The Long Noncoding RNA Landscape of the Ischemic Human Left Ventricle

The Harvard community has made this article openly available. **Please share** how this access benefits you. Your story matters

Citation	Saddic, Louis A., Martin I. Sigurdsson, Tzuu-Wang Chang, Erica Mazaika, Mahyar Heydarpour, Stanton K. Shernan, Christine E. Seidman, et al. 2017. "The Long Noncoding RNA Landscape of the Ischemic Human Left Ventricle." Circulation: Cardiovascular Genetics 10 (1) (January): e001534. doi:10.1161/circgenetics.116.001534.
Published Version	doi:10.1161/CIRCGENETICS.116.001534
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:34298863
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

The Long Noncoding RNA Landscape of the Ischemic Human Left Ventricle

Louis A. Saddic, MD, PhD[1], **Martin I. Sigurdsson, MD, PhD**[1], **Tzuu-Wang Chang, PhD**[1], **Erica Mazaika, BS, MS**[2], **Mahyar Heydarpour, PhD**[1], **Stanton K. Shernan, MD**[1], **Christine E. Seidman, MD**[2], **Jon G. Seidman, PhD**[2], **Sary F. Aranki, MD**[3], **Simon C. Body, MBChB, MPH**[1], and **Jochen D. Muehlschlegel, MD, MMSc**[1]

[1]Department of Anesthesiology, Perioperative and Pain Medicine, Boston, MA

[2]Department of Genetics, Harvard Medical School, Boston, MA

[3]Division of Cardiac Surgery, Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

Abstract

Background—The discovery of functional classes of long noncoding RNAs (lncRNAs) has expanded our understanding of the variety of RNA species that exist in cells. In the heart, lncRNAs have been implicated in the regulation of development, ischemic and dilated cardiomyopathy, and myocardial infarction. Nevertheless, there is a limited description of expression profiles for these transcripts in human subjects.

Methods and Results—We obtained left ventricular tissue from human patients undergoing cardiac surgery and used RNA sequencing to describe a lncRNA profile. We then identified a list of lncRNAs that were differentially expressed between pairs of samples before and after the ischemic insult of cardiopulmonary bypass. The expression of some of these lncRNAs correlates with ischemic time. Coding genes in close proximity to differentially expressed lncRNAs as well as coding genes that have coordinated expression with these lncRNAs are enriched in functional categories related to myocardial infarction including: heart function, metabolism, the stress response, and the immune system.

Conclusions—We describe a list of lncRNAs that are differentially expressed after ischemia in the human heart. These genes are predicted to function in pathways consistent with myocardial injury. As a result, lncRNAs may serve as novel diagnostic and therapeutic targets for ischemic heart disease.

Clinical Trial Registration—https://clinicaltrials.gov/; Unique Identifier: NCT00985049.

Keywords

myocardial ischemia; long noncoding RNA; RNA sequencing

Correspondence: Jochen D. Muehlschlegel MD, MMSc, Brigham and Women's Hospital, CWN L1, Department of Anesthesiology, 75 Francis St., Boston, MA 02115, Tel: (617)-525-8156, Fax: (617)-730-2813, jmuehlschlegel@partners.org.

Journal Subject Terms

Gene Expression and Regulation; Cardiovascular Surgery; Ischemia; Myocardial Infarction

Introduction

Next generation sequencing has expanded our knowledge of the diversity of RNA transcribed by the human genome. We now know that the dogma of genetic information that flows from DNA to RNA and then protein is much more complex than originally envisioned. These new technologies have enabled us to uncover many different categories of RNA transcripts, and most of them are not translated into proteins. These noncoding RNAs are generally classified into either small RNAs, including microRNAs, or long noncoding RNAs (lncRNAs).[1] LncRNAs are RNA transcripts >200 nucleotides in length that have low potential to encode proteins. To date, there are over 50,000 identified lncRNAs in the human genome with some variation based on different annotations. They can be found throughout the genome including within introns of coding genes, antisense transcripts of coding genes, overlapping exons of coding genes or their promoters, or between genes (so called long intergenic RNAs or lincRNAs).[2] Unfortunately, little is known about the function of these transcripts, in part because of the infancy of the field, but also due to the lack of conservation among lncRNAs, the complexity of their isoforms, and limited insight into shared motifs. Nevertheless, studies that have examined the pattern of lncRNA transcription and experiments that have characterized a select number of lncRNAs suggest many potential biological roles. The position of many lncRNAs in promoter and enhancer regions makes them well poised to control transcriptional activation and repression.[3] Some lncRNAs serve as scaffolds for chromatin modifying factors and thus regulate epigenetics. Other lncRNAs, particularly those anti-sense lncRNAs, have been shown to regulate the expression of coding genes by directly binding to components of mRNAs and/or miRNAs.[4]

Differential and tissue specific expression of lncRNAs has been documented for many natural processes such as development[5] and physiology[6] along with several pathological conditions such as cardiovascular disease.[2] Some groups have described the changes in lncRNAs that occur following myocardial infarction in animal models. For example, in a study of acute myocardial ischemia in mice, novel heart specific lncRNAs were identified that are associated with physiological characteristics of heart function and likely mediate the post injury stress response.[7] Human studies, on the other hand, have been limited. One group identified differentially expressed lncRNAs between hearts with either ischemic or non-ischemic cardiomyopathy compared to healthy donors and demonstrated that the expression of some of these lncRNAs returned to levels consistent with healthy hearts following administration of mechanical support.[8] Even more, some studies have also looked at lncRNAs in circulating mononuclear cells[9] and in the plasma[10] of patients after myocardial infarction. Nevertheless, to date there are no studies that have specifically examined the lncRNA profile of acute ischemia in the human heart.

We hypothesized that ischemia alters the expression of lncRNAs in the human heart. During surgery requiring cardiopulmonary bypass (CPB), the heart undergoes many changes similar

to acute myocardial ischemia.[11] We obtained left ventricular (LV) biopsies from a large group of cardiac surgical patients before and after CPB. RNA from these paired samples was sequenced and annotated with known transcripts. We were able to identify a population of differentially expressed lncRNAs and investigated putative functions. In so doing we describe the first lncRNA profile of acute ischemia in the human LV.

Methods

Patients and tissue samples

We prospectively enrolled eighty-five patients undergoing non-emergent aortic valve replacement surgery with CPB (Table 1). Procedures were performed in accordance with the ethical standards of The Partners HealthCare Institutional Review Board which approved this study and written informed consent was obtained from each patient. Punch biopsies (~3–5μg total RNA content) were taken from the LV apex at two time points; immediately after initiation of CPB at the time of routine placement of a surgical vent (pre-ischemic) and after a median of 74 minutes (inter-quartile range 61–93min) (post-ischemia), during which time the heart was arrested with cold blood cardioplegia for myocardial protection.

