tech Solutions Co., Ltd., Beishan Industrial Zone, Yantian District, Shenzhen 518083, China and <sup>4</sup>Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 10021, China

Received October 29, 2015; Revised June 17, 2016; Accepted June 19, 2016

## ABSTRACT

Crosstalk between RNAs mediated by shared microRNAs (miRNAs) represents a novel layer of gene regulation, which plays important roles in development. In this study, we analyzed time series expression data for coding genes and long non-coding RNAs (lncRNAs) to identify thousands of interactions among competitive endogenous RNAs (ceRNAs) in four rhesus tissues. The ceRNAs exhibited dynamic expression and regulatory patterns during each tissue development process, which suggests that ceRNAs might work synergistically during different developmental stages or tissues to control specific functions. In addition, lncRNAs exhibit higher specificity as ceRNAs than coding-genes and their functions were predicted based on their competitive coding-gene partners to discover their important developmental roles. In addition to the specificity of tissue development, functional analyses demonstrated that the combined effects of multiple ceRNAs can have major impacts on general developmental and metabolic processes in multiple tissues, especially transcription-related functions where competitive interactions. Moreover, ceRNA interactions could sequentially and/or synergistically mediate the crosstalk among different signaling pathways during brain development. Analyzing ceRNA interac-

tions during the development of multiple tissues will provide insights in the regulation of normal development and the dysregulation of key mechanisms during pathogenesis.

## INTRODUCTION

Fetal development processes in the uterus or even after birth are among the most critical periods in life, and organs with distinct structures, sizes and diverse functions are formed gradually during these periods (1–3). Elucidating the characteristic molecular events that occur during the development of various organs over the entire period of development can help to understand molecular development as well as improving clinical practice.

In mammals, tissue development is a tightly regulated process, which requires exquisite control over gene expression (4–6). Thus, identifying the regulatory principles that govern tissue development is of great interest to developmental and molecular biologists, as well as clinicians. In particular, microRNAs (miRNAs), a class of small non-coding RNAs, have been studied widely in organ development, inflammation, and tumorigenesis because of their inhibitory effects on targets (7–9). In addition, competing endogenous RNAs (ceRNAs) have emerged recently to add to the complexity of miRNA-mediated gene regulation (10–12). ceRNAs are RNAs that share miRNA recognition elements (MREs) thereby competing for miRNA binding and regulating each other. Numerous recent studies of mammalian cells and tissues have demonstrated the competition

\*To whom correspondence should be addressed. Tel: +86 451 86615922; Fax: +86 451 86615922; Email: lixia@hrbmu.edu.cn

Correspondence may also be addressed to Lianfeng Zhang. Email: zhanglf@cnilas.org

Correspondence may also be addressed to Kaitai Zhang. Email: zhangkt@cicams.ac.cn

Correspondence may also be addressed to Shujun Cheng. Email: chengshj@263.net.cn

<sup>†</sup>These authors contributed to the paper as first authors.

between the protein-coding targets of miRNAs, which supports a regulatory role for mRNAs independent of their protein coding function (13,14). Large-scale analyses of human gliomas have led to the conclusion that thousands of transcripts may act as target decoys, and the combined effects of multiple ceRNAs can have a major impact on gene expression and cellular phenotypes (15,16). However, most ceRNA interactions between mRNAs are linked to various disease states, and few have been linked to development.

In addition, some long non-coding RNAs (lncRNAs), which are recently discovered types of non-coding RNAs, play important roles in the regulation of gene expression by acting as ceRNAs, thereby indicating an additional layer of complexity (17–19). The important roles played by lncRNAs as ceRNAs in development processes are also becoming increasingly evident. C Marcella *et al.* identified a muscle-specific lncRNA, linc-MD1, which governs the time of muscle differentiation by acting as a ceRNA in mouse and human myoblasts (17). Recently, lncRNAs have also been shown to function as an miRNA sponge during human embryonic stem cell self-renewal. Wang *et al.* showed that linc-RoR competes with the mRNAs of core transcription factors to maintain embryonic stem cell pluripotency (19). Interestingly, linc-RoR and its sponge effect can disappear after the embryonic stem cells start to differentiate, thereby indicating dynamic patterns of ceRNA interactions. However, the roles of lncRNA associated-ceRNAs in tissue development are not fully understood, as well as the dynamic patterns of ceRNA interactions during development processes.

According to their competition for targets, the most typical feature of ceRNA pairs is the sharing of MREs. Consequently, ceRNAs exhibit indirect and positively correlated expression mediated by miRNAs, and several powerful tools could be used to quantify dependencies among them, such as correlation coefficient, partial correlation coefficient, mutual information or conditional mutual information (20). In recent years, several computational algorithms have been developed, especially for disease conditions (15,16,21,22), where Ala *et al.* bioinformatically predicted the ceRNA network based on MRE overlap and gene expression correlation (21). In addition, Sumazin *et al.* used mutual and conditional mutual information to describe an extensive and highly interconnected ceRNA network in glioblastoma (15), where there were strong positive correlations among the expression levels of the individual ceRNA components. However, the performance of this method depends heavily on obtaining a large expression profiles of mRNA and miRNA from the same samples, which is relatively difficult to achieve in dynamic studies of development. In addition, Tay *et al.* mainly considered the high overlap of MREs and the similar distribution of MREs in the context of target mRNA sequences, and used this method to predict the ceRNA network for a given mRNA PTEN in human cancers (16). The top ranked ceRNAs for PTEN were also validated experimentally. Despite the growing appreciation of the importance of miRNA mediated RNA crosstalk in development processes, our knowledge of the genome-wide ceRNA interaction networks (ceRNETs) during tissue development is still limited and few modulators of miRNA activity have been characterized.

In this study, to better understand the dynamic processes that occur during tissue development, time-series experiments were performed using four different rhesus tissues with RNA-sequencing technology. A computational algorithm was developed to generate genome-wide ceRNETs among the whole transcriptome, including protein coding genes and lncRNAs. Our algorithm was designed to consider three general characters: shared miRNAs, the similarity of miRNA regulation strength, and the co-expression of transcripts. We also performed a series of computational ceRNA network analyses. In particular, we evaluated the contributions of ceRNA subnetworks in each tissue and their associations among the four tissues during development. Understanding miRNA mediated RNA crosstalks can provide significant insights into gene regulatory networks and their implications for development and disease.

## MATERIALS AND METHODS

### Organ collection and sequencing

All of the rhesus monkeys were raised at the Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences, and Peking Union Medical College. Three different development stages were considered, i.e. fetuses at fetal early development and late development stages, and postnatal infants after several days. We performed caesarean sections to obtain fetuses from pregnant monkeys at given time points. For younger fetuses, lung, liver, brain and colonic mucosal samples were collected from each individual rhesus, brain samples were obtained from the cortex of the right frontal lobe.

The libraries were sequenced with the Illumina HiSeq 2000 platform using the 90-bp pair-end sequencing strategy. The original image data generated by the sequencing machine were converted into sequence data by base calling (Illumina pipeline CASAVA v1.8.2) and then subjected to standard QC criteria to remove all of the reads that satisfied any of the following parameters:

- (1) Reads that aligned to adaptors or primers with no more than two mismatches.
- (2) Reads with >10% unknown bases ( $N$  bases).
- (3) Reads with >50% low quality bases (quality value  $\leq 5$ ) in one read.

### Read mapping and gene expression analysis

The reference genome for the rhesus macaque monkey (version rheMac3) was downloaded from the UCSC Genome Browser (<http://genome.ucsc.edu>). Clean paired-end reads were aligned to the reference genome using TopHat (23). The transcriptome of each sample was constructed using Cufflinks (24). Previously unknown lncRNAs expressed in the rhesus macaque monkey were also identified. Transcripts >200nt were identified as lncRNAs if they did not overlap with known genomic annotations from the Ensembl Database and no coding potential was assessed by PhyloCSF. Next, the reads per kilo base of model per million base pairs sequenced (RPKM) was used to quantify the gene or lncRNA expression levels (25).

### Collection of development-related genes and miRNAs

Embryo development genes were also obtained from the Gene Ontology (GO) database, according to the extracted processes proposed by Gong *et al.* (26). Similarly, brain development-related genes were also obtained from GO, which were annotated with the term ‘brain development.’ In addition, brain-related miRNAs were collected from several recent reviews.

### miRNA target prediction

Known mature miRNA sequences from the rhesus macaque monkey were downloaded from miRBase. We also extracted genome-wide multiple alignments of six primates from the RhesusBase Database (27), which were built from the following genome assemblies: Human, Chimpanzee, Gorilla, Orangutan, Rhesus and Marmoset. The UCSC mappings of the rhesus RefSeq mRNA data (release 6, 5 July 2004) were used as the rhesus genome to define multiple alignments of 3' untranslated regions (UTRs). Multiple alignments of lncRNAs were also obtained based on the genomic coordinates of the lncRNA transcripts. If multiple annotated 3'UTR/lncRNA sequences were available for a coding/lncRNA gene, the longest was used in the analyses. Due to its relatively high precision and previous successful application to the rhesus monkey (28–30), TargetScan 5.0 (31) was used with the default parameters to identify conserved miRNA target sites in the 3' UTR of coding transcripts and full-length lncRNA transcripts. Only MREs (MREs) with context scores <0 were retained.

