ARTICLE
Journal of Cellular Biochemistry 115:910–918 (2014)

Journal of Cellular Biochemistry

Long Noncoding RNAs Expression Profile of the Developing Mouse Heart

Jin Gai Zhu,[1] Ya Hui Shen,[2] Hai Lang Liu,[3] Ming Liu,[3] Ya Qing Shen,[4] Xiang Qing Kong,[3] Gui Xian Song ,[3]** and Ling Mei Qian[3]*

[1]*Department of Pediatrics, Nanjing Maternity and Child Health Care Hospital Affiliated to Nanjing Medical University, Nanjing 210029, People's Republic of China*

[2]*Department of MICU, Nanjing Maternity and Child Health Care Hospital Affiliated to Nanjing Medical University, Nanjing 210029, People's Republic of China*

[3]*Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, People's Republic of China*

[4]*Department of Laboratory Medicine, Taizhou Maternity and Child Health Care Hospital, Taizhou 225300, People's Republic of China*

ABSTRACT

Long noncoding RNAs (lncRNAs) represent a sub-group of noncoding RNAs that are longer than 200 nucleotides. The characterization of lncRNAs and their acceptance as crucial regulators of numerous developmental and biological pathways have suggested that the lncRNA study has gradually become one of the hot topics in the field of RNA biology. Many lncRNAs show spatially and temporally restricted expression patterns during embryogenesis and organogenesis. This study aimed to characterize the lncRNA profile of the fetal mouse heart at three key time points (embryonic day E11.5, E14.5, and E18.5) in its development, by performing a microarray lncRNAs screen. Gene Ontology analysis and ingenuity pathway analysis showed some significant gene functions and pathways were altered in heart development process. We compared lncRNAs profile between the three points (E14.5 vs. E11.5 [early development]; E18.5 vs. E14.5 [later development]). A total of 1,237 lncRNAs were found to have consistent fold changes (>2.0) between the three time points. Among them, 20 dysregulated lncRNAs were randomly selected and confirmed by real-time qRT-PCR. Additionally, bioinformatics analysis of AK011347 suggested it may be involved in heart development through the target gene Map3k7. In summary, this study identified differentially expressed lncRNAs in the three time points studied, and these lncRNAs may provide a new clue of mechanism of normal heart development. J. Cell. Biochem. 115: 910–918, 2014. © 2013 Wiley Periodicals, Inc.

KEY WORDS: HEART DEVELOPMENT; LONG NONCODING RNA; MICROARRAY; BIOINFORMATICS

The heart is the first functional organ during mouse embryonic development. More than any other organ, the heart has to maintain a high level of function throughout the lifespan of the organism, starting from the early primitive heart tube, to formation of the heart chambers, and throughout life [Liu and Olson, 2010]. Heart development requires precise temporal spatial regulation of gene expression, in which the highly conserved modulation networks of transcription factors (TFs) accurately control the signaling pathways required for normal cardiovascular development. Therefore, the regulatory networks that control the development and adaptations of

Jin Gai Zhu and Ya Hui Shen have contributed equally to this work.

Conflict of interest: No conflict of interest exists in the submission of this manuscript.

Grant sponsor: Priority Academic Program Development of Jiangsu Higher Education Institutions; Grant sponsor: National Natural Science Foundation of China; Grant number: 81370278; Grant sponsor: Natural Science Foundation of Jiangsu Province, China; Grant number: SBK201340683; Grant sponsor: Talent Foundation of Jiangsu Province, China; Grant number: WSN-020; Grant sponsor: Postgraduate Research and Innovation Project in Jiangsu Province, China; Grant number: CXZZ13_0575.

* Correspondence to: Ling Mei Qian, Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, People's Republic of China. E-mail: lmqian@njmu.edu.cn

** Correspondence to: Gui Xian Song, Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, People's Republic of China. E-mail: songguixian1983@163.com

Manuscript Received: 26 September 2013; Manuscript Accepted: 4 December 2013

Accepted manuscript online in Wiley Online Library (wileyonlinelibrary.com): 16 December 2013

DOI 10.1002/jcb.24733 • © 2013 Wiley Periodicals, Inc.

the heart have been under active investigation [Wilczynski and Furlong, 2010]. Thanks to the development of new molecular and biologic techniques in the past decade, we have witnessed significant progress in elucidation of the molecular mechanisms of heart formation. In particular, several TFs such as NKX2.5 [Lints et al., 1993], Tbx5 [Bruneau et al., 2001], and GATA4 [Watt et al., 2004] have been identified as being essential for heart development. Nevertheless, the upstream regulators as well as interacting partners and downstream targets/effectors of the handful of TFs remain largely unknown.

Long noncoding RNAs (LncRNAs) are transcripts of at least 200 nucleotides transcribed from all over the genome [Mattick, 2009]. From a genetic point of view, lncRNAs fall into one or more of five broad categories: (1) sense or (2) antisense, when overlapping one or more exons of another transcript on the same or opposite strand, respectively; (3) bidirectional, when the expression of it and a neighboring coding transcript on the opposite strand is initiated in close genomic proximity; (4) intronic, when derived from an intron of a second transcript; or (5) intergenic, when it lies as an independent unit within the genomic interval between two genes [Ponting et al., 2009]. Recently, it became obvious that lncRNAs play an important role in regulating gene expression at various levels, including chromatin modification, transcription, and post-transcriptional processing [Mercer et al., 2009; Wilusz et al., 2009]. For example, the lncRNAs Xist (X inactive-specific transcript) or HOTAIR (HOX Antisense Intergenic RNA) interact with chromatin remodeling complexes to induce heterochromatin formation in specific genomic loci leading to reduced target gene expression [Rinn et al., 2007; Gupta et al., 2010; Tsai et al., 2010]. LncRNAs can also function by regulating transcription through a variety of mechanisms that includes interacting with RNA-binding proteins, acting as a co-activator of TFs, or repressing a major promoter of their target genes [Feng et al., 2006; Martianov et al., 2007; Wang et al., 2008]. In addition to chromatin modification and transcriptional regulation, lncRNAs can modulate gene expression at the post-transcriptional level or splicing level [Beltran et al., 2008; Faghihi et al., 2008; Tripathi et al., 2010].