RNA sequencing

Tissue samples were immediately placed in RNAlater® (Ambion, Life Technologies, USA) at +4C for 48 hours and then frozen at −80C until RNA extraction. Total RNA was isolated with Trizol and RNA quality was assessed using the Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA). Ribosomal RNA was removed by performing 1–2 washings of RNA annealed to poly-T oligo beads (Invitrogen, Life technologies, Grand Island, NY). RNAs were fragmented and then reverse transcribed using random hexamers (Invitrogen). Double-stranded DNA (dsDNA) synthesis was performed using Pol I and RNA-ase H. Short fragments were purified with QiaQuick PCR extraction kit (Qiagen, Hilden, Germany) and resolved with EB buffer for end reparation and poly(A) addition followed by ligation with sequencing adaptors for cluster generation and non-strand tagged sequencing on the Illumina HiSeq 2000 (Illumina, San Diego, CA). Read length was 100 base pairs. Paired-end sequencing produced on average 82 million reads per sample.

Alignment

Raw reads produced by the Illumina sequencer imaging files were filtered to remove reads containing adaptor sequences, containing >5% unknown nucleotides, or having >50% of reads with base quality scores < 5. Cleaned and trimmed mRNA reads were aligned to the human reference genome (UCSC hg19) using Tophat version 2.0.5 under default settings[12]. BAM files were subsequently sorted by chromosome number and converted to SAM files using SAMtools.[13] Reads were counted with HTSeq-count[14] using the GENCODE release 19[15] annotation file amended to include mitochondrial long noncoding RNAs found in the NONCODE 2016 annotation file.[16] As many lncRNAs overlap with coding regions in both the sense and anti-sense orientation, the "union" default option for HTSeq-count was chosen to avoid miscounting reads from mRNAs as lncRNAs.

Differential expression

Read counts were subjected to paired differential expression analysis using the R[17] package DESeq2.[18] Genes were considered significant if their log2 fold change (post-ischemic/pre-ischemic) was >0.58 or <−0.58, their mean FPKM (fragments per kilobase per million mapped reads) across all samples was greater than 0.01, and their *p*-value adjusted for multiple comparison (Benjamini Hochberg) was <0.01 (Figure S1). The same methodology was used for the subgroup analysis on diabetic and non-diabetic patients with each cohort analyzed separately.

Neighboring coding gene enrichment analysis

All coding genes located within 100 kb (kilobases) of each differentially expressed lncRNA were considered neighboring. Identification of these genes was facilitated by the GREAT database.[19] Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was conducted using the R package clusterProfiler.[20] The *p*-values are corrected for multiple comparison using the Benjamini Hochberg method.

Co-expressed lncRNA-mRNA studies

The correlation between the expression of all differentially expressed lncRNAs and all mRNAs from all 85 paired samples (170 samples total) was determined using Partek® Genomics Suite™ (PGS) version 6.6 (Partek Inc., MO, USA). FPKM (fragments per kilobase per million mapped reads) was used for mRNA and lncRNA expression. These normalized values were obtained by extracting read counts from DESeq2 and gene lengths from the annotation file using the R package GenomicFeatures[21]. Those coding genes with absolute Spearman's rank correlation coefficient greater than 0.7 were subjected to GO and KEGG analysis with clusterProfiler as described above. Variance stabilized gene expression values were obtained for all lncRNA and mRNAs in the annotation file using DESeq2. Fold change (post-ischemic/pre-ischemic) was calculated for all RNAs from all 85 patients. Spearman's rank correlation coefficients for co-expressed lncRNAs and mRNAs, and enrichment terms for mRNAs were then determined using similar methods described above. A density plot was generated in JMP® Version 10 (SAS Institute Inc., Cary, NC) using absolute Spearman's rank correlation coefficients derived from the co-expression of each lncRNA-neighboring coding gene pair compared to correlations from the co-expression of randomly selected lncRNA-mRNA pairs. Significance was determined using a standard *t*-test.

Transcription factor binding site enrichment analysis

Over-represented transcription factor binding sites was performed on the 2kb region upstream from the transcriptional start site of all differentially expressed lncRNAs using the TRANSFAC® database (www.biobase-international.com/transcription-factor-binding-sites from BIOBASE Corporation).[22] The background sequences were randomly generated and default settings were used for all other parameters; the significance was set at a *p*-value < 0.01. Enrichment of E2F1 binding in HeLa-S3 cells was extracted from publically available peak calls through the ENCODE database.[23]

Droplet Digital polymerase chain reaction (ddPCR™)

Validation of selected lncRNA and mRNA transcripts was performed on 12 pairs of pre-ischemic and post-ischemic samples using the Bio-Rad QX100™ Droplet Digital PCR (ddPCR™) system (Pleasanton, CA).[24] cDNA was generated from patient samples as described above. PCR reactions were constructed using custom probe and primer sequences (Table S1) and absolute concentration of transcripts was obtained according to the manufacturer's specifications. Relative expression was calculated using the house keeping gene *tumor protein, translationally-controlled 1* (*TPT1*). Statistical significance was determined using a paired *t*-test. Stripcharts and statistics were generated using the R package ggplots2.[25]

Results

The landscape of lncRNA expression before and after ischemia

Our annotation file contained 13,871 lncRNAs. The most highly expressed lncRNAs in our LV samples have previously been identified as highly expressed in other transcriptome profiles of the human LV.[8, 26] Over half of the top 15 abundant lncRNAs are located between coding genes and include well known lncRNAs such as *MALAT1, H19, NEAT1*, and *DANCR* (Table 2). *H19*, which has been shown to play a role in cardiac necrosis[27], met statistical significance for up-regulation following ischemia in our samples albeit with a low fold change (log2 fold change 0.33).

Differential expression analysis between pre- and post-ischemic samples was performed using a paired analysis. 128 lncRNAs with log2 fold change >0.58 or <−0.58 (which corresponds to an absolute fold change >1.5) and adjusted *p*-value < 0.01 were designated as significant (Table S2). Of these, 97 genes were up-regulated and 31 genes were down-regulated after ischemia (Figure 1A). Correcting for ischemic time yielded similar results (data not shown). LncRNAs are classified in the GENCODE annotation file according to their location relative to coding genes. This classification includes: antisense transcripts (long noncoding RNAs that overlap coding genes but are expressed from the opposite strand), intronic (lncRNAs that overlap with intronic sequences of coding genes) sense overlapping (lncRNAs that at least in part overlap with exons on the same strand as a coding gene) and lincRNAs (located between coding genes). The majority of the differentially expressed lncRNAs identified in this study were either antisense (44%) or lincRNAs (43%) (Figure 1B). Furthermore, these lncRNAs were typically smaller in size (greater than 58% were less than 1kb in length) and expressed at lower levels (70% had a mean FPKM across all samples of <0.1) (Figure 1B) as compared to their coding gene counterparts. This is also consistent with other profiles of human and mouse lncRNA expression in the heart.[8, 28]

Interestingly, for some differentially expressed lncRNAs the magnitude of fold change was correlated with the extent of injury. For example, there is a significant correlation between ischemic time and expression log2 fold change among all patients in this study for the upregulated lncRNA *RP11-64B16.4* (Spearman's rank correlation coefficient 0.34 and *p*-value 1.67×10^{-3}). In fact, this relationship was stronger than the correlation between

ischemic time and post-operative day 1 CKMB for all patients (Spearman's rank correlation coefficient 0.15 and *p*-value 0.18) (Figure 1C).

