Long Non-coding RNAs in Endothelial Biology

Tyler Weirick, Giuseppe Militello and Shizuka Uchida*

Cardiovascular Innovation Institute, University of Louisville, Louisville, KY, United States

In recent years, the role of RNA has expanded to the extent that protein-coding RNAs are now the minority with a variety of non-coding RNAs (ncRNAs) now comprising the majority of RNAs in higher organisms. A major contributor to this shift in understanding is RNA sequencing (RNA-seq), which allows a largely unconstrained method for monitoring the status of RNA from whole organisms down to a single cell. This observational power presents both challenges and new opportunities, which require specialized bioinformatics tools to extract knowledge from the data and the ability to reuse data for multiple studies. In this review, we summarize the current status of long non-coding RNA (lncRNA) research in endothelial biology. Then, we will cover computational methods for identifying, annotating, and characterizing lncRNAs in the heart, especially endothelial cells.

Keywords: bioinformatics, databases, lncRNAs, miRNAs, RNA-seq, RNA editing, RNA modifications

INTRODUCTION

The development of next generation sequencing (NGS) and RNA sequencing (RNA-seq) has significantly improved the understanding of transcriptomes. For example, we now know that most of the human genome is transcribed (Lander et al., 2001), yet only a small percent of these RNAs code for protein (Weirick et al., 2016a). When the human genome was annotated (i.e., giving the definition to the genome by naming a particular gene and its corresponding exons), it was originally thought that the number of protein-coding genes in humans should be more than those of lower organisms (e.g., yeast, plants, fishes, amphibians) (Mercer et al., 2011; Ezkurdia et al., 2014). However, when the numbers of protein-coding genes are compared among species, the number of human genes is not more than those of lower organisms (Figure 1A). Given that humans are able to carry out more complex tasks than lower organisms, the question remains in the field: What aspect of our genome allows for the increased complexity? One school of thoughts is that proteins can be modified for various biological processes (e.g., phosphorylation of a protein for its activation). Another school suggests for the increased variety of isoforms resulting from one gene due to the alternative splicing events. In both schools, the ultimate final products are proteins as we know more about proteins than RNAs. Last school postulates that the increased number of ncRNAs (especially, IncRNAs) is at the base of the highest complexity in human, although it is highly subjective as the number of ncRNAs depends on how well the organism is studied as C. elegans has more IncRNAs than any other organisms (Figure 1B). At the moment, it is most likely that the combination of these schools of thoughts may yield important answers to the question.

In recent years, many review articles about IncRNAs in the heart are published (Geisler and Coller, 2013; Archer et al., 2015; Devaux et al., 2015; Iyer et al., 2015; Ounzain et al., 2015a; Rizki and Boyer, 2015; Ballantyne et al., 2016; Busch et al., 2016; Lorenzen and Thum, 2016; Uchida and Bolli, 2017; Viereck and Thum, 2017; Sallam et al., 2018). Furthermore, there are large amounts of screening data available for IncRNAs.
expressed in the heart (Ounzain et al., 2014, 2015b,c; Kurian et al., 2015) even single-cell RNA-seq data (Chan et al., 2016; Delaughter et al., 2016; King et al., 2017; Lescroart et al., 2018; Skelly et al., 2018). However, these data sets were mostly generated for specific purposes and have largely not been analyzed for lncRNAs. Here, we will focus on endothelial cells (ECs), an important cell type in cardiovascular medicine. Furthermore, we will cover computational methods for identifying, annotating, and characterizing lncRNAs in the heart, especially ECs.

lncRNAs IN ENDOTHELIAL CELLS

Vessels deliver metabolites and oxygen to the tissue and export waste products to sustain the well-being of an organism (Asahara et al., 2011). After tissue injury [e.g., myocardial infarction (MI)], ECs migrate to the site of injury to re-establish the capillary network through a process called “angiogenesis” (Jakobsson et al., 2010; Oka et al., 2014). Furthermore, ECs contribute to the multicellular communications that maintain the balance between the regeneration and dysfunctional or maladaptive healing (Cines et al., 1998; Libby, 2012; Kluge et al., 2013; Eelen et al., 2018). Although advances have been made to understand angiogenesis, the recent emergence of IncRNAs has added another layer of complexity to the genetic network of angiogenesis. To date, a number of IncRNAs are identified and characterized (Table 1) as ECs can be found throughout the human body. For example, MALAT1 regulates endothelial cell function and vessel growth via cell cycle control (Michalik et al., 2014); the histone demethylase JARID1B controls the lncRNA MANTIS, which regulates EC function and vessel growth by binding to the chromatin modifying enzyme BRG1 (Leisegang et al., 2017); and several lncRNAs bind miRNAs to function as miRNA sponges (He et al., 2015, 2017; Huang et al., 2015; Yan et al., 2015; Lu et al., 2016; Ming et al., 2016; Ma Y. et al., 2017; Sun et al., 2017; Zhang B. Y. et al., 2017; Bao et al., 2018).