Increasing studies have indicated that lncRNAs are crucial regulators of numerous developmental and biological pathways in the eukaryotic genome [Yan and Wang, 2012]. However, study of lncRNAs, which comprise the bulk of the noncoding transcriptome, is still in its infancy in heart. It is known that many lncRNAs show spatially and/or temporally restricted expression patterns [Mercer et al., 2009; Cabili et al., 2011]. Thus, by characterizing the spatial and temporal expression profiles of lncRNAs in the developing heart, we can improve our understanding of heart development and gene regulation.

The mouse heart shows great similarity to the human heart, with respect to anatomy, growth and development, making the mouse a suitable experimental model for biomedical research [Wessels and Sedmera, 2003]. The formation of the mouse heart starts to take shape at approximately E10.0, and at E11.5–E18.5, the tube undergoes a complex series of movements and tissue remodeling events that lead to the formation of the four-chambered heart [Savolainen et al., 2009]. Based on this developmental timeline, we selected three key time points (E11.5, E14.5, and E18.5) representing the process of normal heart development, to perform an lncRNAs screening by comprehensive lncRNAs microarray technique in C57BL/6 mice. The aim of this study was to explore the dynamic expression profile of lncRNAs in embryonic heart development, and offer a foundation for future functional analysis.

MATERIALS AND METHODS

EXPERIMENTAL ANIMALS

The Nanjing Medical University Animal Care and Use Committee approved the experimental protocols used in this study. Pathogen-free male and female C57BL/6J mice were obtained from the animal center of the Nanjing Medical University. The animals were housed in individual cases in a temperature-controlled room with a 12-h light/dark cycle. At the age of 6 months, the males and females were mated. Pregnancy was detected by visual inspection of a distended abdomen. At E11.5, E14.5, and E18.5, pregnant mice were sacrificed with CO_2, embryos were collected and fetal hearts dissected and pooled within each age group for further analysis.

HEMATOXYLIN AND EOSIN (H&E) STAINING

Collected fetal hearts were washed with cold PBS and then fixed in formalin overnight at 4°C. Sections (7 μm) of paraformaldehyde-fixed heart tissue were obtained and stained with H&E for morphological analysis. H&E sections were viewed under a light microscope at magnifications of 40× to observe changes in fetal heart development at the three experimental time points.

RNA EXTRACTION

To extract RNA, frozen tissues were ground into powder with mortar and pestle and resuspended in Trizol reagent (Invitrogen, Carlsbad, CA). The RNA purification was performed on the RNA-containing aqueous phase with RNeasy minikit (Qiagen). After elution with RNase-free water and treatment with turbo DNase (Ambion), the RNA is ready for all kinds of applications. Quantification and quality check were performed with Nanodrop and Agilent 2100 Bioanalyzer (Agilent Technologies), respectively.

MICROARRAY ANALYSIS

Arraystar Mouse LncRNA Microarray v 2.0 is designed for the global profiling of mouse lncRNAs and protein-coding transcripts. Each transcript is represented by a specific exon or splice junction probe which can identify individual transcript accurately. Positive probes for housekeeping genes and negative probes are also printed onto the array for hybridization quality control. Briefly, mRNA was purified from total RNA after removal of rRNA (mRNA-ONLY^{TM} Eukaryotic mRNA Isolation Kit, Epicentre). Then, each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcripts without 3′ bias utilizing a random priming method. The labeled cRNAs were hybridized onto the Mouse LncRNA Array v 2.0 (8 × 60 K, Arraystar). After having washed the slides, the arrays were scanned by the Agilent Scanner G2505C. Agilent Feature Extraction software (version 11.0.1.1) was used to analyze acquired array images. Quantile normalization and subsequent data processing were performed using the GeneSpring GX v11.5.1 software package

(Agilent Technologies). Finally, three samples were hybridized, two biological replicates for each condition (E11.5, E14.5, and E18.5, respectively). Differentially expressed lncRNAs with statistical significance were identified through Volcano Plot filtering. The threshold we used to screen up or downregulated lncRNAs is fold change >2.0 and P-value < 0.05.

GO AND PATHWAY ANALYSIS

Previous studies have shown that mammalian lncRNAs are preferentially located next to genes with developmental functions [Mercer et al., 2009; Cabili et al., 2011]. For each lncRNA locus, the nearest protein-coding neighbor within <100 kb was identified. For antisense overlapping and intronic overlapping lncRNAs, overlapping gene was identified. Pathway analysis and GO analysis were applied to determine the roles of these closest coding genes played in biological pathways or GO terms. GO analysis was applied to analyze the main function of the closest coding genes according to the GO database which provides the key functional classifications for the National Center for Biotechnology Information (NCBI) [Ashburner et al., 2000]. Generally, Fisher′s exact test and v2 test were used to classify the GO category, and the false discovery rate (FDR) was calculated to correct the P-value. Gene networks and canonical pathways representing key genes were identified using the curated ingenuity pathway analysis (IPA) database according to KEGG, Biocarta, and Reatome, as previously described [Han et al., 2006]. We again turned to the Fisher′s exact test and v2 test to select the significant pathway, and the threshold of significance was defined by P-value and FDR.