### Constructing ceRNETS related to the development of rhesus monkey tissues

*Overview of the processes used to identify ceRNA interaction pairs.* Supplementary Figure S1 illustrates the method used in this study. After obtaining the miRNA-mRNA and miRNA-lncRNA regulatory data, we employed the following principles to generate ceRNA pairs that occurred during development in each rhesus monkey tissue. A central tenet of the ceRNA characteristics is that *trans*-regulatory ceRNA crosstalk is enhanced by the high regulatory similarity of miRNAs between transcripts and strong co-expression during rhesus monkey tissue development. First, for each RNA-RNA pair (including mRNA-mRNA pairs, lncRNA-lncRNA pairs and lncRNA-mRNA pairs), the significance of shared miRNAs was computed. If two RNA components were significantly regulated by common miRNAs, we also calculated the similarity of the miRNA regulatory pattern between them according to a set of rules. The RNA pairs that passed through these two filters were considered candidate reciprocal ceRNA interaction pairs. Moreover, to identify ceRNA-ceRNA pairs implicated in rhesus tissue development, we examined whether putative ceRNA pairs were positively coexpressed in the time series expression data for each tissue. Finally, the ceRNETS for each tissue were constructed by assembling all of the ceRNA pairs, where the nodes represent mRNAs or lncRNAs and the edges represented their competing interactions.

*Evaluating the significance of sharing miRNAs.* Due to the computationally prohibitive burden of testing all possible combination of RNA/miRNAs/RNA triplets, we first required that RNA-RNA pairs shared a statistically significant number of common miRNAs. For a given RNA pair of RNA *X* and *Y*, we identified the miRNAs that regulated them, and a hypergeometric test was used to measure whether these two RNA components significantly shared miRNAs. The probability *P* was calculated according to:

$$P = 1 - F(N_{XY} - 1 | N, N_X, N_Y) = 1 - \sum_{t=0}^{N_{XY}-1} \frac{\binom{N_X}{t} \binom{N - N_X}{N_Y - t}}{\binom{N}{N_Y}}$$

where *N* is the number of all miRNAs (default background distribution), *N<sub>X</sub>* and *N<sub>Y</sub>* represent the total number of miRNAs that regulate RNA *X* and *Y*, respectively, and *N<sub>XY</sub>* is the number of miRNAs shared between RNA *X* and *Y*, which had to be at least three. We controlled for multiple hypotheses using the false discovery rate (FDR), and only pairs passing an FDR of 0.1 were considered to be significantly co-regulated. Thus, there were 714 061 RNA pairs, which accounted for 8.5% of the RNA pairs sharing at least three miRNAs.

*Screening candidate ceRNA pairs with high similarity of miRNA regulation pattern.* In addition to significantly sharing miRNAs, the specific details of MREs may influence the ceRNA crosstalk, such as the MRE densities of the shared miRNAs and their distribution throughout the target RNA (10,11,16,21). Thus, we also evaluated the trans-regulatory efficiency of two RNAs based on their MRE language according to four rules. The power of these indexes were demonstrated by the validated ceRNA interactions for PTEN in humans (details in Text S1). If the similarity of miRNA regulation is higher, then two RNAs are more likely to compete with each other. Thus, the significant candidate RNA pairs identified as described above were assessed further. For a given RNA pair of RNA *X* and *Y*, the following four measures were calculated to determine whether the RNA *Y* is a modulator of RNA *X*.

- (1) Accumulating evidences suggest that ceRNA cross-regulation increases with the number of shared miRNAs (21), which are often considered to identify ceRNA interactions for a specific RNA in biological experiments (14,16). Thus, when RNA *Y* is regulated by more shared miRNAs with RNA *X*, then it may be a more effective ceRNA. The proportion of miRNAs that RNA *Y* shares with RNA *X* was calculated as follows.

$$R_1 = \frac{N_{XY}}{N_X}$$

- (2) miRNAs are predicted to target tens to hundreds of RNAs, but they do not exert the same degree of repression on all of them, where this is an underlying association with the number of MREs (10). The number of MREs has an independent and significant effects on the targeting efficiency and increasing the number of MREs for the same or different miRNAs improve the binding

strength (31). Moreover, ceRNA crosstalk depends on the MREs located on each transcript, which combinatorially comprise the foundation of these co-regulatory interactions (11). This measurement represents the relationship between the total number of MREs predicted in *Y* compared with the total number of miRNAs that give rise to these MREs, thereby favoring cases where each targeting miRNA give rise to more than one MRE in *Y*:

$$R_2 = \frac{NMRE_{XY} - N_{XY} + 1}{NMRE_{XY}}$$

where  $NMRE_{XY}$  is the number of MREs in *Y* for miRNAs that regulate both RNA *X* and *Y*.

- (3) The effectiveness of a ceRNA also depends on the global sponge activity for shared miRNAs (10), which can be measured as the density of the predicted MREs for the shared miRNAs. This index represents the normalized regulatory strength in *Y* for miRNAs that regulate *X*, which is similar to the expression level of a transcript quantified by RPKM, and it favors cases where more MREs are located within shorter distances:

$$R_3 = \sum_1^{N_{XY}} \frac{NMRE_{XY_i} * 1000}{D_i * NMRE_i}$$

where  $NMRE_{XY_i}$  represents the number of MREs in *Y* for one shared miRNA *i*.  $D_i$  is the distance between the leftmost and rightmost MREs in *Y*, and  $NMRE_i$  is the total number of MREs in all of the considered genes for miRNA *i*.

- (4) In addition to the number of MREs, the MRE distribution throughout the target RNA influences the miRNA targeting efficacy. The effect can be antagonism when two target sites are nearly overlapping (16,32). Thus, an even distribution of MREs might favor the capacity of individual miRNAs to function collectively in synergistic manner. This index measures the distribution of the MREs over the distance that they span in *Y* in order to evaluate the extent of the evenness of the distribution:

$$R_4 = \frac{\sum_1^{N_{XY}} D_i^2}{\sum_1^{N_{XY}} \sum_1^{NMRE_{XY_i}-1} D_{ij}^2}$$

where  $D_{ij}$  represents the distance between any two successive MREs in *Y* for miRNA *i*, which favors situations where the MREs tend to be more evenly distributed.

To further determine whether the miRNA regulation pattern of RNA *Y* was significantly similar to that of RNA *X*, randomization tests were performed by randomly selecting RNA pairs similar to a pair of *X* and *Y* in term of the 3'UTR length and GC content. We recalculated these four indexes for each random RNA pair. If the pair had no common miRNAs, the values of the four indexes were set to 0. This procedure was repeated 1000 times. The significant *P*-value for each index was given as the frequency of the index values in random conditions, which was greater than the value in the actual condition. In this study, we aimed to identify reciprocal ceRNA interactions, so we also determined whether RNA *X* was also a modulator of RNA *Y* based on the four indexes and the significance of each index was evaluated as described above. Finally, the global statistical sig-

nificance of the reciprocal interaction was determined using Fisher's method, which combines the *P*-values for the two competing directions. Benjamini-Hochberg multiple testing correction was used to estimate the FDR, where only RNA pairs that passed an FDR of 0.1 were considered to be candidate reciprocal ceRNA interactions.

*Discovering ceRNA interactions that underlie tissue development in rhesus monkeys.* To characterize the functions of specific ceRNAs and their competing relationships in the context of tissue development in rhesus monkeys, we examined whether putative ceRNA pairs were positively co-expressed during tissue development based on time-series of gene expression data. RNA *Y* was considered to be co-expressed with RNA *X* if *Y* was within the top five percentile for all of the correlation values calculated with respect to *X*, and vice versa. Thus, candidate ceRNA pairs with strong positive co-expression in both directions were implicated in tissue development.

Finally, for each tissue, after assembling all of the ceRNA-ceRNA pairs, we generated the ceRNET related to the development of this tissue. A node represented an RNA (coding RNA or lncRNA), and two nodes were connected by an edge if the corresponding RNA pair had a reciprocal competing interaction; otherwise, no edge was added.

#### Identification of ceRNAs with developmental stage-specific expression pattern

General changes in trends of ceRNA expression during tissue development were analyzed using the k-means clustering method, where a Silhouette function was employed to optimize the selection of *K* to minimize the distances among data within clusters while maximizing the distance between clusters (details in Text S2) (33,34). For colon and liver tissues, *K* = 3 obtained a maximal silhouette score according to both statistics, so this was used. In addition, *K* = 3 or *K* = 4 obtained similar high silhouette scores, and after observing the expression features of ceRNAs based on heatmap figures, *K* = 4 was considered to be more suitable for brain and lung tissues. In each cluster, ceRNAs exhibited similar expression levels, which might play important roles in specific developmental stages.

#### Interaction preferences

To evaluate how closely different clusters of ceRNAs or ceRNAs within the same cluster interacted competitively with each other, we employed the concept of an interaction preference index (35). An interaction preference index quantitatively assesses the extent of interactions by ceRNAs between two clusters or within the same clusters in the actual ceRNET of each tissue compared with random cases. Random cases represented the average number of interactions by the considered ceRNAs in 10 000 degree-preserving random networks. When the interaction preference index was significantly >1, this indicated that the interactions were enhanced. By contrast, interactions were considered to be suppressed if this measure was significantly <1.

### Enriched functional categories for ceRNA clusters

A cumulative hypergeometric distribution test was used to identify the significantly overrepresented biological function categories for specific ceRNA clusters, where the functional annotations of genes were obtained from the GO and KEGG databases. Specifying genes expressed in the relevant tissue were defined as background genes. Functional categories with adjusted *P*-values <0.05 and annotated by at least three genes were considered in our analyses.