Neighboring coding genes function in the stress and immune response

Since many studies have demonstrated that lncRNAs can be co-expressed, regulate, and share similar function as their neighboring coding genes, we analyzed the gene ontology enrichment of all neighboring coding genes for all differentially expressed lncRNAs. We chose a reference range of 100 kb for this analysis based on reports which have shown a decrease in strength of expression correlations between lncRNAs and nearby mRNAs beyond this range.[29] Within 100 kb, we identified 277 unique coding genes. In the GO category biological processes, these genes were significantly enriched in "hydrogen peroxide metabolic process", "response to stress", "response to stimulus", and "immune system process". In the GO category molecular function, these genes were significantly associated with "serine-type endopeptidase activity" and "protein binding". Finally, we also looked at the KEGG pathway enrichment of neighboring genes and found an association with "tight junctions" and "leukocyte transendothelial migration" (Figure 2A). These findings are consistent with many other enrichment studies on direct mRNA transcriptomes of ischemic hearts in human and animal models.[30–32] Moreover, there is a high correlation between the expression of lncRNAs and their neighboring mRNA pairs (Figure 2B). For example, the expression of the down-regulated lncRNA *RP11-171A24.3* and its neighboring gene *RORB* have a strong positive correlation (Spearman's rank correlation coefficient 0.79) (Figure 2C).

mRNAs co-expressed with ischemic lncRNAs regulate metabolism and heart physiology

Co-expression studies with coding genes may also uncover potential functions of lncRNAs. We thus generated a list of coding genes that were co-regulated with our pool of differentially expressed lncRNAs. Associations were considered significant if the absolute Spearman's rank correlation coefficient was greater than 0.7, a threshold chosen based on similar comparisons made by other groups.[33, 34] Interestingly, only 6 out of the 128 differentially expressed lncRNAs had significant correlations with 215 unique coding genes (Table S3). Even more, all 6 of these lncRNAs were down-regulated with ischemia, they tended to have high levels of expression, and all of their relationships with coding genes were positive (that is no significant correlations were uncovered with a down-regulated lncRNA and an up-regulated mRNA). Several of these co-expressed coding genes have been shown to play a role in myocardial ischemia including *FEM1A*[35] (co-expressed with *RP3-527G5.1*) and *ADIPOQ*[36] (co-expressed with *TRHDE-AS1*). Even more, these two particular coding genes were also significantly down-regulated following ischemia in our study (log2 fold change −0.67 and −0.88 and adjusted *p*-value 1.14×10^{-12} and 4.29×10^{-07}, respectively). Another notable gene is *PHACTR1* which is co-expressed with two lncRNAs, *RP3-527G5.1* and *RP11-371M22.1*. This gene was also significantly down-regulated in our study although to a lesser degree than *FEM1A* and *ADIPOQ* (log2 fold change −0.51 and *p*-value 1.95×10^{-13}). *PHACTR1* harbors an intronic SNP, rs12526453, which has been shown in genome-wide association studies to be linked to myocardial ischemia and coronary artery disease.[37] In rats, *PHACTR1* is also down-regulated following myocardial infarction and

may play a role in regulating the relative expression of actin isoforms in the heart.[38] ddPCR™ was used to confirm the differential expression of these transcripts (Figure 3A–E).

The functional categories of all co-expressed genes was further analyzed with gene ontology analysis. In the GO category biological process, there were several terms linked to metabolism and cardiac function (Figure 4A). In the GO category cellular component there were several terms linked to components of the mitochondria and sarcomere (Figure 4B). In the GO category molecular function, there were several terms linked to nucleotide binding (Figure 4C). Finally, the top KEGG pathways associated with the co-expressed coding genes were "dilated cardiomyopathy", "adrenergic signaling", "carbon metabolism", "insulin secretion", and "citrate cycle" (Figure 4D). Once again these associations have been demonstrated for ischemic mRNA profiles of the heart in many other studies.[30–32, 39]

We repeated these correlation studies by first obtaining variance stabilized gene expression levels for all samples and then computing the fold change for each RNA species between pre- and post-ischemic subjects. Subsequently, we re-ran our correlation matrix where the fold change of each differentially expressed lncRNA is compared to the fold change of each mRNA for all 85 patients. The candidate coding genes and enrichment results were highly similar to the original correlation studies using FPKM calculated abundances from all 170 samples (Table S4).

Transcription factor binding site enrichment of differentially expressed lncRNAs is linked to ischemic injury

Transcription factors play a central role in the changes in gene expression following a potent stimulus such as ischemia. As a result, we sought to determine if there exists an over-representative collection of conserved transcription factor binding sites in the differentially expressed lncRNAs uncovered in this study. To this end, we chose the 2kb region upstream of each lncRNA as the promoter region for this study. There were a total of 30 transcription factors with significantly enriched binding sites compared to a random background sequence. Among the list were E2F and XBP-1 which have been shown in prior studies to play a role in ischemic/reperfusion injury[40, 41], along with Myogenin (MYOG), which plays a critical role in muscle development[42] (Figure S2).

Discussion

In this study we describe the first lncRNA profile of acute ischemia in the human heart. Our list of differentially expressed transcripts offers a unique and essential contribution to previously described profiles in that our study involves a high number of paired left ventricular samples from human subjects immediately following ischemic insult. This design not only enabled us to make cross-species comparisons, but also permitted us to refine candidate lncRNAs that may have bonafide significant functions in the pathology of ischemia in the heart. According to the Genotype-Tissue Expression (GTEx) project and NONCODE database, many of the top differentially expressed lncRNAs demonstrate enriched expression in the heart compared to other tissues including: *RP13-270P17.1, RP11-371M22.1, RP11-517P14.2, RP11-968A15.2*, and *RP11-64B16.4*. Furthermore, the expression of three lncRNAs shown to be strongly correlated with coding genes in our

study, namely *RP11-517P14.2, RP3-527G5.1*, and *RP11-371M22.1*, had the highest expression in the heart compared to other tissues.[16, 43] Also, the list of differentially expressed lncRNAs lacked conservation. Of the 121 differentially expressed lncRNAs identified in the NONCODE database, 79% were conserved in primates, but only 29% were conserved in rodents and 1% in birds. These characteristics fit well with a predicted tissue specific pattern of functional lncRNAs in the heart.[2, 44]