QUANTITATIVE REAL-TIME PCR

cDNA was synthesized from 1 μg total RNA using the AMV Reverse Transcriptase Kit (Promega, Madison, WI). Real-time PCR was performed using the SYBR green method in an Applied Biosystems 7300 Sequence Detection System (ABI 7300 SDS; Foster City, CA), following the manufacturer′s protocols. The PCR conditions included a denaturation step (95°C for 10 min), followed by 40 cycles of amplification and quantification (95°C for 15 s, 60°C for 1 min). Relative gene expression levels were quantified based on the cycle threshold (Ct) values and normalized to the reference gene glyceraldehyde three-phosphate dehydrogenase (GAPDH). Each sample was measured in triplicate, and the gene expression levels were calculated by the $2^{-\Delta\Delta Ct}$ method. The sequences of the primers used are shown in Table I.

BIOINFORMATICS ANALYSIS

Noncoding regions harbor transcriptional regulatory elements; however, it can be challenging to distinguish these using only the primary sequences as a guide. This protocol describes how to use maps of various epigenetic phenomena to aid in the identification of noncoding regulatory elements. Trimethylation of lysine four (H3K4me3) tends to mark promoters, whereas mono-methylation of the same lysine subunit (H3K4me1) tends to mark enhancers [Barski et al., 2007]. Acetylation of lysine 27 (H3K27ac) also associates with enhancers; however, it appears to have some specificity for those that are active rather than those that are merely "poised" for activity [Creyghton et al., 2010]. TF occupancy of a particular sequence can be another useful indicator of regulatory function, especially in combination with the aforementioned marks [Mortlock and Pregizer, 2012]. Typically, TF binding sites are less broad and can thus be more helpful in precisely mapping the boundaries of a regulatory element. In addition, regulatory sequences are often hypersensitive to DNase treatment when applied to native chromatin. The ENCODE Integrated Regulatory track on the UCSC genome browser makes it easy to scan for all such regions described above. The methods used to gather data for this track, and much more, are described in an extensive publication by the ENCODE Project Consortium (2011).

TABLE I. The Primers Used in This Study

Gene name	Forward (5′–3′)	Reverse (5′–3′)
AK020289	CAGAAGGCTGTCAGATGG	CTGCGGCTGTAAGTAAGA
AK157663	TGGGCGGGCTTTCTATCA	AAACGGTTGCTCCCTTGC
ENS129245	GAGGGACAAGCGACAAAG	AGGGCTCAGACTCAAGACAC
AK053631	ATACTAGGATTGCCTGAG	TGGTGCCTTTAGACTTAC
AK081087	GTATGGCAGGGCACAGGA	CCCACGGACTTAGACAAC
AK050713	TTCAGAGCTTTGTGGCTTCT	GTGTTGGCTACCCGACCT
AK045554	GGTCCTGCCTTCTGGTTC	ATGCCTTGTCTTAGTCTGGTTT
AK138404	GCTTGGCTTCAGGTCATC	ACCGCCGAGTGAGTATGA
BC024929	CAGCAACTGACGGAAATG	TCTTTGGTTGGTGGCTCA
AK008015	CACTGGCAAGGCTGTTTC	CCTGGCCGGTGATCTTAT
AK029733	AGAAGCAATCGGAGTGAG	CAACCAGTCCCAAGTAGC
AK043920	CGTAATCGTCGCACAGAG	AGACATTGAGATCGTGGTAAA
AK052505	TGAGGACTCGCAGGACAC	CCCACGAAACTGAGGAAA
AK137959	ATTATTACAGCCTGGTCC	CTTCGTGTTCTTCCTCCT
AK021352	ATCCAGATTGGGACCTCA	GGCTACGCCTTCTCATAC
BC049716	TACGGCGAACTGGAGATG	TCCCAAATGCTGGACAAC
AK085135	TTACTCACCTCTACCCTCA	AACTAAACCCAAACCACTC
AK156749	GCAGGAGAAGCTAAGAGT	TATGGTGTCCAGGAAGTG
AK013988	TGTTCTGATCCTCCACCA	ACACTGAGATTCGCTTCC
DT903035	TATTAACTGAAGTAAGCCAAGC	AAAGGGATTCCCGTGTAG

RESULTS

HISTOLOGY

A series of hearts were collected at each time point for histological analysis using H&E (Fig. 1). The typical features of the developing heart were observed at each time point: in the E11.5 group, the endocardial, myocardial, and epicardial layers had matured; in the E14.5 group, the myocardium had completed development; in the E18.5 group, the endocardial cushions appeared to fuse, and the sepals, aorta, and tracheal were also visible.

PROFILE OF MICROARRAY DATA

Arraystar Mouse LncRNA Microarray v2.0 is designed for the global profiling of mouse lncRNAs and protein-coding transcripts. 31,423