## RESULTS

### Complex competitive interactions among coding-RNAs and lncRNAs during tissue development

To systematically detect the expression of both lncRNAs and coding genes during tissue development in rhesus monkeys, we first analyzed time-series expression data by RNA-seq for four tissues, i.e. brain, lung, colon and liver. The samples ranged from the early fetal to postnatal stages. In total, 2613 lncRNAs were detected in four tissue types. We used the TargetScan algorithm to obtain genome-wide prediction of miRNA regulation throughout the rhesus monkey transcriptome, including protein-coding RNAs and lncRNAs (see the Materials and Methods). Using *in silico* approaches, it has been estimated that miRNAs regulate >60% of human protein-coding genes, and we predicted that approximately 46.86% of the coding genes in the rhesus monkey were regulated by miRNAs. In addition, 76.16% of the lncRNAs also interacted with miRNAs, thereby indicating the more extensive regulation of miRNAs to lncRNAs in the rhesus monkey. Without considering the conservation of miRNA regulations, we also found that 90.91% of lncRNAs were regulated by miRNAs, which is also higher than that for mRNAs. In addition, the regulatory density was calculated as the number of miRNAs per 1000 bp, which showed that the regulatory density was lower in lncRNAs than mRNAs (Supplementary Figure S2A and B). Therefore, these results indicate that there is a higher fraction but lower regulatory density in lncRNAs than mRNAs. Moreover, all of the lncRNAs shared at least one miRNA with mRNAs, where the proportion reached 77.06% for those sharing more than five miRNAs, thereby supporting the ‘sponge’/‘decoy’ function roles for lncRNAs and reducing the regulatory effect of miRNAs on the targeted mRNAs. These observations might suggest that miRNAs and their targets are connected in complex ceRNAs.

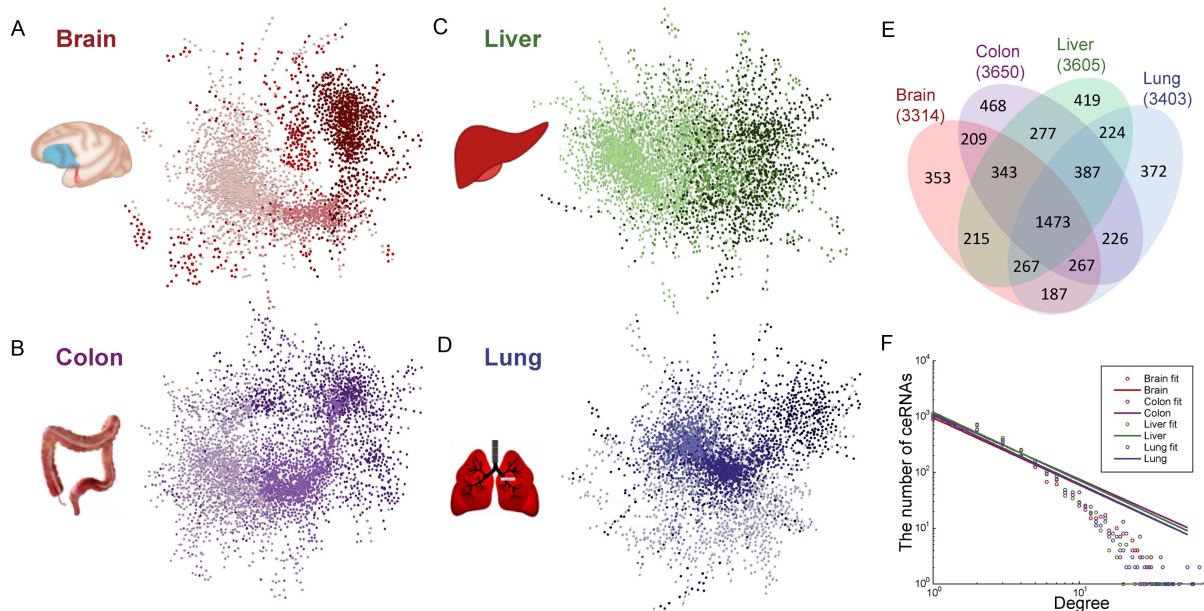
To systemically explore target competition during organ development, we employed a computational method to identify genome-wide ceRNA–ceRNA interactions involved in each tissue development process, where we considered both the similarity of the miRNA regulatory patterns and the similarity in expression during tissue development (see the Methods). Based on ceRNA interaction pairs, four ceRNAs were constructed (Figure 1A–D). Moreover, few ceRNA pairs were co-regulated by transcription factors, which indicates that the correlated expression of ceRNA pairs might not be due to shared regulation at the transcriptional level, but instead it may be attributable mainly to shared regulation at the post-transcriptional level (details in Text S3). Three types of ceRNA interactions were

present in each network: mRNA–mRNA, lncRNA–mRNA and lncRNA–lncRNA interactions. In general, the sizes of the four ceRNAs were similar (Figure 1E and Supplementary Table S1), where the number of ceRNA interactions accounted for only 0.015% of all the candidate RNA combinations, which involved ~19.14% of the lncRNAs or mRNAs. These results suggest that a specific proportion of mRNAs and lncRNAs may act as ceRNAs during rhesus tissue development. Moreover, over half of the ceRNAs were present in at least two ceRNAs and they could directly modulate the expression of other transcripts (Figure 1E). In particular, 1473 ceRNAs were shared by four tissues, which suggests that some of the coding transcripts and lncRNAs tend to act as common ceRNAs that regulate developmental processes in four tissues.

As shown in Figure 1A–D, most of the ceRNAs in each ceRNA were interconnected and they could engage in crosstalk directly with each other by sharing numerous MREs. In addition, all of the ceRNA interactions in each tissue formed a large connecting subnetwork, thereby indicating that the ceRNAs can also communicate indirectly via their shared ceRNA partners. Therefore, the ceRNAs had complex combinations in terms of competition, and thus alterations in one ceRNA might have major effects on the entire ceRNA. Next, we examined the degree distributions of these ceRNAs, which detected power law distributions (Figure 1F). Therefore, like many large-scale biological networks, the ceRNAs had scale-free characteristics, which indicates that the ceRNAs were not random but instead they were characterized by a core set of organizing structural principles that distinguished them from randomly linked networks (36). Understanding tissue development in the rhesus monkey in the context of these network principles allows us to investigate some fundamental properties of mRNAs and lncRNAs.

### ceRNAs exhibit dynamic expression and interaction patterns during tissue development

Next, we investigated how the global expression patterns of protein coding transcripts and lncRNAs in the ceRNAs were coordinated during the development processes in the four tissue types. We found that most of the ceRNAs in each ceRNA (95.23–99.35%) were expressed during the time course of the corresponding tissue ( $PRKM > 1$ ). To further extract additional biological information from the multi-dimensional transcriptome data set, we performed clustering analysis using the *k*-means method to identify co-expressed ceRNA clusters (Figure 2). We found that most of the ceRNAs exhibited dynamic and stage-specific expression, thereby demonstrating the temporally related expression of both mRNAs and lncRNAs. The ceRNAs in the ceRNAs constructed for brain and lung tissues were globally grouped into four clusters, and those in other two ceRNAs were grouped into three main clusters. These ceRNA clusters corresponded to the highest expression levels during early prenatal development, medium-term fetal development, and the postnatal infancy period. Interestingly, most of the ceRNAs in each tissue exhibited their highest expression levels during the fetal development period. This trend has also been detected during the development