Among the top differentially expressed lncRNAs is the upregulated lncRNA *RP13-270P17.1*. This lncRNA lies between the two myosin light chain genes *MYL12A* and *MYL12B*. It overlaps portions of the *MYL12A* gene on the opposite strand and lies directly upstream of the *MYL12B* transcriptional start site. As a result it is poised to regulate these two muscle specific genes. The expression of this lncRNA was also highly enriched in the heart in the GTEx database.[43] *SNORD3A* had the highest fold change and lowest *p*-value in our candidate gene list. It encodes a C/D box small nucleolar RNA and thus its upregulation following ischemia could play a central role in RNA biogenesis during acute ischemia[45]. Furthermore, *RP11-64B16.4* was also strongly upregulated in our study. The expression of this lncRNA is also highly enriched in heart tissue[43]. Even more, it is an anti-sense lncRNA located within the *HSPB8* gene. *HSPB8* encodes a heat shock protein which has been shown by others to promote potent cardioprotection following ischemia.[46] As a result, *RP11-64B16.4* may have a similar function either through its regulation of *HSPB8* or simply through sharing its transcriptional machinery.

The expression of some lncRNAs, including the upregulated genes *RP11-64B16.4* and *RP13-270P17.1*, had stronger correlations to ischemic time than post-operative day 1 CKMB, a marker of ischemia commonly used in clinical practice. Interestingly, *RP13-270P17.1* was also shown to be up-regulated in failing human right ventricles compared to non-failing donor tissue[28]. Even more, according to the GTEx project, *RP13-270P17.1* is expressed primarily in the heart and testes with very little expression in whole blood.[43] As a result, the presence of *RP13-270P17.1* in blood in the setting of symptoms of acute coronary syndrome could be highly suggestive of myocardial ischemia. These findings are purely observational and not conclusive at this time. Further studies are needed to validate the diagnostic potential of these lncRNAs.

Given the lack of conservation, smaller size, and limited knowledge of functional motifs, it can be more difficult to predict functions of lncRNAs using a bioinformatic approach. Nevertheless, many studies have demonstrated a tendency for lncRNAs to be co-expressed with neighboring coding genes. Potential mechanisms to explain this phenomenon include shared regulatory sequences or feed-forward/feed-back loops in the local environment where the lncRNA directly regulates the expression of its neighbor.[47] Ontology and pathway analysis of neighboring genes from our experiment showed enrichment in categories that resemble profiles of coding genes from many human and animal models of myocardial ischemia.[30–32, 39] This includes such categories as stress, tight junctions, and leukocyte migration. The observation that those lncRNAs that changed acutely during ischemia are located in close proximity to coding genes that define the signature of myocardial infarction, suggests that these lncRNAs could function as early initiators of transcriptional regulation triggered by this insult. This is supported by the correlated expression between many

neighboring coding genes and their lncRNA neighbors. Even more, transcription factor binding site enrichment of these lncRNAs demonstrated an over-representation of transcription factors known to play a role in ischemia/reperfusion injury, including E2F family members and XBP-1. Publically available Chip-seq data from the ENCODE project for E2F1 binding in HeLa-S3 cells demonstrated peaks in the promoter region of 3 differentially expressed lncRNAs: *RP11-968A15.2, RP13-895J2.7*, and *RP11-517P14.2*.[23] Of note, expression of the lincRNA *RP11-517P14.2* is highly correlated with coding genes in our study. *RP11-517P14.2* was also identified as being up-regulated in human hearts with ischemic cardiomyopathy compared to healthy donor hearts.[8]

We also investigated the pool of coding genes in our samples that were co-expressed with all differentially expressed lncRNA. Ontology and pathway analysis of these highly correlated coding genes were similarly enriched in categories linked to myocardial infarction including several metabolic pathways and functions specific to cardiomyocytes. One possible explanation for the function of these co-expressed lncRNAs includes operating as competitive inhibitors of miRNAs.[4] This function would also fit with the tendency for lncRNAs to be preferentially positively correlated instead of negatively correlated with mRNAs, a pattern which has been demonstrated by other studies[29]. Moreover, according to the miRcode[48] database, many of the differentially expressed lncRNAs identified here are predicted to interact with many miRNAs that have been implicated in the regulation of myocardial ischemia including miR-1, miR-15, miR-21, miR-24, miR-34, miR-133, and miR-199a[49] (data not shown). The consequences of these interactions in the ischemia phenotype require further investigation.

Future studies will be directed at validating these differentially expressed lncRNAs and determining if the degree of their expression changes following ischemia is affected by conditions such as coronary artery disease, gender, and diabetes. For example, a subgroup analysis on patients with and without diabetes identified 81 lncRNAs with changes in expression following ischemia that may depend on the glycemic environment of the patient (Table S5 and S6).

Since our ischemic time was a median of 74 minutes, there may not have been enough damage or time to allow for large changes in gene expression. Furthermore, by adding a fold change cutoff we may also be excluding those lncRNAs that have significant functions yet small changes in expression. These characteristics of the experimental design may increase the number of false negatives. Nevertheless, we focused our attention on those lncRNAs with the most robust changes in gene expression in order to limit the number of false positives and thus improve the accuracy of our functional predictions, which can be quite difficult for lncRNAs. We would also like to mention that due to the experimental design, non polyadenylated lncRNAs cannot be analyzed in this current study and thus future work will have to be done to provide a comprehensive description of the lncRNA pool of the human left ventricle. Regardless of this caveat, we uncovered many lncRNAs with potential functional roles downstream of the ischemic-reperfusion pathway.