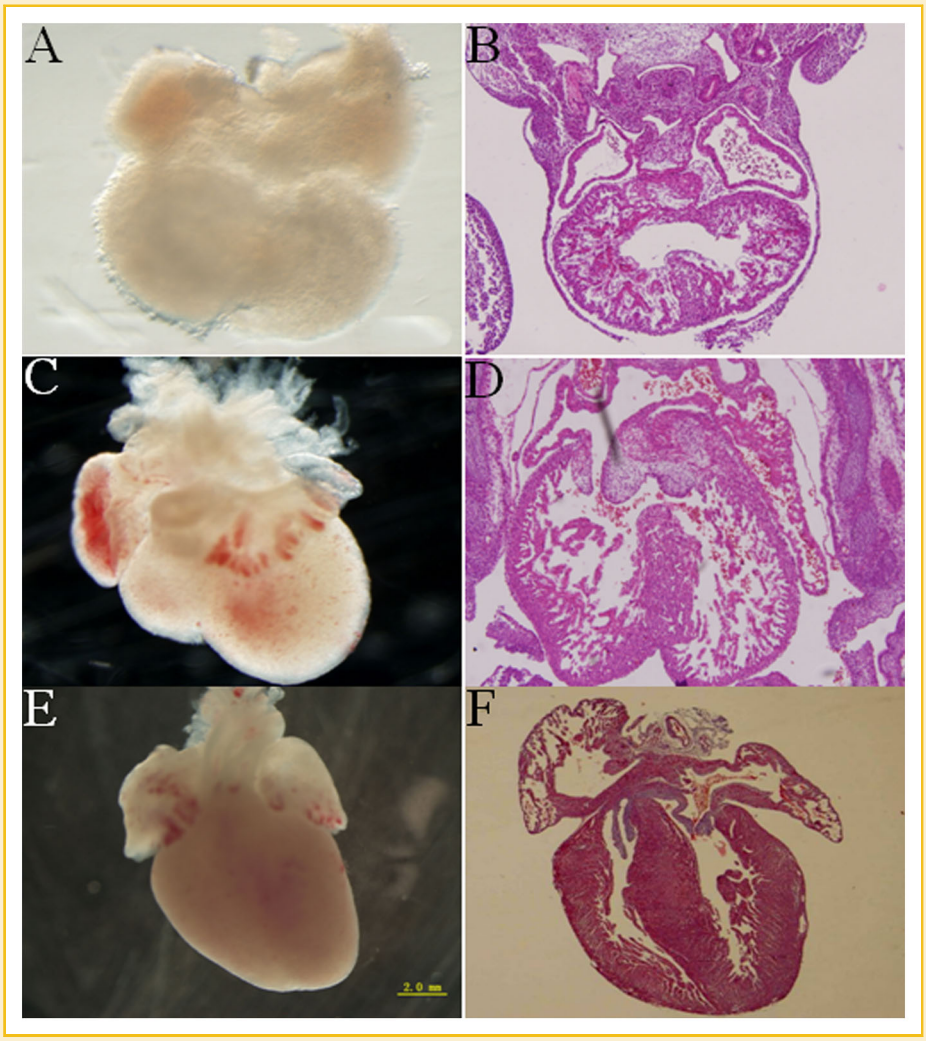

Fig. 1. Morphological stages of heart development at the three time points used for lncRNA profiling. Images of the whole heart (A, C, and E) and corresponding tissue sections stained with H&E (B, D, and F) are shown for each time point studied (E11.5, E14.5, and E18.5, from top to bottom).

lncRNAs and 25,376 coding transcripts can be detected by the second-generation lncRNA microarray. The lncRNAs are carefully collected from the most authoritative databases such as RefSeq, UCSC Knowngenes, Ensembl and many related literatures (Fig. 2A). Hierarchical clustering was performed to show the distinguishable lncRNAs and mRNAs expression pattern among samples (Fig. 2B). The scatterplot is a visualization that is useful for assessing the variation (or reproducibility) between chips (Fig. 2C). The lncRNAs in this microarray are mainly between 200 and 3,000 bp in length (Fig. 2D).

MICROARRAY VALIDATION BY qRT-PCR

We set a threshold as fold change >2.0 and found that 454 lncRNAs were upregulated in early development and then downregulated in later development; 741 lncRNAs were downregulated in early development and then upregulated in later development; 16 lncRNAs were consistently upregulated and 26 lncRNAs were consistently downregulated (Table II). To validate microarray analysis findings, we randomly selected 20 lncRNAs from the differentially expressed lncRNAs with fold change >3 and analyzed their expression by real-time PCR in expanded heart samples. Our data confirmed to be consistent with the microarray results (Fig. 3).

GO AND PATHWAY ANALYSIS

The GO project (http://www.geneontology.org) is a collaborative effort to construct and use ontologies to facilitate the biologically meaningful annotation of genes and their products in a wide variety of organisms and is the key functional classification system of NCBI. In our survey of existing data, when comparing E14.5 to E11.5, the associated gene functions of upregulated lncRNAs mainly involved in (1) cell differentiation; (2) positive regulation of cellular process; (3) regulation of cell cycle; (4) organ development; (5) tissue development; and (6) muscle structure development. The associated gene functions of downregulated lncRNAs mainly involved in (1) cellular process; (2) regulation of cell fate commitment; (3) regulation of cellular metabolic process; (4) cellular metabolic process; (5) cellular