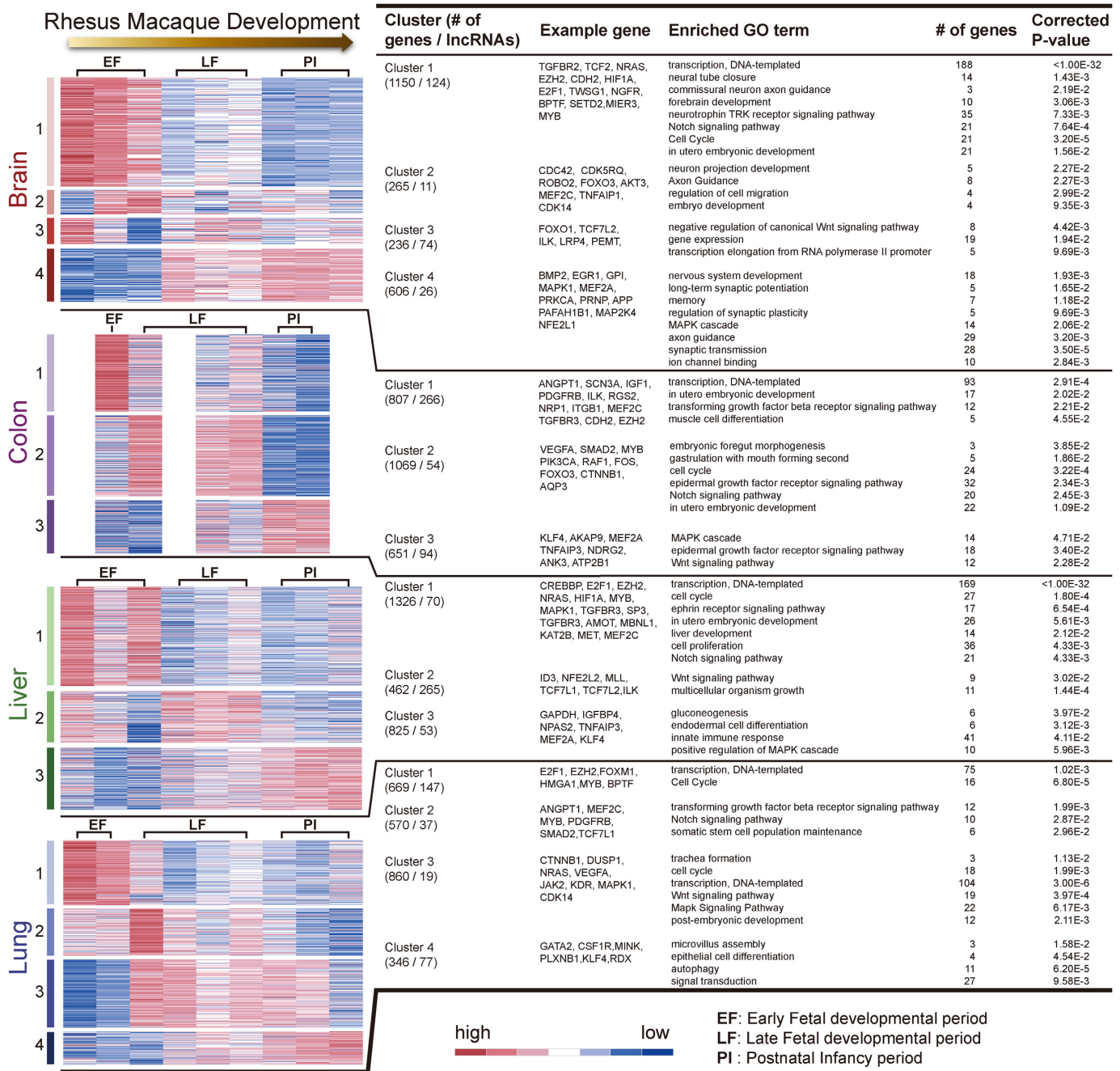


**Figure 1.** The four ceRNETs obtained for different tissues. (A–D) The four ceRNETs obtained for brain, liver, colon, and lung tissues. The color gradient represents different clusters. (E) Venn diagram showing the overlapping relationships of the ceRNAs in the four ceRNETs. (F) The degree distributions of the four ceRNETs.

of human brain tissue, where the highest changes in gene expression occur during fetal stage, before declining into infancy (3). Our results confirmed the expression of key coding genes during brain development, such as HIF1A, E2F1 and TWG1, as well as known colon-related coding genes (including MYB and RNF43). We also identified several other stage-specific ceRNAs, which represent potentially new regulators of tissue development.

We considered that ceRNAs with similar temporal expression patterns would share a common miRNA regulatory pattern and exhibit close competing interactions. Thus, we defined an interaction preference index to detect significantly over- or under-represented interaction patterns (see the Materials and Methods). We found that ceRNAs within the same cluster had a significantly high density of competing interactions (Table 1,  $P < 0.0001$ ), thereby implying that intra-cluster crosstalk was enhanced and that ceRNAs with major roles during the same period were more likely to competitively regulate each other. However, the interaction density between different ceRNA clusters appeared to be extremely low and all were significantly under-represented (Table 1,  $P < 0.0001$ ), which suggests that the crosstalk among different ceRNA clusters was repressed. Moreover, we found that the crosstalk between clusters not related to adjacent development stages was repressed more strictly than that between clusters with adjacent development stages. For example, in the four tissues, the interaction preference indexes were lowest between clusters during early prenatal development and clusters during the postnatal infancy period, compared with the scores for clusters in adjacent developmental stages. These results suggest that ceRNAs involved in the same development stage tend to engage in crosstalk with each other, whereas those operating in different stages tend to avoid interactions.

Moreover, we performed GO and pathway analyses for each co-expressed ceRNA clusters, which showed that the stage-specific clusters were enriched in terms of the expected functional gene sets (Figure 2 and Supplementary Table S2). Various general functions related to development and metabolism were also identified in all four tissues, such as transcription, development, cell cycle, notch signaling pathway and MAPK cascade. Similarly, within the same tissue, we also discovered several consistent functions that operated during tissue development in both the fetal and postnatal-infant periods. For example, during brain tissue development from the fetal to infant periods, ceRNAs participated in many more specific functions that contributed to continuous brain development as well as the common functions described above, such as signaling transduction, cell cycle, axon guidance, autophagy, and embryonic development. In addition, expected GO terms for different brain development stages were also enriched for these stage-specific clusters. The largest cluster for brain tissue, Cluster 1, exhibited a common developmental trend with the highest expression levels during the early stage of prenatal development (Figure 2), followed by a progressive decline in the expression levels with age until postnatal infancy. This cluster also comprised a large number of known early neurogenic factors as prominent members. Enrichment analysis showed that these genes were enriched for functional categories related to neural differentiation, such as neuronal stem cell population maintenance ( $\text{adj}P = 0.046$ ) and neural tube closure ( $\text{adj}P = 1.43\text{e}-3$ ). By contrast, another cluster exhibited changes in the opposite direction (relative to the former cluster), with the highest expression levels during the postnatal infancy period. The functional categories in this cluster included many functions related to ion transport and homeostasis, as well as calmodulin binding ( $\text{adj}P = 9.86\text{e}-4$ ). Moreover, the ceRNAs in this cluster might



**Figure 2.** Development-dependent clusters with significantly enriched canonical functions related to development and metabolism. *K*-means clustering analysis grouped the ceRNAs in the ceRNETS into different clusters. ceRNAs in the ceRNETS for the brain and lung were globally grouped into four clusters, whereas those in the liver and colon ceRNETS were mainly grouped into three clusters.

regulate long-term synaptic potentiation (LTP) or synaptic plasticity to further improve the development of the brain after birth. For example, LTP ( $adjP = 0.0165$ ) or memory ( $adjP = 0.0118$ ) were only enriched for this cluster, where these function are essential for learning and memory. In particular, *NRGN* and four calmodulin-dependent protein kinases were most highly expressed during the infant period, which are important for memory, as well as neuropsychiatric disorders (37). Overall, these observations suggest that ceRNAs may regulate various biological processes dur-

ing different developmental stages to promote tissue development. The ceRNA clusters obtained included genes encoding molecules with previously assigned functions during tissue development (e.g. *MEF2C*), as well as many genes with functions that have not been studied during tissue development to the best our knowledge, especially lncRNAs, thereby providing a rich resource for discovering additional molecular participants in tissue development.

**Table 1.** The interaction preference of different ceRNA clusters in four tissues

Brain	C1	C2	C3	C4	Colon	C1	C2	C3
C1	1.68 <sup>a</sup>	0.26 <sup>b</sup>	0.19 <sup>b</sup>	0.003 <sup>b</sup>	C1	2.42 <sup>a</sup>	0.23 <sup>b</sup>	0.18 <sup>b</sup>
C2		5.71 <sup>a</sup>	0.11 <sup>b</sup>	0.39 <sup>b</sup>	C2		2.09 <sup>a</sup>	0.16 <sup>b</sup>
C3			9.96 <sup>a</sup>	0.47 <sup>b</sup>	C3			3.96 <sup>a</sup>
C4				4.10 <sup>a</sup>				
Lung	C1	C2	C3	C4	Liver	C1	C2	C3
C1	4.21 <sup>a</sup>	0.43 <sup>b</sup>	0.01 <sup>b</sup>	0.31 <sup>b</sup>	C1	1.68 <sup>a</sup>	0.22 <sup>b</sup>	0.18 <sup>b</sup>
C2		3.28 <sup>a</sup>	0.27 <sup>b</sup>	0.02 <sup>b</sup>	C2		3.69 <sup>a</sup>	0.43 <sup>b</sup>
C3			2.01 <sup>a</sup>	0.29 <sup>b</sup>	C3			3.27 <sup>a</sup>
C4				7.71 <sup>a</sup>				

Note: The upper triangular and the diagonal show the ratio = (number of interactions in real condition)/(average number of interactions in random networks).

<sup>a</sup>Significantly over-represent.

<sup>b</sup>Significantly under-represent.