In conclusion, we present the first lncRNA profile of the human LV following the early stages of ischemia. This profile may have utility in predicting the extent of myocardial

injury and can be used to describe several potential mechanisms of interaction between lncRNAs and the coding genes that define the ischemic signature. Further studies, both in humans and in animal models, will help validate the functional significance of these candidate noncoding transcripts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Atkinson SR, Marguerat S, Bahler J. Exploring long non-coding RNAs through sequencing. Semin Cell Dev Biol. 2012; 23:200–205. [PubMed: 22202731]

2. Uchida S, Dimmeler S. Long noncoding RNAs in cardiovascular diseases. Circ Res. 2015; 116:737–750. [PubMed: 25677520]

3. Ounzain S, Pezzuto I, Micheletti R, Burdet F, Sheta R, Nemir M, et al. Functional importance of cardiac enhancer-associated noncoding RNAs in heart development and disease. J Mol Cell Cardiol. 2014; 76:55–70. [PubMed: 25149110]

4. Kornienko AE, Guenzl PM, Barlow DP, Pauler FM. Gene regulation by the act of long non-coding RNA transcription. BMC Biol. 2013; 11:59. [PubMed: 23721193]

5. Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, et al. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. Cell. 2013; 152:570–583. [PubMed: 23352431]

6. Moran I, Akerman I, van de Bunt M, Xie R, Benazra M, Nammo T, et al. Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. Cell Metab. 2012; 16:435–448. [PubMed: 23040067]

7. Ounzain S, Micheletti R, Beckmann T, Schroen B, Alexanian M, Pezzuto I, et al. Genome-wide profiling of the cardiac transcriptome after myocardial infarction identifies novel heart-specific long non-coding RNAs. Eur Heart J. 2015; 36:353–368a. [PubMed: 24786300]

8. Yang KC, Yamada KA, Patel AY, Topkara VK, George I, Cheema FH, et al. Deep RNA sequencing reveals dynamic regulation of myocardial noncoding RNAs in failing human heart and remodeling with mechanical circulatory support. Circulation. 2014; 129:1009–1021. [PubMed: 24429688]

9. Vausort M, Wagner DR, Devaux Y. Long noncoding RNAs in patients with acute myocardial infarction. Circ Res. 2014; 115:668–677. [PubMed: 25035150]

10. Kumarswamy R, Bauters C, Volkmann I, Maury F, Fetisch J, Holzmann A, et al. Circulating long noncoding RNA, LIPCAR, predicts survival in patients with heart failure. Circ Res. 2014; 114:1569–1575. [PubMed: 24663402]

11. Feng J, Bianchi C, Sandmeyer JL, Li J, Sellke FW. Molecular indices of apoptosis after intermittent blood and crystalloid cardioplegia. Circulation. 2005; 112:I184–189. [PubMed: 16159813]

12. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14:R36. [PubMed: 23618408]

13. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]

14. Anders S, Pyl PT, Huber W. HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015; 31:166–169. [PubMed: 25260700]

15. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012; 22:1760–1774. [PubMed: 22955987]

16. Zhao Y, Li H, Fang S, Kang Y, Wu W, Hao Y, et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic Acids Res. 2016; 44:D203–208. [PubMed: 26586799]

17. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2015.

18. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. [PubMed: 25516281]

19. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010; 28:495–501. [PubMed: 20436461]

20. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16:284–287. [PubMed: 22455463]

21. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013; 9:e1003118. [PubMed: 23950696]

22. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res. 2006; 34:D108–110. [PubMed: 16381825]

23. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]

24. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem. 2011; 83:8604–8610. [PubMed: 22035192]

25. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag; New York: 2009.

26. Zangrando J, Zhang L, Vausort M, Maskali F, Marie PY, Wagner DR, et al. Identification of candidate long non-coding RNAs in response to myocardial infarction. BMC Genomics. 2014; 15:460. [PubMed: 24917243]

27. Wang J-X, Zhang X-J, Li Q, Wang K, Wang Y, Jiao J-Q, et al. MicroRNA-103/107 Regulate Programmed Necrosis and Myocardial Ischemia/Reperfusion Injury Through Targeting FADD. Circ Res. 2015; 117:352–363. [PubMed: 26038570]

28. Di Salvo TG, Guo Y, Su YR, Clark T, Brittain E, Absi T, et al. Right ventricular long noncoding RNA expression in human heart failure. Pulm Circ. 2015; 5:135–161. [PubMed: 25992278]

29. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012; 22:1775–1789. [PubMed: 22955988]

30. Harpster MH, Bandyopadhyay S, Thomas DP, Ivanov PS, Keele JA, Pineguina N, et al. Earliest changes in the left ventricular transcriptome postmyocardial infarction. Mamm Genome. 2006; 17:701–715. [PubMed: 16845475]

31. Zhang T, Zhao LL, Cao X, Qi LC, Wei GQ, Liu JY, et al. Bioinformatics analysis of time series gene expression in left ventricle (LV) with acute myocardial infarction (AMI). Gene. 2014; 543:259–267. [PubMed: 24704022]

32. Muehlschlegel JD, Christodoulou DC, McKean D, Gorham J, Mazaika E, Heydarpour M, et al. Using next-generation RNA sequencing to examine ischemic changes induced by cold blood cardioplegia on the human left ventricular myocardium transcriptome. Anesthesiology. 2015; 122:537–550. [PubMed: 25581909]

33. Jordan IK, Marino-Ramirez L, Wolf YI, Koonin EV. Conservation and coevolution in the scale-free human gene coexpression network. Mol Biol Evol. 2004; 21:2058–2070. [PubMed: 15282333]

34. Liu MX, Chen X, Chen G, Cui QH, Yan GY. A computational framework to infer human disease-associated long noncoding RNAs. PLoS One. 2014; 9:e84408. [PubMed: 24392133]

35. Cambier L, Lacampagne A, Auffray C, Pomies P. Fem1a is a mitochondrial protein up-regulated upon ischemia-reperfusion injury. FEBS Lett. 2009; 583:1625–1630. [PubMed: 19406122]

36. Shibata R, Sato K, Pimentel DR, Takemura Y, Kihara S, Ohashi K, et al. Adiponectin protects against myocardial ischemia-reperfusion injury through AMPK- and COX-2-dependent mechanisms. Nat Med. 2005; 11:1096–1103. [PubMed: 16155579]

37. Myocardial Infarction Genetics C. Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009; 41:334–341. [PubMed: 19198609]

38. Kelloniemi A, Szabo Z, Serpi R, Napankangas J, Ohukainen P, Tenhunen O, et al. The Early-Onset Myocardial Infarction Associated PHACTR1 Gene Regulates Skeletal and Cardiac Alpha-Actin Gene Expression. PLoS One. 2015; 10:e0130502. [PubMed: 26098115]

39. McCully JD, Bhasin MK, Daly C, Guerrero MC, Dillon S, Liberman TA, et al. Transcriptomic and proteomic analysis of global ischemia and cardioprotection in the rabbit heart. Physiol Genomics. 2009; 38:125–137. [PubMed: 19454556]

40. Thorp EB. The Myocardial Unfolded Protein Response during Ischemic Cardiovascular Disease. Biochem Res Int. 2012; 2012:583170. [PubMed: 22536506]

41. Angelis E, Zhao P, Zhang R, Goldhaber JI, Maclellan WR. The role of E2F-1 and downstream target genes in mediating ischemia/reperfusion injury in vivo. J Mol Cell Cardiol. 2011; 51:919–926. [PubMed: 21964190]

42. Blais A, Tsikitis M, Acosta-Alvear D, Sharan R, Kluger Y, Dynlacht BD. An initial blueprint for myogenic differentiation. Genes Dev. 2005; 19:553–569. [PubMed: 15706034]