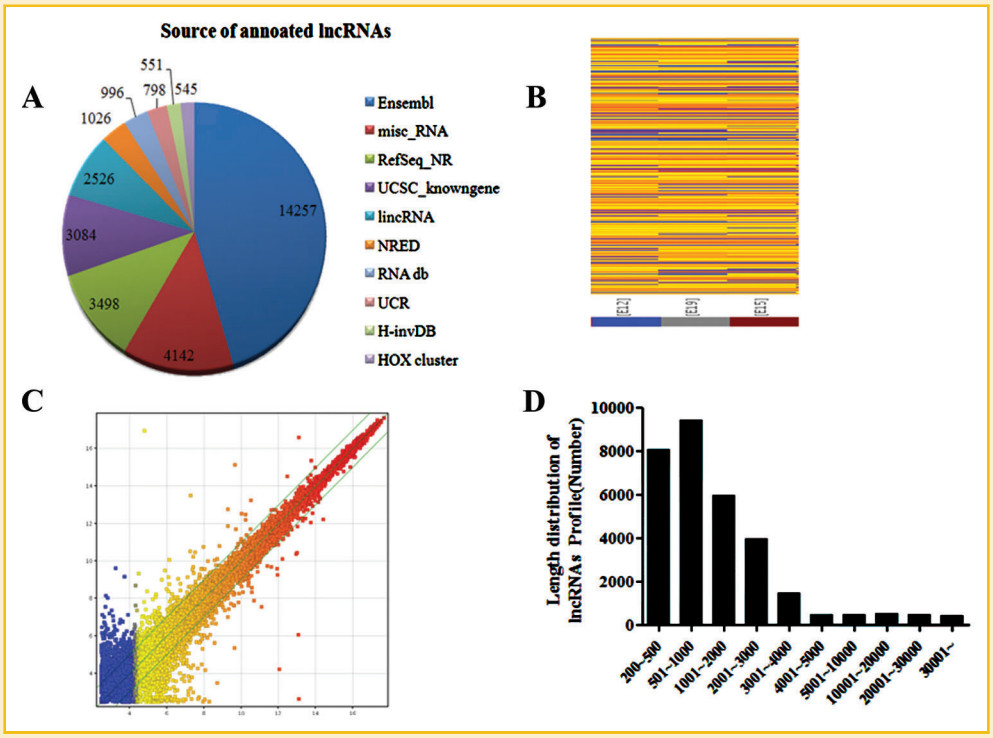

Fig. 2. Profile of microarray. A: Microarray v2.0 recovered the vast majority of expressed RefSeq transcripts; 31,423 lncRNAs and 25,376 coding transcripts can be detected using this microarray. The lncRNAs are carefully collected from the most authoritative databases such as RefSeq, UCSC Knowngenes, Ensembl and related literature. B: Hierarchical clustering was performed to show the distinguishable lncRNAs and mRNA expression pattern among samples. C: The scatterplot is a visualization of the variation (or reproducibility) between chips. D: The length distribution of lncRNAs in microarray 2.0.

macromolecule metabolic process; and (6) primary metabolic process (Fig. 4A). Meanwhile, when comparing E18.5 to E14.5, the associated gene functions of upregulated lncRNAs mainly involved in (1) cellular process; (2) developmental process; (3) metabolic process; (4) regulation of gene expression; (5) gene expression; and (6) organ morphogenesis. The associated gene functions of downregulated lncRNAs mainly involved in (1) multicellular organismal process; (2) positive regulation of MAP kinase activity; (3) cell–cell signaling; (4) system process; (5) muscle tissue development; and (6) localization (Fig. 4B).

IPA was used to identify pathways and gene net works represented among the sets of protein-coding mRNAs identified in the VSD gene expression signature. In our survey of existing data, when comparing E14.5 to E11.5, the associated genes of upregulated lncRNAs mainly involved the following pathways: (1) mTOR signaling pathway; (2) dilated cardiomyopathy; (3) hypertrophic cardiomyopathy; (4) hepatitis C; (5) non-homologous end-joining; and (6) cell adhesion molecules. The associated genes of downregulated lncRNAs mainly involved (1) pentose phosphate pathway; (2) hedgehog-signaling pathway; (3) apoptosis; (4) glycerophospholipid metabolism; (5) pentose phosphate pathway; and (6) valine, leucine, and isoleucine degradation (Fig. 5A). While comparing E18.5 to E14.5, the associated genes of upregulated lncRNAs mainly involved the following pathways: (1) dilated cardiomyopathy; (2) cell adhesion molecules; (3) Wnt-signaling pathway; (4) calcium-signaling pathway; (5) hedgehog-signaling pathway; and (6) GnRH-signaling pathway. The associated genes of downregulated lncRNAs mainly involved (1) calcium-signaling pathway; (2) acute myeloid leukemia; (3) dilated cardiomyopathy; (4) glycosaminoglycan biosynthesis; (5) pentose phosphate pathway; and (6) apoptosis (Fig. 5B).

BIOINFORMATICS ANALYSIS OF AK011347

AK011347 is a bidirectional lncRNA, transcribed from 294 bp downstream of the Map3k7 gene (Fig. 6A). The integrated regulation track is actually several separate tracks, collectively referred to as a "super-track." Thus, the "Layered H3K4me1, H3K4me3" tracks summarize covalent histone modification mapping, the "TF binding sited by ChIP" track summarizes TF mapping, and the "DNase Clusters" track summarizes DNase hypersensitivity mapping (Fig. 6B).

DISCUSSION

Heart function requires sophisticated regulatory networks to orchestrate organ development, physiological responses, and environmental adaptation. Until recently, it was thought that these regulatory networks were composed solely of protein-mediated transcriptional control and signaling systems. However, it is becoming evident that RNA, long considered functioning primarily as the platform for protein production, may in fact play a major role in