### lncRNAs are associated with the development of specific tissues according to ceRNA interactions

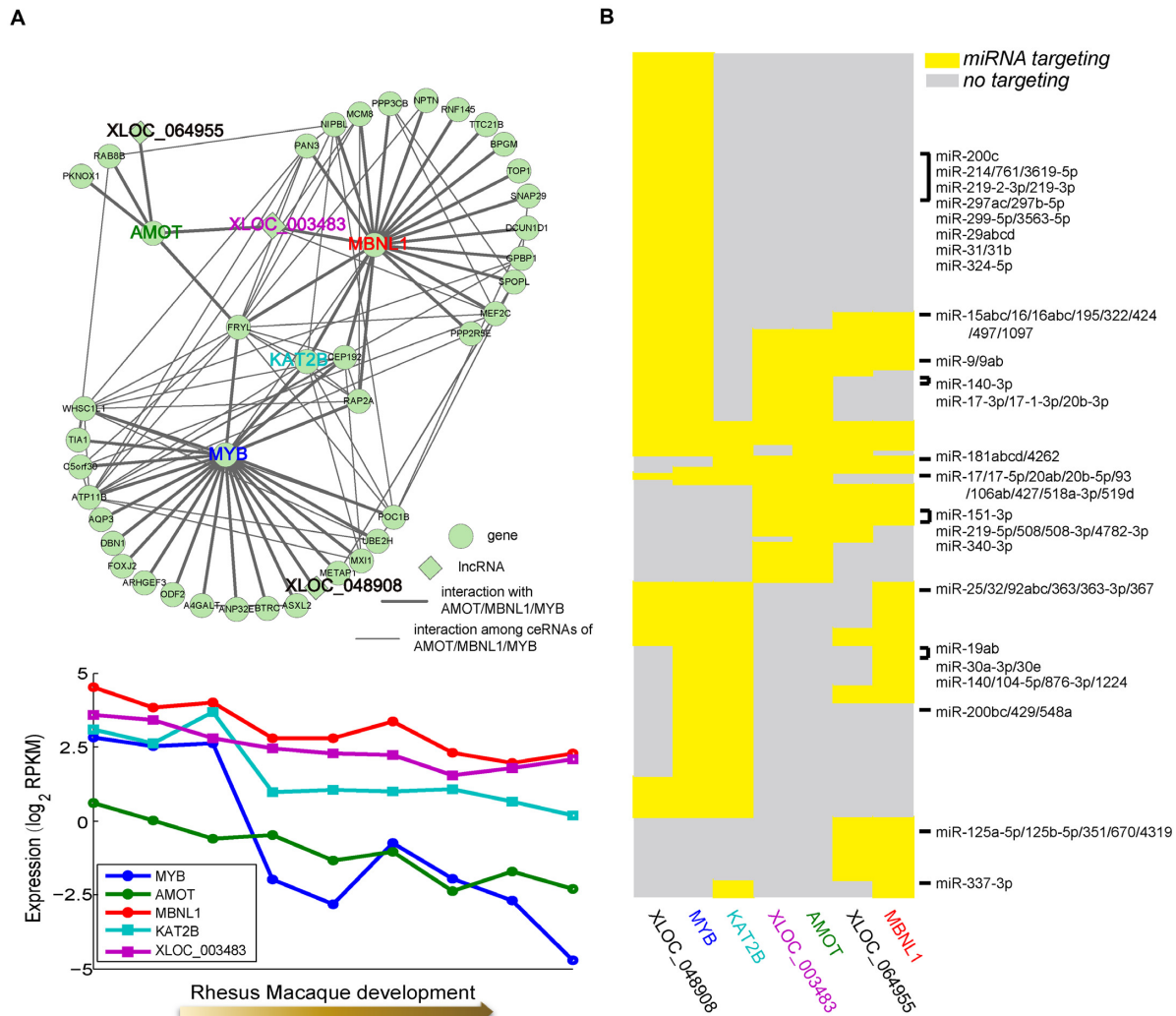
In mammalian cell types, lncRNAs are dynamically expressed and they function during tissue development. However, lncRNAs that act as ceRNAs have not yet been implicated in tissue development in the rhesus monkey, especially during prenatal development. As shown in Figure 2, lncRNAs exhibited remarkable stage-specific expression patterns in our study system. In agreement with the expression trends of mRNAs, most of the lncRNAs were highly expressed during the early prenatal development of different tissues but not after birth. In addition, we identified lncRNAs with dynamic expression throughout various developmental stages using an extraction of differential gene expression (EDGE-) based methodology (38), which showed that 68.37% of the lncRNAs in the brain ceRNET exhibited stage-specific expression. Our results indicate that lncRNAs exhibited slightly but significantly higher temporal specificity than that under random conditions according to the randomization tests ( $P = 0.0047$ , more details in Text S4), thereby suggesting that they might play important roles in determination and/or functions during specific developmental stages. We also explored the tissue specificity of lncRNAs that act as ceRNAs by calculating the number of tissues in which they occurred. We found that 15.50–19.97% of the lncRNAs functioned as ceRNAs in only one tissue, which was about two times higher than the proportion of mRNAs. Moreover, about one-third of the lncRNAs operated in more than two tissues, whereas half of the mRNAs were shared by three tissues (Supplementary Figure S3A and B). After deleting the RNAs with high tissue-specific expression based on the tissue specificity index ( $>0.7$ ) (39), we obtained the similar results (Supplementary Figure S3C and D). Thus, we concluded that for the lncRNAs that act as ceRNAs during mammalian development exhibit high tissue specificity.

Many previous studies have discovered that ceRNA interactions have immediate pragmatic applications in uncovering the functions of lncRNAs that are poorly annotated (10,11). Considering the spatial and temporal specificity of lncRNAs, we expected that if lncRNAs function as ceRNAs to regulate tissue development, then their compet-

ing interaction partners should have functions related to this process. Thus, we detected all of the lncRNA-related competitive interactions with known embryo development genes. It is showed that many lncRNAs could be involved in embryo development by directly regulating the expression of known embryo development genes, such as MEF2C, TGFBR3 and TCF7, thereby suggesting that these lncRNAs may modulate genes expression via shared MREs during tissue development. For example, in the brain ceRNET, lncRNA XLOC\_062139 interacted with 10 known brain development-related genes. In addition, XLOC\_062139 was regulated by seven known brain-related miRNAs. Thus, we inferred that XLOC\_062139 may be involved as a potential regulator during brain development.

To predict the further genome-wide the functions of lncRNAs during tissue development, we determined the enrichment of GO terms for coding genes directly and indirectly interacted with lncRNAs, which demonstrated the enrichment of genes involved in development, morphogenesis, and transcriptional processes (Supplementary Table S3). For example, the function of ‘*In utero* embryonic development’ ( $\text{adj}P = 0.0297$ ) was significantly enriched for the interactive partners of lncRNAs during early prenatal liver development, especially three genes (MYB, AMOT and MBNL1) that directly interacted with three lncRNAs (XLOC\_003483, XLOC\_064955 and XLOC\_048908), as shown in Figure 3A. We also found that several other competition partners (e.g. KAB2T, MEF2C and RAP2A) were located at the central in the subnetwork, thereby implying that they might be related with utero embryonic development (Figure 3A). KAB2T, also known as PCAF, has a histone acetyl transferase activity with core histones and nucleosome core particles, where it plays a direct role during transcriptional regulation. We found that this gene interacted with two genes annotated in utero embryonic development. Growing evidence suggests that changes in transcriptional regulation comprise an important part of the genetic basis of the evolution of development. In addition, the expression of PCAF in hepatocellular carcinoma differs significantly from that in the adjacent liver tissues, and it is correlated with tumor staging, thereby suggesting its important roles in the liver. Next, we aimed to discover the poten-





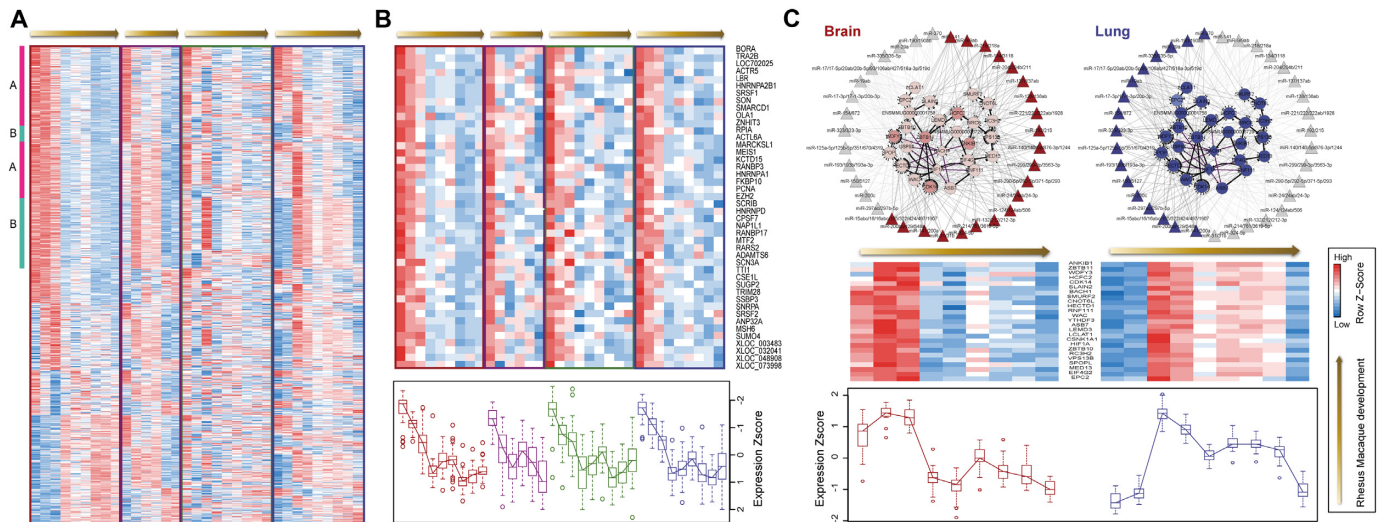
**Figure 3.** IncRNAs with characteristic expression pattern related to tissue development. (A) Three lncRNAs interacted closely with the three coding genes (MYB, AMOT and MBNL1) and their partners in the ceRNETS. Most of these ceRNAs were highly expressed during the early prenatal liver development. (B) Three lncRNAs interacted with the three coding genes by competitively binding many development-related miRNAs. Labels show the development-related miRNAs.

tial functions of lncRNAs by analyzing the organization of the ceRNA subnetwork (Figure 3A). We found that three lncRNAs interacted closely with these three coding genes by competitively binding many development-related miRNAs (Figure 3B), e.g. the lncRNA XLOC\_003483, which interacted with two of the three genes and it was highly expressed during early prenatal development of the liver. This suggests that this lncRNA may play important roles during early prenatal liver development. Identifying of these lncRNAs as ceRNAs allows us to conclude that lncRNAs play key roles in complex regulatory networks, which may improve our understanding of rhesus monkey development.