43. Consortium GT. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013; 45:580–585. [PubMed: 23715323]

44. Greco S, Gorospe M, Martelli F. Noncoding RNA in age-related cardiovascular diseases. J Mol Cell Cardiol. 2015; 83:142–155. [PubMed: 25640162]

45. Dieci G, Preti M, Montanini B. Eukaryotic snoRNAs: a paradigm for gene expression flexibility. Genomics. 2009; 94:83–88. [PubMed: 19446021]

46. Depre C, Wang L, Sui X, Qiu H, Hong C, Hedhli N, et al. H11 kinase prevents myocardial infarction by preemptive preconditioning of the heart. Circ Res. 2006; 98:280–288. [PubMed: 16373598]

47. Villegas VE, Zaphiropoulos PG. Neighboring gene regulation by antisense long non-coding RNAs. Int J Mol Sci. 2015; 16:3251–3266. [PubMed: 25654223]

48. Jeggari A, Marks DS, Larsson E. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. Bioinformatics. 2012; 28:2062–2063. [PubMed: 22718787]

49. Boon RA, Dimmeler S. MicroRNAs in myocardial infarction. Nat Rev Cardiol. 2015; 12:135–142. [PubMed: 25511085]

Clinical Perspective

Advances in genomic technology have led to the classification of a novel set of RNA transcripts that do not encode proteins, namely long noncoding RNAs (lncRNAs). The expression of these lncRNAs tends to be highly tissue and disease specific. The discovery of differentially expressed lncRNAs in the human heart following acute ischemia may lead to novel diagnostic and therapeutic targets for ischemic heart disease.