TABLE II. Differentially Expressed lncRNAs in the Developing Heart

Regulation	Top 10 lncRNAs	Chromosomal localization	RNA length	Start locus	Stop locus	Associated gene name	Relationship
14.5 vs. E11.5 up;	DT903035	chr1	555	145668786	145669341	Uchl5	Intergenic
E18.5 vs. E14.5 up	BC049716	chr6	836	127726100	127728683	Prmt8	Intergenic
	ENSMUST00000119471	chr4	1,293	118614664	118615957	Olfr1328	Intergenic
	ENSMUST00000118140	chrX	502	91233656	91234158	Zfx	Intergenic
	AK085135	chr7	2508	149644122	149646629	Tnni2	Intergenic
	AK013988	chr6	1,679	56834421	56836101	Nt5c3	Antisense overlap
	uc008hzy.1	chr19	2,158	57191265	57193423	Ablim1	Sense overlap
	uc008mey.1	chr2	2,364	127028302	127033990	Snrnp200	Bidirectional
	MM9LINCRNAEXON10678	chr4	364	88584091	88584455	Mir31	Intergenic
	NR_029457	chrX	2,380	166412974	166416849	Mid1	Antisense overlap
E14.5 vs. E11.5 down;	AK045554	chr10	2,022	101577076	101579098	Mgat4c	Sense overlap
E18.5 vs. E14.5 down	AK050713	chr12	2,122	110907078	110909200	Rian	Intergenic
	AK032574	chr6	2,723	80464587	80467308	Lrrtm4	Intergenic
	uc007xfq.1	chr15	5,217	88985978	88997196	Plxnb2	Sense overlap
	BC024929	chr1	1,078	9629335	9630663	Mybl1	Intergenic
	AK008015	chr9	364	60181948	60182775	Thsd4	Antisense overlap
	AK019733	chr10	355	36983199	36983554	Marcks	Intergenic
	ENSMUST00000140148	chr5	1,353	22056254	22207317	Orc5	Bidirectional
	Gm16133	chr13	563	63434444	63441267	Fancc	Antisense overlap
	AK033527	chrX	2,534	131067485	131070020	Taf7l	Intergenic
E14.5 vs. E11.5 up;	uc008pdt.1	chr3	1,066	51068790	51144562	Elf2	Sense overlap
E18.5 vs. E14.5 down	AK029733	chr9	2,974	31085605	31088577	Prdm10	Intergenic
	AK052505	chr12	1,100	103150876	103151976	Trip11	Sense overlap
	AK043920	chr12	2,414	17182362	17184770	Kcnf1	Antisense overlap
	uc007mnp.1	chr11	515	117630419	117631858	Tmc6	Sense overlap
	AK137959	chr16	2,288	56191202	56193464	Slc9a10	Intergenic
	Gm12316	chr11	406	70424713	70426713	Mink1	Antisense overlap
	AK037866	chr1	2,904	165856589	165859494	Scyl3	Bidirectional
	AK077024	chr9	2,768	103459419	103468221	Tmem108	Antisense overlap
	ENSMUST00000167330	chr11	1,457	115476775	115488021	Slc25a19	Antisense overlap
E14.5 vs. E11.5 down;	AK083183	chr9	2,759	95464522	95467281	Paqr9	Intergenic
E18.5 vs. E14.5 up	Gm14508	chr5	901	116459324	116466114	Prkab1	Antisense overlap
	AK020289	chr11	896	22760360	22761256	B3gnt2	Bidirectional
	uc009qrj.1	chr9	3,917	74800788	74879892	BC031353	Sense overlap
	Gm12940	chr4	574	126712693	126714257	Zmym1	Intergenic
	uc008dwu.1	chr17	963	95234244	95235217	Rik	Bidirectional
	AK020577	chr13	765	108642170	108642935	Zswim6	Sense overlap
	Gm16156	chr8	713	108458232	108469743	Slc12a4	Antisense overlap
	BC037032	chr15	3,300	3971688	3977405	Oxct1	Antisense overlap
	AK032255	chr9	1,041	105371436	105372477	Atp2c1	Sense overlap

most, if not all, aspects of gene regulation, especially the epigenetic processes that underpin organogenesis. These include not only well-validated classes of regulatory RNAs, such as microRNAs, but also tens of thousands of lncRNAs that are differentially expressed across the entire genome of humans and other animals. In the present study, we have characterized the lncRNAs expression profile in the developing mouse heart from E11.5 to E18.5 using microarray analysis. After confirmation of microarray by RT-PCR, significant differences in lncRNA expression profiles were observed in cardiac tissues among three key development time points analyzed (E11.5, E14.5, and E18.5), thus suggesting numerous lncRNAs were involved in heart development.

The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism [Ashburner et al., 2000]. In our survey of existing data, the main involved biology process of dysregulated lncRNAs included many closely connected to heart development, such as "regulation of cell cycle," "cell differentiation," "regulation of cell fate commitment," "organ development," and "muscle structure development". However, perhaps the most important challenge today is that the knowledge embedded in pathways regarding how various genes interact with each other is not currently exploited. Microarray technology makes it possible to measure the expression levels of almost all the coding genes and therefore facilitate the identification of genes and pathways that are related to disease initiation and development. Based on our data, the associated genes of differentially expressed lncRNAs in heart development mainly involved some pathways that play important roles in DNA damage-repair, energy metabolism, and apoptosis. For instance, Wnt-signaling pathway has been recognized for its function in embryonic development. The embryonic processes it controls include body axis patterning, cell fate specification, cell proliferation, and cell migration. These processes are necessary for proper formation of important tissues including bone, heart, and muscle [Cadigan and Nusse, 1997].

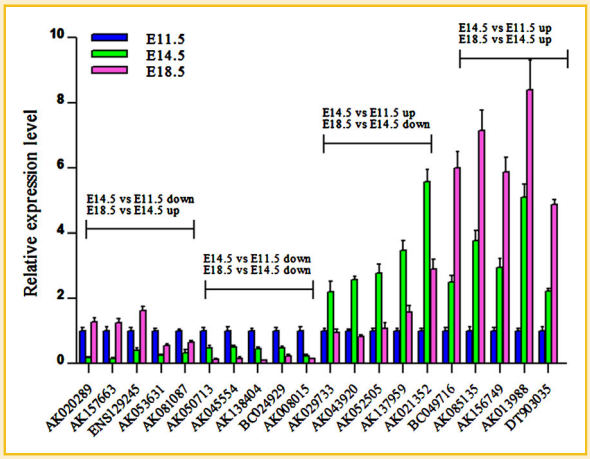

Fig. 3. Twenty lncRNAs were randomly selected and confirmed by qRT-PCR. Our qRT-PCR data are consistent with the microarray results. Data are expressed as the mean ± standard deviation (n = 6) of three independent experiments.

LncRNAs may be expressed preferably during particular developmental time points, or within certain tissues. Moreover, within a developmental framework, lncRNAs may exhibit dynamic expression patterns [Mercer et al., 2009; Cabili et al., 2011]. A major focus of our study was to define the repertoire of lncRNAs expressed at different time points of heart development. In terms of broad classes with expression that changed during development, four major expression profiles were identified: lncRNAs that were upregulated in early development and then downregulated in later development; lncRNAs that were downregulated in early development and then upregulated in later development; lncRNAs that were consistently upregulated or downregulated in heart development. Our comparative clustering analyses have shown that there were numerous lncRNAs were differentially expressed at different time points. These differential expression patterns may indicate their involvement in heart development.