#### Shared ceRNA interactions in multiple tissues during development

We discovered that certain proportions of mRNAs and lncRNAs acted as ceRNAs during the development of multiple tissues in the rhesus monkey. Next, we systemically

explored the shared competitive rules during tissue development. We found that some of the common ceRNAs tended to have similar expression patterns in all four tissues throughout the development process (Figure 4). In particular, one sub-group comprising 44 ceRNAs (Figure 4B) exhibited consistent high expression during early development, but lower expression after birth. Moreover, we assembled the competitive interactions of these ceRNAs in at least two ceRNETS, where they formed a tightly connecting subnetwork (Supplementary Figure S4). Next, we performed function enrichment analysis of the ceRNA interaction subnetwork to explore the biological processes that might be mediated by ceRNAs. We found that this class of genes was enriched for multiple categories related to transcription and cell cycle (Supplementary Table S4). Remarkably, thirty-five nodes were transcription factors, although the proportion of TFs in the entire ceRNA networks was low (Supplementary Table S5). These results demonstrate that transcription factors and ceRNA networks are intertwined.



**Figure 4.** Shared competitive rules in tissue development. (A) The expression patterns of common ceRNAs in all four tissues throughout the development process. (B) Forty-four ceRNAs with consistently high expression levels in the four tissues during early development, but lower expression after birth. The z-scores for these ceRNA expression levels in each developmental stage. (C) A ceRNA module shared by brain and lung developmental processes. Red triangular nodes denote brain-related miRNAs and blue triangular nodes are lung-related. Arrowed lines represent regulation from miRNAs to ceRNAs. Black solid lines represent ceRNA interactions. The expression patterns of these ceRNAs were distinct during brain and lung development.

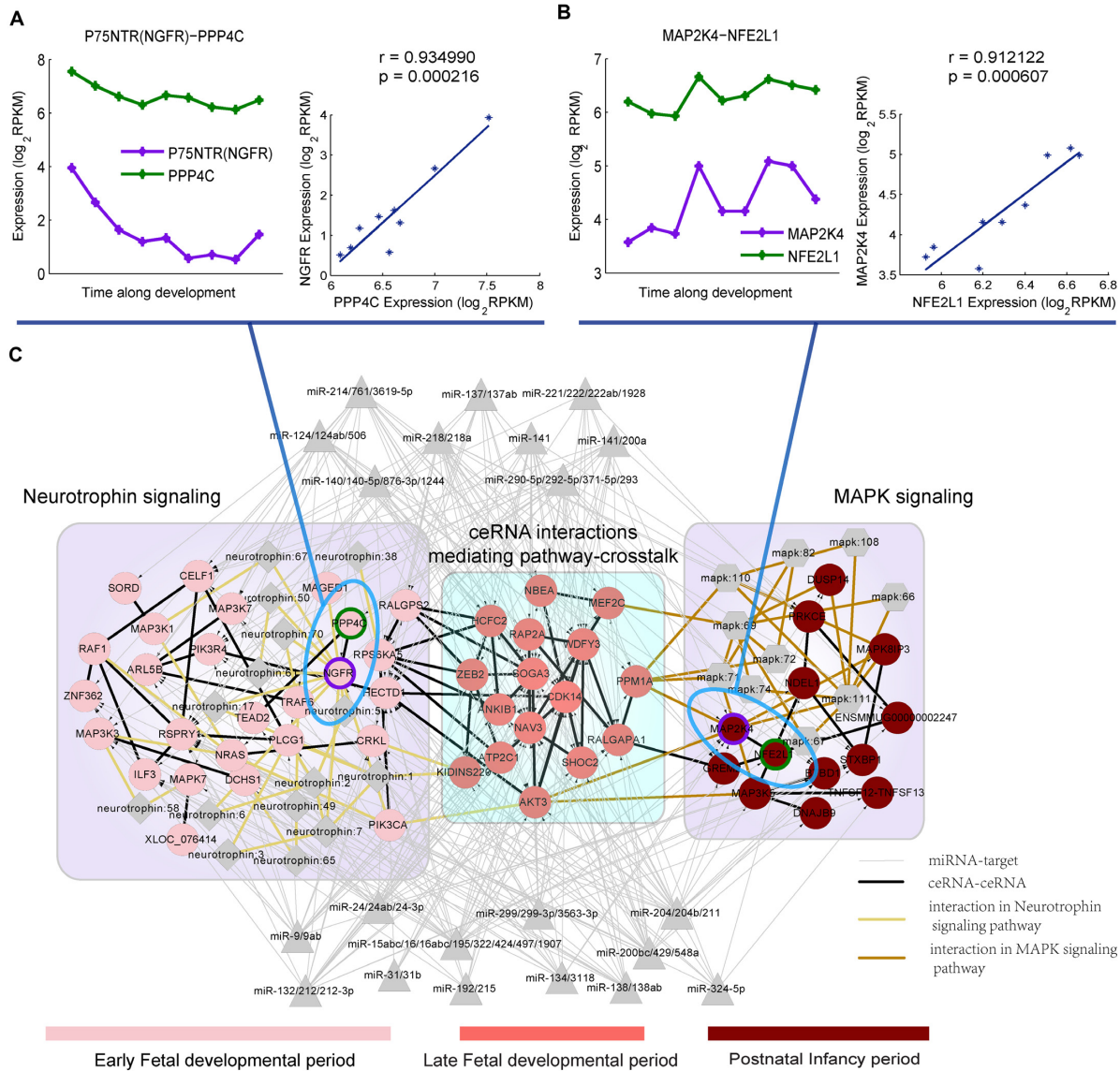
Ala *et al.* previously analyzed the relationship between transcriptional regulation and ceRNA networks in human cell lines, and concluded that the efficiency of ceRNA-mediated cross-regulation depends on miRNA fluctuations, whereas the activity of TFs toward their targets is more stable and independent of the miRNA expression levels (21). Therefore, we concluded that the regulatory potency of a TF's regulatory potency might be much larger than anticipated in transcription-related functions during tissue development.

By analyzing the ceRNETS, we identified a ceRNA module shared by brain and lung development processes (Figure 4C). Among these ceRNA interactions, 67% or 81% were mediated by 16 brain-related or 11 lung-related miRNAs, respectively. In particular, in the context of lung development, 14 (51.85%) ceRNA interactions were mediated by a key miRNA cluster, miR-17-92. It is shown that the miR-17-92 cluster has a high expression level during the early stages of lung development, but it declines as development proceeds (40). These ceRNA interactions may play important roles in fine-tuning the regulation of miRNA expression during lung development. Furthermore, these same ceRNA interactions were mediated by distinct miRNAs during brain development (Figure 4C). For example, miR-137 has important roles in inhibiting neural stem cell proliferation and promoting neuronal differentiation, which could regulate nine ceRNAs in this module. In addition, the expression patterns of these ceRNAs were distinct during brain and lung development, where most of the ceRNAs were highly expressed during early prenatal brain development, whereas they were highly expressed during the medium-term or postnatal infancy period in lung development (Figure 4C). These results indicate that the ceRNAs can regulate the general and common functions in a more subtle and specific manner during tissue development by affecting miRNA mediated regulation and/or expression patterns.

#### ceRNA-mediated pathway crosstalk during brain development

As described above, we found that ceRNA–ceRNA interactions were generally associated with gene pairs within the same functions. In addition, it is important to understand ceRNA interactions and the crosstalk between pathways to elucidate the functions of cells and more complex systems. For example, many different types of neurons are precisely connected during brain development, where they are guided by multiple corresponding active signaling pathways. Remarkably, we found that some ceRNA interactions only became active during specific development stage and they mediated the crosstalk among key pathways in brain development. As shown in Figure 5C, during early stage of brain development, the neurotrophin signaling pathway appeared to be fine-tuned by 19 pairs of ceRNA interactions. And the MAP kinase (MAPK) signaling pathway was then activated during the postnatal infancy period. Neurotrophins comprise a family of proteins that regulate the development, maintenance, and function of vertebrate nervous systems (41). The neurotrophins p75NTR (NGFR) was found to be fine-tuned by ceRNA-PPP4C. We found that PPP4C was highly expressed during the early stage (Figure 5A), where it had a high capacity for sponging the miRNAs (miR-128/128ab, miR-27abc/27a-3p, miR-486-3p, miR-612/3150a-3p and miR-939/1343) shared with p75NTR. This might ensure that sufficient p75NTR is present to activate the downstream pathways.

In addition, neurotrophins are known to activate many signaling pathways, including the MAPK, PI-3 kinase, and Jun kinase cascades (41,42). We found that several ceRNA interactions were involved during the middle stage of brain development, including AKT3-CDK14 and MEF2C-NAV3. The MAPK signaling pathway was then activated



**Figure 5.** ceRNA interactions that mediated the crosstalk between the neurotrophin signaling pathway and the MAPK signaling pathway during brain development. (A) Expression levels of P75NTR (NGFR) and PPP4C during each brain development stage, and the expression relationship between these two genes. (B) Expression levels of MAP2K4 and NFE2L1 during each brain development stage and their expression relation. (C) ceRNAs involved in key pathways during brain development. The ceRNA nodes in the left shadow participated in the neurotrophin signaling pathway. Light pink nodes were highly expressed during early fetal development. The ceRNA nodes in the right shadow participated in the MAPK signaling pathway. Red nodes were highly expressed in the postnatal infancy period. The ceRNA nodes in the middle shadow mediated pathway-crosstalk. Pink nodes were highly expressed in the late fetal developmental period.

in the postnatal infancy period. MAPKs are intracellular signaling pathways that play pivotal roles in many essential cellular processes such as proliferation and differentiation. It has been shown that the expression of MAP2K4 increases gradually in the central nervous systems during late embryogenesis, with the highest levels at E18 and after birth. In addition to the direct physical interactions between upstream protein kinases with MAP2K4, we suggest that the ceRNAs interacting with MAP2K4 may also be regulators of its expression, such as NFE2L1, which was also highly expressed during the postnatal infancy period (Figure 5B), and it shared 20 miRNAs with MAP2K4. Thus, miR-mediated interactions provide a mechanism related to

high expression levels. Overall, our results suggest that during brain development, the cell responses involve a discrete set of ceRNA interactions, which act sequentially and/or synergistically and they may also mediate crosstalk between pathways in the context of development.