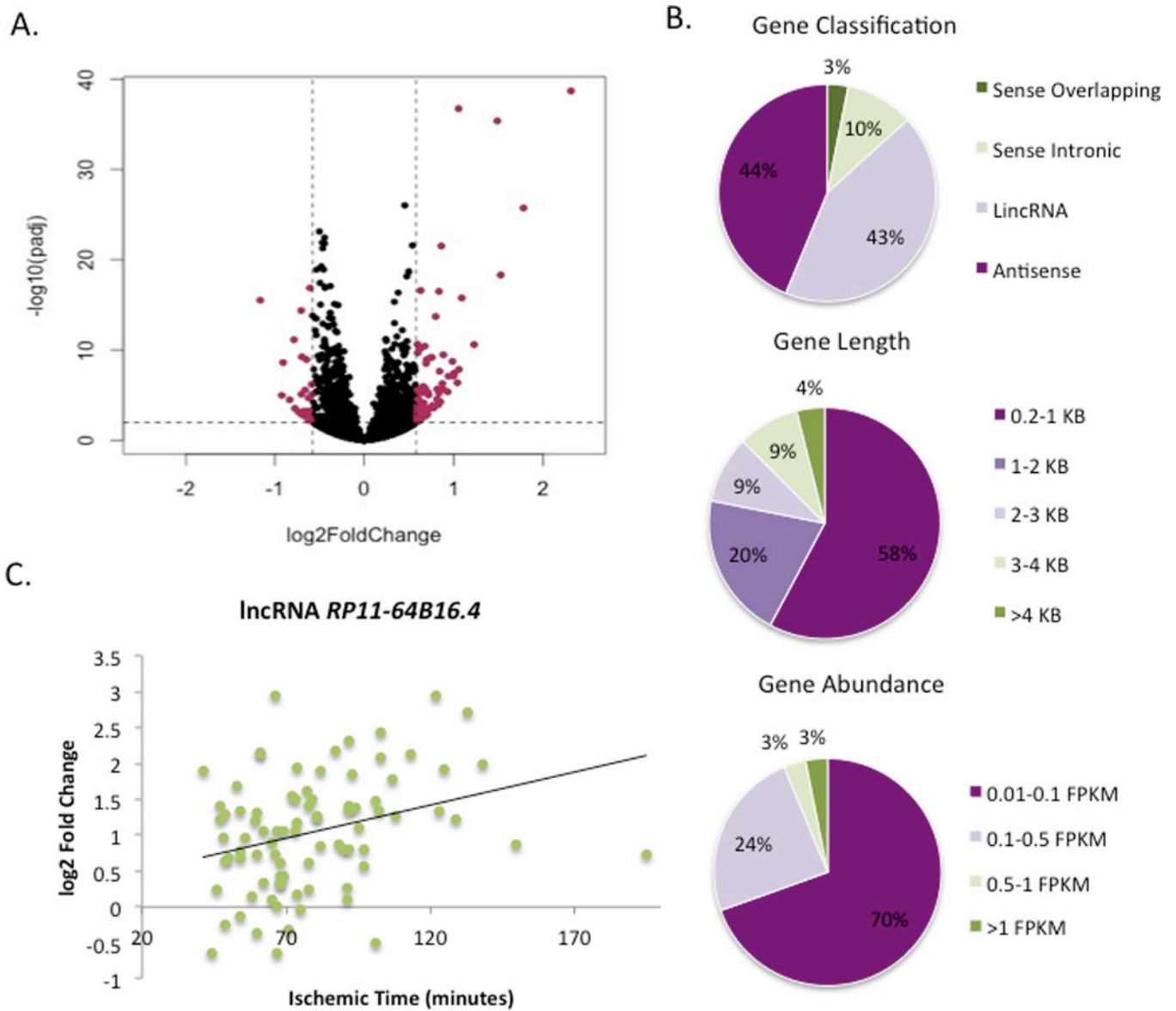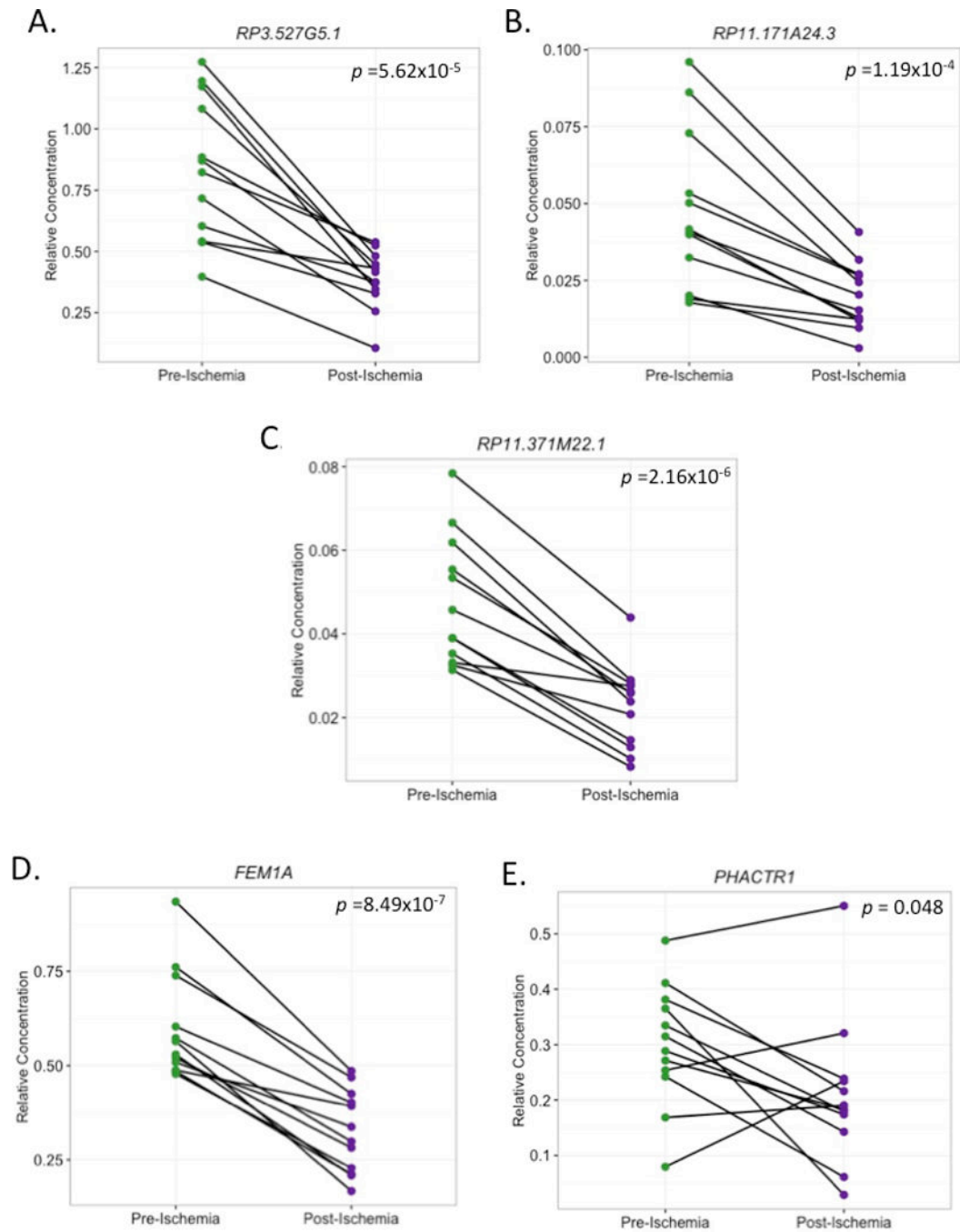
Figure 1.
A. Volcano plot of all lncRNAs which depicts log2 fold change versus −log10 adjusted *p*-value (padj). Significant differentially expressed genes are located between the vertical and horizontal dotted lines and are highlighted in purple. B. Gene classification, length, and abundance distribution of all differentially expressed lncRNAs. C. Scatter plot of log2 fold change expression of *RP11-64B16.4* versus ischemic time (in minutes).
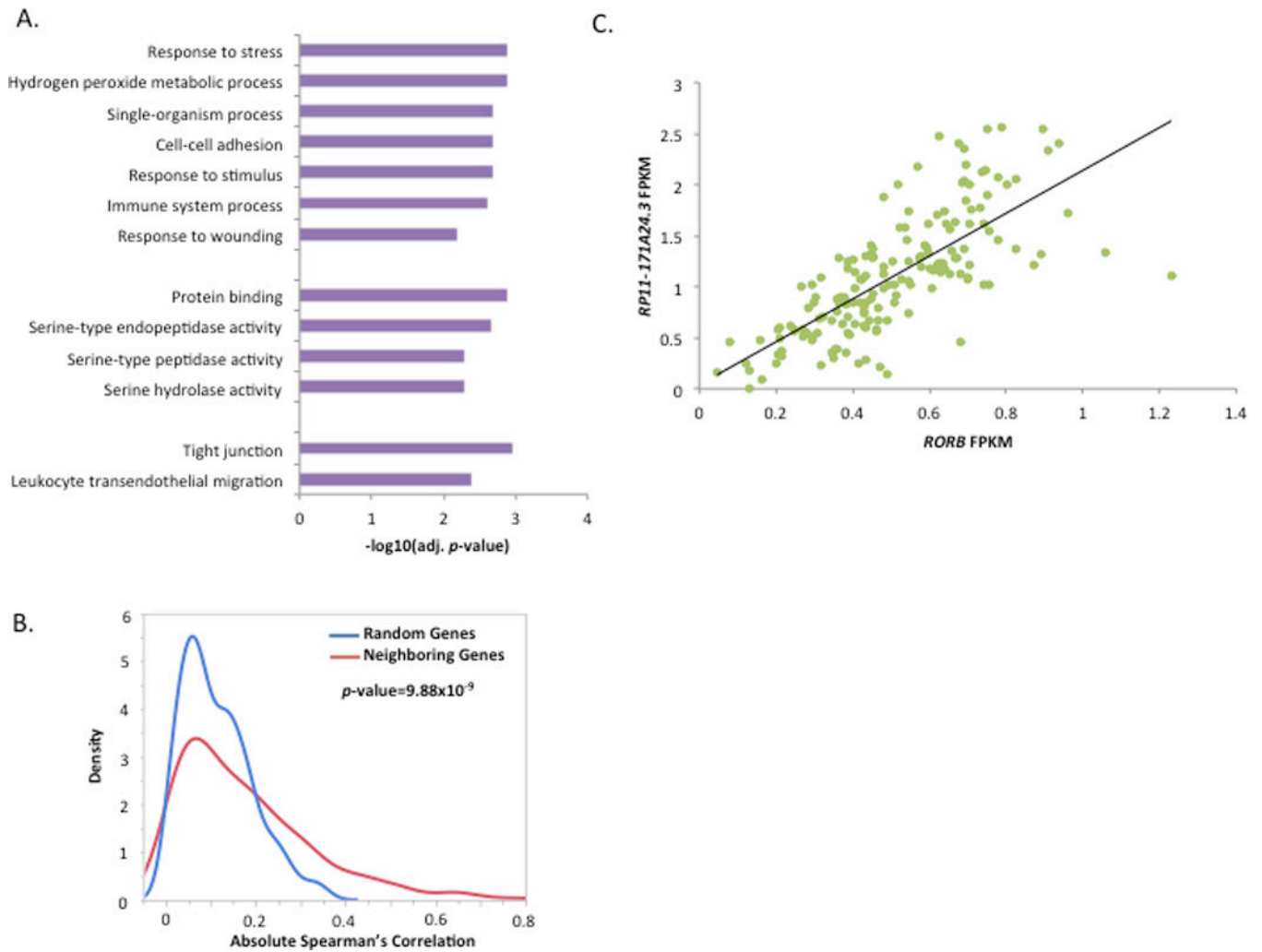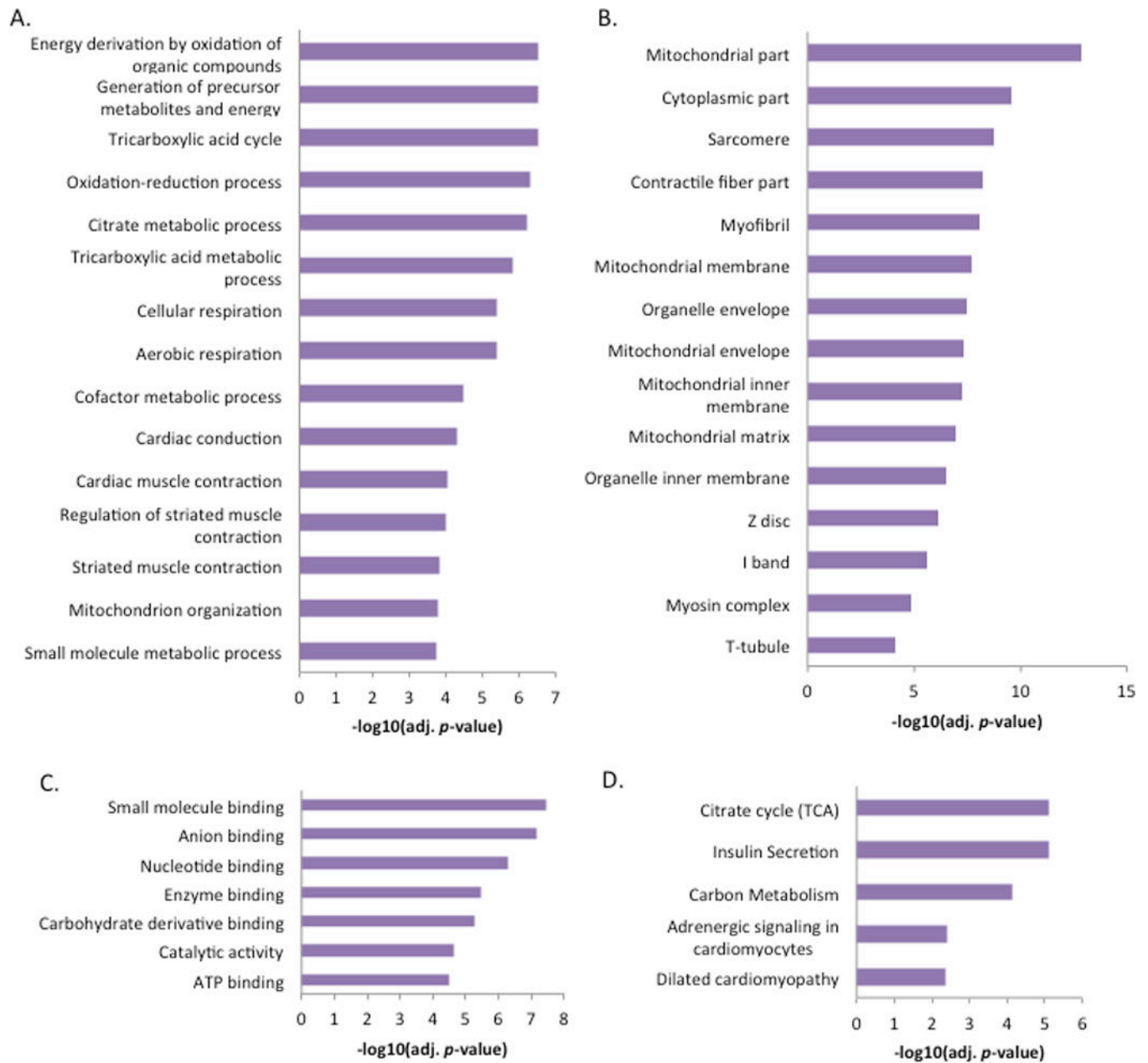
Figure 2.
A. *Neighboring* gene enrichment by Gene Ontology (GO) biological process, GO molecular function, and KEGG pathway. Terms are ranked by significance using -log10 adjusted *p*-value. B. Density plot comparing the Spearman's rank correlation coefficients from the expression of all *neighboring* lncRNA-mRNA pairs with the coefficients from the expression of a random selection of lncRNA-mRNA pairs. *p*-value denotes a standard *t*-test. C. Correlation of expression between the lncRNA *RP11-171A24.3* and its *neighboring* gene *RORB*. FPKM=fragments per kilobase per million mapped reads.