The language used by lncRNAs to interact with network components is still largely elusive. Unlike the well-studied miRNAs, lncRNAs do not seem to function via a common pathway; therefore, no predictions can be made about their function based on their primary sequence or secondary structure. A major challenge lies in decoding the functional elements and modules in the primary sequence of noncoding genes, including structural motifs and regulatory elements that define their roles [Guil and Esteller, 2012]. Currently, there are no features of either the genome or epigenome that can be used to unequivocally identify regulatory elements. Nevertheless, some features, such as DNase hypersensitivity, TF occupancy, and histone modifications, seem to be more reliable indicators of regulatory function than others [Mortlock and Pregizer, 2012]. Based on our analysis, for AK011347, we identified a region of overlap between the various tracks, which indicated that it may be function as regulatory element. Furthermore, the associated gene of AK011347 is Map3k7. The protein encoded by this gene is a member of the serine/threonine protein kinase family. This kinase mediates the signaling transduction induced by TGF-beta and morphogenetic protein, and controls a variety of cell functions including transcription regulation and apoptosis [Choi et al., 2012]. Therefore, AK011347 may be involved in heart development through the target gene Map3k7. Nevertheless, identification of its function requires future loss-of-function or gain-of-function analysis.

In conclusion, our study provides an expression profile of lncRNAs in the developing heart, as well as a series of differentially expressed miRNAs between key developmental time points. We believe that these lncRNAs likely play an important role in heart development, thus additional studies may clarify the mechanism(s) of normal heart development, and provide a physiological basis for future investigations on congenital heart disease.

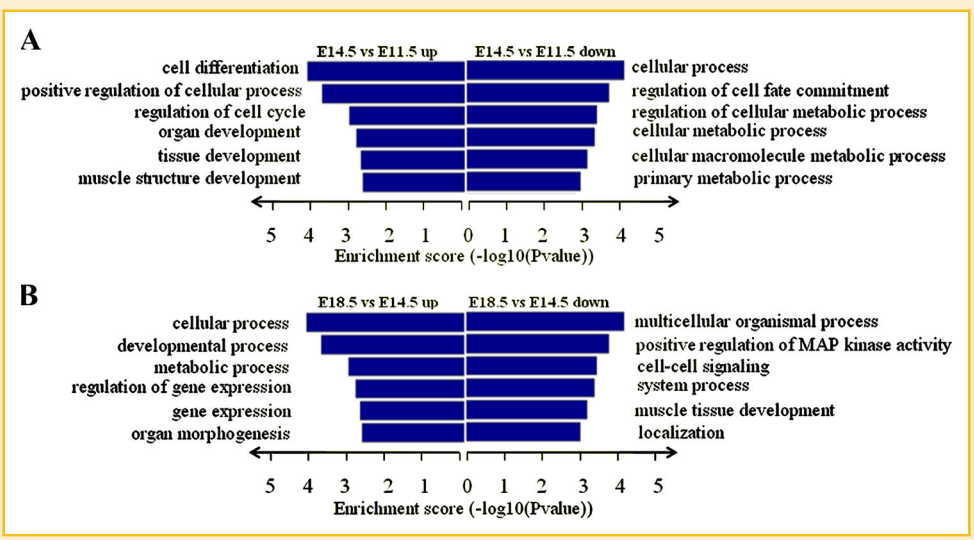

Fig. 4. GO analysis. The first six Go ID that exhibited significant differences between E14.5 and E11.5 are listed in A; the first six Go ID that exhibited significant differences between E18.5 and E14.5 are listed in B (left and right panels show the associated coding genes of upregulated and downregulated lncRNAs, respectively).

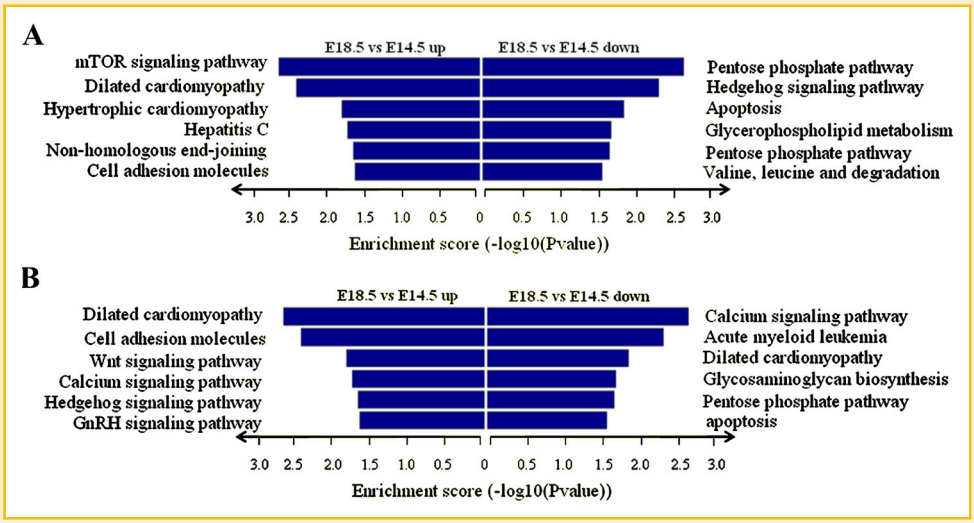

Fig. 5. Pathway analysis. The first five pathway that exhibited significant differences between E14.5 and E11.5 are listed in A; the first five pathways that exhibited significant differences between E18.5 and E14.5 are listed in B (left and right panels show the associated coding genes of upregulated and downregulated lncRNAs, respectively).