## DISCUSSION

In this study, we developed a computational algorithm to generate genome-wide ceRNETs for the whole transcriptome including protein coding genes and lncRNAs, which allowed us to hear and translate this new ‘language’ mediated by MREs for communicating between transcripts dur-

ing the development of different tissues. We found that indirect ceRNA interactions were common and most of the interactions were connected by sharing numerous MREs, thereby forming four robust biological networks in each tissue. Moreover, the ceRNAs exhibited both dynamic expression and regulation patterns during the developmental processes. Genes that are highly expressed in particular developmental stages or cell types are functionally active. Thus, in the four tissues, the ceRNAs exhibited stage-specific expression, where most had their highest expression levels in the fetal development period. These trends extended from brain tissue to the other three tissue types, which might be a general trend in tissue development. We also found that ceRNA interactions within the same development period were enhanced, whereas the ceRNAs that operated in different stages tended to avoid interactions, especially the ceRNA clusters in the early prenatal stage and those in the infant stage. Further functional analysis showed that the dynamic expression and regulation pattern were related to the expected tissue- or phase-specific functions. Therefore, these results suggest that ceRNAs might work synergistically during different developmental stage to control specific functions. In particular, ceRNA interactions also mediated the coordination of different functions during brain development.

Crosstalk between coding and noncoding RNAs could answer questions related to the evolution of development. It is known that human lncRNAs are under weak selective constraints and many are primate specific (43). A previous study showed that 63% of human lincRNAs have orthologues in the rhesus monkey, which is higher than the number found in the mouse and rat (44). Therefore, the rhesus monkey might be a better model for analyzing the roles of lncRNAs during development. In addition, transcriptome analysis during development depends mainly on existing annotations, but RNA-seq now allows the flexible and potentially unbiased characterization of the transcriptome including protein-coding genes and lncRNAs. According to previous studies, lncRNAs exhibit more tissue-specific expression than protein-coding genes (43), and we also observed that the competitive interactions of lncRNAs were also more tissue-specific. In addition, it should be noted that in lung tissue, the expression levels of 27.5% of the lncRNAs increased during the middle to late fetal stages, although most of the lncRNAs were highly expressed during early prenatal development in different tissues (Figure 2). This proportion is slightly but significantly higher than that determined under random conditions by the randomization tests ( $P = 0.0044$ , more details in Text S4). Thus, we conclude that these lncRNAs might be involved in lung adaptation before and after birth.

Our functional analysis of ceRNAs in different tissues discovered many general developmental and metabolic processes, such as transcriptional-related function. Moreover, some of the ceRNA interactions were also shared by multiple tissues. Recently, ceRNAs were examined in the context of other cellular networks, which showed that they are tightly integrated within the transcriptional regulatory network (21). During the developmental processes, we found that ceRNA interactions and transcriptional regulatory networks were also intertwined because transcription fac-

tor RNAs and their regulatory target mRNAs engaged in crosstalk with other RNA transcripts. This suggests the intriguing possibility that a transcription factor's regulatory potency is much larger than anticipated, which has important implications for cellular and organismal physiology, as well as for the pathogenesis of cancer and other diseases where aberrant transcription factor expression occurs.

Transcriptome analysis of brain tissues can provide fundamental insights into development and disease. Based on analyses of other publicly available microarray data for genes and miRNAs in the postnatal brains of rhesus monkeys and humans (GSE22570 and GSE18069), we found that the ceRNA pairs identified in rhesus monkey were robust and conserved during the human brain developmental process, thereby providing further insights into understand the mechanism of human brain development (details in Text S5). The conservation of ceRNA crosstalk suggests that it may represent an important and widespread layer for RNA regulation. In addition, perturbations of the ceRNA network may disrupt normal development and even cause the development-related diseases. It is known that developmental brain disorders often involve genes that are differentially expressed during fetal development compared with the postnatal period. Indeed, autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental disorder in which hundreds of genes have been implicated, but it remains unclear how these genes may converge to affect brain development, which is critical for achieving a mechanistic understanding of ASD. By assessing the enrichment of an independent set of probable ASD (pASD) genes, derived from a recent study (45), we found that the pASD genes had significant more ceRNA partners in the ceRNET (Supplementary Figure S5,  $P = 0.00054$ , Wilcoxon's rank-sum test). In particular, 81.9% of the ASD genes were highly expressed during the fetal stage and they interacted with each other. These pASD genes with high connectivity may play key roles in dysregulated cellular processes. For example, the SETD2 gene has important roles in chromatin structure modulation and it interacted with 33 ceRNAs in the brain development associated ceRNET. De novo variants of this gene were identified in two separate studies of ASD (46,47). Another key gene is MIER3, but no genetic variant has been found of this gene. However, we found that this gene interacts with 23 ceRNAs in the ceRNET, thereby indicating that it has important roles during brain development, which may also be implicated in ASD. Evidences suggests that this gene may play key roles in transcriptional regulation, and it was recently shown that transcriptional dysregulation is involved in ASD. These results suggest that ceRNETs will be useful for understanding a considerable proportion of human diseases with genetic components, and analyses of the structure of ceRNETs will allow us to rapidly screen putative disease-associated ceRNAs *in silico*.

We considered that our method might be a good choice for identifying ceRNA interactions during embryo development. Our method determines the global binding patterns of MREs to predict ceRNA interactions by considering both the relatively small size of embryo development data sets and the successful identification of ceRNA interactions in humans without miRNA expression (16,21). The enrichment analysis of shared miRNAs and the additional

four rules were designed to assess the global similarity of miRNA regulation between two RNAs, which might be a useful method for predicting ceRNA interactions because the predicted targets could be true targets with higher probability than random predictions (48,49). The integration of miRNA expression data is not suitable for our current method, because correlation with expression levels are generally calculated between miRNAs and the predicted target RNAs, whereas the detail information related to MREs is discarded. Moreover, the robust and conservation analysis of ceRNA interactions in the brain provides additional support for our method. Moreover, an intended use of our ceRNETS is to discover putative ceRNA interactions as candidates for experimental validation. However, some challenges still hinder the application of the ceRNA model to explore the molecular mechanisms of embryo development. For example, any ceRNA interaction study relies to some extent on target prediction. Thus, improving the accuracy of the predicted miRNA–target interactions is often the first step toward identifying ceRNA crosstalk. Other different contributory factors that may modulate ceRNA interactions also await further characterization, such as subcellular localization, RNA editing, and the negative correlation of expression of ceRNA pairs (10). These challenges should be addressed in future ceRNA prediction analyses, which may potentially add further complexities to the dynamics of ceRNA regulation during primary embryo development. We expect that the identification of miRNA and ceRNA interactions will become progressively more accurate as our knowledge of miRNA regulatory mechanism increases with the size of the molecular profiles during primary embryo development.

In summary, the application of ceRNA network analysis to transcriptomes obtained during tissue development provides a novel approach for understanding gene functionality which can yield many new insights into the structure of molecular pathways. Analyzing ceRNA interactions in the context of tissue development will provide insights into the regulation of normal cell development, as well as the dysregulation of key mechanisms of pathogenesis. Furthermore, this analysis sheds light on the potential functions of non-coding RNAs and the regulatory roles of mRNAs independent of their protein coding roles, thereby identifying regulatory networks that might have been overlooked by conventional protein-coding studies. The ceRNA networks discovered in this study still need to be examined further by analyzing the expression profiles of more developmental samples, but our results have exciting implications for understanding the mechanisms of development and the identification of new therapeutic approaches for complex diseases.

## SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

## FUNDING

National High Technology Research and Development Program of China (863 Program) [2014AA021102 and 2014AA020602]; National Program on Key Basic Research

Project (973 Program) [2014CB910504]; National Natural Science Foundation of China [91439117, 61473106, 61203264, 31571331 and 61502126]; Weihai Youth Science Fund Project of Harbin Medical University. Funding for open access charge: National High Technology Research and Development Program of China (863 Program) [2014AA021102 and 2014AA020602]; National Program on Key Basic Research Project (973 Program) [2014CB910504].

*Conflict of interest statement.* None declared.