Figure 3.
Droplet Digital PCR of A.) *RP3.527G5.1* B.) *RP11.171A24.3* C.) *RP11.371M22.1* D.) *FEM1A* E.) *PHACTR1* on 12 paired pre- and post-ischemic samples. *p*-value was generated from a paired *t*-test.

Figure 4.
Over-represented Gene Ontology (GO) biological process (A), GO cellular component (B), GO molecular function (C), and KEGG pathway (D) terms of genes with mRNA expression that is highly correlated with the expression of differentially expressed lncRNAs. Terms are ranked by significance using −log10 adjusted *p*-value.

Table 1

Patient demographics and clinical characteristics (N=85)

Demographic	Statistic
Age (years)	71 (64–81)[*]
Male gender	51 (60)[†]
Caucasian Descent	84 (99)[†]
BMI (kg/m^2)	30 (26–34)[*]
Diabetes	37 (44)[†]
CAD	40 (47)[†]
Post-operative Day 1 CKMB (µg/L)	26.95 (22.4–34.0)[*]
Aortic cross-clamp (minutes)	74 (61–93)[*]
LV ejection fraction (%)	60 (55–65)[*]

[*] Median value (inter-quartile range)

[†] Number of patients (percent of patients)

BMI: body mass index; CAD: coronary artery disease; CKMB: creatine kinase MB fraction; LV: left ventricle)

Table 2

Top 15 most abundant lncRNAs

LncRNA	Mean FPKM	Chromosome	Strand
NONHSAG053897.1	956.9	chrM	+
MALAT1	378	chr11	+
H19	99.8	chr11	−
RP3-527G5.1	51.9	chr6	−
LINC00657	46.4	chr20	−
NEAT1	42.1	chr11	+
RP11-480G7.1	26.3	chr16	−
RP11-553L6.5	20.1	chr3	−
AC053503.6	19.9	chr2	−
RP11-532N4.2	15.1	chr6	+
CTD-2083E4.7	12.8	chr5	+
DANCR	12.6	chr4	+
RP11-264A11.1	11.8	chr10	−
LINC00984	10.4	chr15	+
RP11-432J24.5	10.3	chr10	−

lncRNA: long noncoding RNA; FPKM: fragments per kilobase per million mapped reads