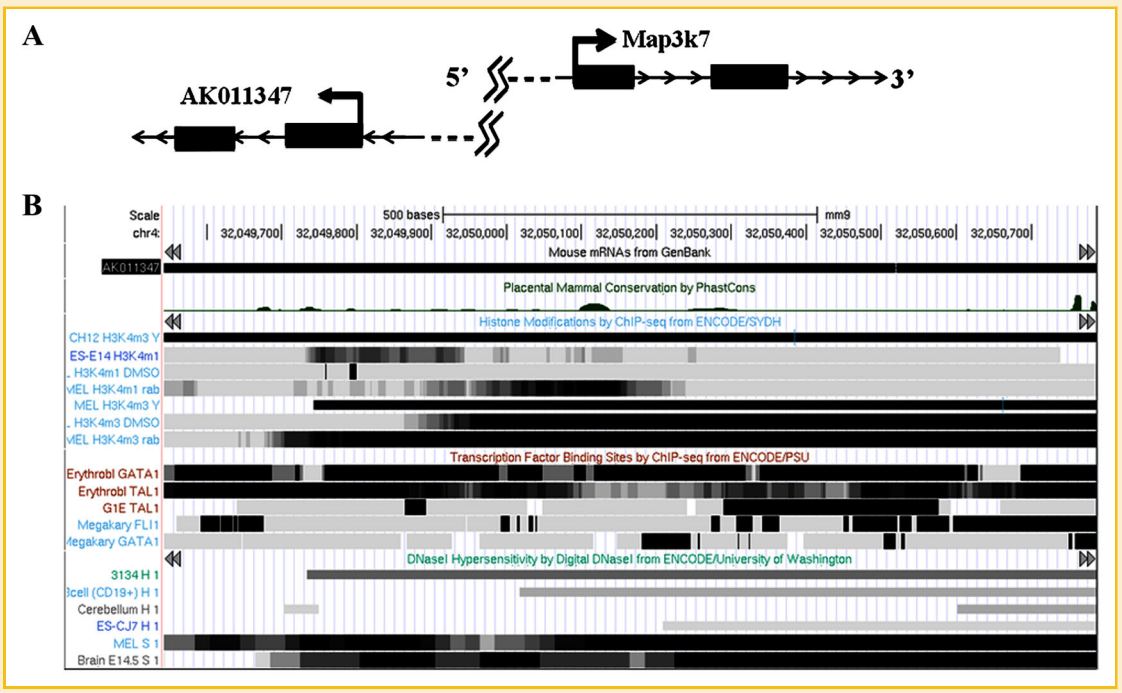

Fig. 6. Bioinformatics analysis of AK011347. A: AK011347 is a bidirectional lncRNA, transcribed from 294 bp downstream of the Map3k7 gene. B: Several tracks of interest, including conservation, histone markings, DNase hypersensitivity, and TFBS are displayed. Integrated regulation track data for a region spanning the AK011347 are shown.

REFERENCES

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene Ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129:823–837.

Beltran M, Puig I, Pena C, Garcia JM, Alvarez AB, Pena R, Bonilla F, de Herreros AG. 2008. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev 22:756–769.

Bruneau BG, Nemer G, Schmitt JP, Charron F, Robitaille L, Caron S, Conner DA, Gessler M, Nemer M, Seidman CE, Seidman JG. 2001. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. Cell 106:709–721.

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25:1915–1927.

Cadigan KM, Nusse R. 1997. Wnt signaling: A common theme in animal development. Genes Dev 11:3286–3305.

Choi ME, Ding Y, Kim SI. 2012. TGF-beta signaling via TAK1 pathway: Role in kidney fibrosis. Semin Nephrol 32:244–252.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci USA 107:21931–21936.

Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, Finch CE, St Laurent G, III, Kenny PJ, Wahlestedt C. 2008. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. Nat Med 14:723–730.

Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD. 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev 20:1470–1484.

Guil S, Esteller M. 2012. Cis-acting noncoding RNAs: Friends and foes. Nat Struct Mol Biol 19:1068–1075.

Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY. 2010. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 464:1071–1076.

Han T, Wang J, Tong W, Moore MM, Fuscoe JC, Chen T. 2006. Microarray analysis distinguishes differential gene expression patterns from large and small colony thymidine kinase mutants of L5178Y mouse lymphoma cells. BMC Bioinformatics 7(Suppl2):S9.

Lints TJ, Parsons LM, Hartley L, Lyons I, Harvey RP. 1993. Nkx-2.5: A novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. Development 119:419–431.

Liu N, Olson EN. 2010. MicroRNA regulatory networks in cardiovascular development. Dev Cell 18:510–525.

Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A. 2007. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature 445:666–670.

Mattick JS. 2009. The genetic signatures of noncoding RNAs. PLoS Genet 5: e1000459.

Mercer TR, Dinger ME, Mattick JS. 2009. Long non-coding RNAs: Insights into functions. Nat Rev Genet 10:155–159.

Mortlock DP, Pregizer S. 2012. Identifying functional annotation for noncoding genomic sequences. Curr Protoc Hum Genet. Chapter 1:Unit 1.10.

Ponting CP, Oliver PL, Reik W. 2009. Evolution and functions of long noncoding RNAs. Cell 136:629–641.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129:1311–1323.

Savolainen SM, Foley JF, Elmore SA. 2009. Histology atlas of the developing mouse heart with emphasis on E11.5 to E18.5. Toxicol Pathol 37:395–414.

Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, Blencowe BJ, Prasanth SG, Prasanth KV. 2010. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell 39:925–938.

Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. Science 329:689–693.

Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, Tempst P, Rosenfeld MG, Glass CK, Kurokawa R. 2008. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature 454: 126–130.

Watt AJ, Battle MA, Li J, Duncan SA. 2004. GATA4 is essential for formation of the proepicardium and regulates cardiogenesis. Proc Natl Acad Sci USA 101:12573–12578.

Wessels A, Sedmera D. 2003. Developmental anatomy of the heart: A tale of mice and man. Physiol Genomics 15:165–176.

Wilczynski B, Furlong SA. 2010. Challenges for modeling global gene regulatory networks during development: Insights from *Drosophila*. Dev Biol 340:161–169.

Wilusz JE, Sunwoo H, Spector DL. 2009. Long noncoding RNAs: Functional surprises from the RNA world. Genes Dev 23:1494–1504.

Yan B, Wang Z. 2012. Long noncoding RNA: Its physiological and pathological roles. DNA Cell Biol 31(Suppl1):S34–S41.