## REFERENCES

1. Assou, S., Boumela, I., Haouzi, D., Anahory, T., Dechaud, H., De Vos, J. and Hamamah, S. (2011) Dynamic changes in gene expression during human early embryo development: from fundamental aspects to clinical applications. *Hum. Reprod. Update*, **17**, 272–290.
2. Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J. *et al.* (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.*, **20**, 1131–1139.
3. Colantuoni, C., Lipska, B.K., Ye, T., Hyde, T.M., Tao, R., Leek, J.T., Colantuoni, E.A., Elkhouloun, A.G., Herman, M.M., Weinberger, D.R. *et al.* (2011) Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature*, **478**, 519–523.
4. Somel, M., Guo, S., Fu, N., Yan, Z., Hu, H.Y., Xu, Y., Yuan, Y., Ning, Z., Hu, Y., Menzel, C. *et al.* (2010) MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.*, **20**, 1207–1218.
5. Oldham, M.C., Konopka, G., Iwamoto, K., Langfelder, P., Kato, T., Horvath, S. and Geschwind, D.H. (2008) Functional organization of the transcriptome in human brain. *Nat. Neurosci.*, **11**, 1271–1282.
6. Li, F., Xiao, Y., Huang, F., Deng, W., Zhao, H., Shi, X., Wang, S., Yu, X., Zhang, L., Han, Z. *et al.* (2015) Spatiotemporal-specific lncRNAs in the brain, colon, liver and lung of macaque during development. *Mol. bioSyst.*, **11**, 3253–3263.
7. Somel, M., Liu, X., Tang, L., Yan, Z., Hu, H., Guo, S., Jiang, X., Zhang, X., Xu, G., Xie, G. *et al.* (2011) MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. *PLoS Biol.*, **9**, e1001214.
8. Xu, J., Li, C.X., Li, Y.S., Lv, J.Y., Ma, Y., Shao, T.T., Xu, L.D., Wang, Y.Y., Du, L., Zhang, Y.P. *et al.* (2011) MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. *Nucleic Acids Res.*, **39**, 825–836.
9. Li, Y., Xu, J., Chen, H., Bai, J., Li, S., Zhao, Z., Shao, T., Jiang, T., Ren, H., Kang, C. *et al.* (2013) Comprehensive analysis of the functional microRNA-mRNA regulatory network identifies miRNA signatures associated with glioma malignant progression. *Nucleic Acids Res.*, **41**, e203.
10. Tay, Y., Rinn, J. and Pandolfi, P.P. (2014) The multilayered complexity of ceRNA crosstalk and competition. *Nature*, **505**, 344–352.
11. Salmena, L., Poliseno, L., Tay, Y., Kats, L. and Pandolfi, P.P. (2011) A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*, **146**, 353–358.
12. Ergun, S. and Oztuzcu, S. (2015) Oncocers: ceRNA-mediated cross-talk by sponging miRNAs in oncogenic pathways. *Tumour Biol.*, **36**, 3129–3136.
13. Arancio, W., Carina, V., Pizzolanti, G., Tomasello, L., Pitrone, M., Baiamonte, C., Amato, M.C. and Giordano, C. (2015) Anaplastic thyroid carcinoma: a ceRNA analysis pointed to a crosstalk between SOX2, TP53, and microRNA biogenesis. *Int. J. Endocrinol.*, 439370.
14. Poliseno, L. and Pandolfi, P.P. (2015) PTEN ceRNA networks in human cancer. *Methods*, **77–78**, 41–50.
15. Sumazin, P., Yang, X., Chiu, H.S., Chung, W.J., Iyer, A., Llobet-Navas, D., Rajbhandari, P., Bansal, M., Guarnieri, P., Silva, J. *et al.* (2011) An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. *Cell*, **147**, 370–381.
16. Tay, Y., Kats, L., Salmena, L., Weiss, D., Tan, S.M., Ala, U., Karreth, F., Poliseno, L., Provero, P., Di Cunto, F. *et al.* (2011) Coding-independent

- regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell*, **147**, 344–357.
17. Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., Tramontano, A. and Bozzoni, I. (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*, **147**, 358–369.
  18. Cheng, E.C. and Lin, H. (2013) Repressing the repressor: a lincRNA as a MicroRNA sponge in embryonic stem cell self-renewal. *Dev. Cell*, **25**, 1–2.
  19. Wang, Y., Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., Wu, M., Xiong, J., Guo, X. and Liu, H. (2013) Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev. Cell*, **25**, 69–80.
  20. Zhao, J., Zhou, Y., Zhang, X. and Chen, L. (2016) Part mutual information for quantifying direct associations in networks. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 5130–5135.
  21. Ala, U., Karreth, F.A., Bosia, C., Pagnani, A., Taulli, R., Leopold, V., Tay, Y., Provero, P., Zecchina, R. and Pandolfi, P.P. (2013) Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 7154–7159.
  22. Xu, J., Li, Y., Lu, J., Pan, T., Ding, N., Wang, Z., Shao, T., Zhang, J., Wang, L. and Li, X. (2015) The mRNA related ceRNA-ceRNA landscape and significance across 20 major cancer types. *Nucleic Acids Res.*, **43**, 8169–8182.
  23. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.
  24. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. and Pachter, L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.*, **7**, 562–578.
  25. Mortazavi, A., Schwarz, E.M., Williams, B., Schaeffer, L., Antoshechkin, I., Wold, B.J. and Sternberg, P.W. (2010) Scaffolding a *Caenorhabditis* nematode genome with RNA-seq. *Genome Res.*, **20**, 1740–1747.
  26. Gong, B., Liu, T., Zhang, X., Chen, X., Li, J., Lv, H., Zou, Y., Li, X. and Rao, S. (2011) Disease embryo development network reveals the relationship between disease genes and embryo development genes. *J. Theor. Biol.*, **287**, 100–108.
  27. Zhang, S.J., Liu, C.J., Shi, M., Kong, L., Chen, J.Y., Zhou, W.Z., Zhu, X., Yu, P., Wang, J., Yang, X. *et al.* (2013) RhesusBase: a knowledgebase for the monkey research community. *Nucleic Acids Res.*, **41**, D892–D905.
  28. Warnefors, M., Liechti, A., Halbert, J., Valloton, D. and Kaessmann, H. (2014) Conserved microRNA editing in mammalian evolution, development and disease. *Genome Biol.*, **15**, R83.
  29. Chuang, T.J. and Chiang, T.W. (2014) Impacts of pretranscriptional DNA methylation, transcriptional transcription factor, and posttranscriptional microRNA regulations on protein evolutionary rate. *Genome Biol. Evol.*, **6**, 1530–1541.
  30. Alexiou, P., Maragkakis, M., Papadopoulos, G.L., Reczko, M. and Hatzigeorgiou, A.G. (2009) Lost in translation: an assessment and perspective for computational microRNA target identification. *Bioinformatics*, **25**, 3049–3055.
  31. Garcia, D.M., Baek, D., Shin, C., Bell, G.W., Grimson, A. and Bartel, D.P. (2011) Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat. Struct. Mol. Biol.*, **18**, 1139–1146.
  32. Saetrom, P., Heale, B.S., Snove, O. Jr, Aagaard, L., Alluin, J. and Rossi, J.J. (2007) Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic Acids Res.*, **35**, 2333–2342.
  33. Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A. and Rinn, J.L. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.*, **25**, 1915–1927.
  34. Rousseeuw, P.J. (1987) Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Computat. Appl. Math.*, **20**, 53–65.
  35. Li, Y., Xu, J., Chen, H., Zhao, Z., Li, S., Bai, J., Wu, A., Jiang, C., Wang, Y., Su, B. *et al.* (2013) Characterizing genes with distinct methylation patterns in the context of protein-protein interaction network: application to human brain tissues. *PLoS One*, **8**, e65871.
  36. Barabasi, A.L., Gulbahce, N. and Loscalzo, J. (2011) Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.*, **12**, 56–68.
  37. Ohi, K., Hashimoto, R., Yasuda, Y., Fukumoto, M., Yamamori, H., Umeda-Yano, S., Fujimoto, M., Iwase, M., Kazui, H. and Takeda, M. (2013) Influence of the NRGN gene on intellectual ability in schizophrenia. *J. Hum. Genet.*, **58**, 700–705.
  38. Leek, J.T., Monsen, E., Dabney, A.R. and Storey, J.D. (2006) EDGE: extraction and analysis of differential gene expression. *Bioinformatics*, **22**, 507–508.
  39. Chen, G., Qiu, C., Zhang, Q., Liu, B. and Cui, Q. (2013) Genome-wide analysis of human SNPs at long intergenic noncoding RNAs. *Hum. Mutat.*, **34**, 338–344.
  40. Bonauer, A. and Dimmeler, S. (2009) The microRNA-17-92 cluster: still a miRacle? *Cell Cycle*, **8**, 3866–3873.
  41. Ascano, M., Bodmer, D. and Kuruvilla, R. (2012) Endocytic trafficking of neurotrophins in neural development. *Trends Cell Biol.*, **22**, 266–273.
  42. Bartkowska, K., Turlejski, K. and Djavadian, R.L. (2010) Neurotrophins and their receptors in early development of the mammalian nervous system. *Acta Neurobiol. Exp. (Wars)*, **70**, 454–467.
  43. Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G. *et al.* (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.*, **22**, 1775–1789.
  44. Washietl, S., Kellis, M. and Garber, M. (2014) Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six mammals. *Genome Res.*, **24**, 616–628.
  45. Parikhshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S. and Geschwind, D.H. (2013) Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*, **155**, 1008–1021.
  46. O’Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., Levy, R., Ko, A., Lee, C., Smith, J.D. *et al.* (2012) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*, **485**, 246–250.
  47. O’Roak, B.J., Vives, L., Fu, W., Egerton, J.D., Stanaway, I.B., Phelps, I.G., Carvill, G., Kumar, A., Lee, C., Ankenman, K. *et al.* (2012) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science*, **338**, 1619–1622.
  48. Wang, J., Ma, R., Ma, W., Chen, J., Yang, J., Xi, Y. and Cui, Q. (2016) LncDisease: a sequence based bioinformatics tool for predicting lincRNA-disease associations. *Nucleic Acids Res.*, **44**, e90.
  49. Bartel, D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215–233.