In addition to the above IncRNAs, we recently reported the presence of an emerging class of IncRNAs called “circular RNAs (circRNAs)” in ECs (Boeckel et al., 2015). CircRNAs are byproducts of splicing events (more specifically, "backsplicing") of mostly protein-coding genes (Jeck et al., 2013; Jeck and Sharpless, 2014; Boeckel et al., 2015), are stable and localized predominantly in the cytoplasm (Nigro et al., 1991; Cocquerelle et al., 1993). When some circRNAs are knocked down, there are phenotypes observed, which may not be observed when the parental transcripts of circRNAs are knocked down (Boeckel et al., 2015; Gerstner et al., 2016). For example, the hypoxia-regulated circRNA cZNF292, which is derived by backsplicing of ZNF292 protein-coding gene, exhibits proangiogenic activities (Boeckel et al., 2015). Some studies suggest that circRNAs function as miRNA sponges (Hansen et al., 2013; Memczak et al., 2013; Geng et al., 2016; Liu Q. et al., 2016; Zheng et al., 2016). However, recent comprehensive bioinformatics analysis (Guo et al., 2014) and our biological validation experiments
| IncRNA     | Organism       | Tissue                                      | Cell type | Function                                                                                           | References               |
|------------|----------------|---------------------------------------------|-----------|----------------------------------------------------------------------------------------------------|--------------------------|
| ALT1       | Human          | N/A                                         | HUVEC     | Interacts with ACE2 and CUL1 to control the expression of Cyclin D1 possibly via ubiquitination and degradation. | Li et al., 2017b         |
| ASncmtRNA-2| Human; mouse   | Aortas of old mice                          | HUVEC     | Might be involved in the RS establishment by participating in the cell cycle arrest in G2/M phase, possibly through the production of miR-4485 and -1973. | Bianchessi et al., 2015  |
| GAS5       | Human          | Atherosclerotic plaques                     | HUVEC     | Can be transferred from macrophages to EC in exosomes to induce apoptosis of ECs.                   | Chen et al., 2017        |
| GATA6-AS   | Human; mouse   | Cell-based Xenograft model                  | HUVEC     | Binds the lysyl oxidase LOXL2 to impair its function as H3K4me3 deaminase.                         | Neumann et al., 2018     |
| H19        | Human          | Brain; Glioma tissue specimens             | HBMVEC    | Knockdown of H19 suppressed glioma induced angiogenesis by inhibiting miR-29a, which may modulate the onset of glioma by regulating biological behaviors of glioma vascular ECs. | Jia et al., 2016         |
| H19        | Human          | N/A                                         | HUVEC     | Is contained in exosomes released by CD90+ cancer cells to promote angiogenic phenotype and cell-to-cell adhesion in ECs. | Conigliaro et al., 2015  |
| HIF1A-AS2  | Human; rat     | Permanent middle cerebral artery occlusion model | HUVEC     | Facilitates the up-regulation of HIF-1α by sponging miR-153-3p, thereby promoting angiogenesis in hypoxia. | Li et al., 2017a         |
| HOTAIR     | Human          | N/A                                         | HBMVEC    | Is contained in the glioma cell-derived extracellular vesicles and transmitted into ECs.            | Ma X. et al., 2017       |
| HOTAIR     | Human          | Atherosclerotic plaques                     | HUVEC; HAEC | Positively regulates proliferation and migration of ECs.                                         | Peng et al., 2017        |
| HOTTIP     | Human          | CAD and normal arterial tissues             | HUVEC     | Its overexpression induces β-catenin expression and enhances the downstream protein c-Myc expression in ECs to affect cell proliferation and migration. | Liao et al., 2018        |
| IGF2AS     | Rat            | Heart                                       | mMVE      | Reciprocal regulation IGF2AS and IGF2 is critical in modulating angiogenic development in myocardial tissues in type 2 diabetes. | Zhao et al., 2017        |
| LEENE      | Human; mouse   | Thoracic aorta and aortic arch              | HUVECs; HAoEC | Serves as a guide to facilitate RNA Pol II binding to the promoter of eNOS.                       | Miao et al., 2018        |
| LINC00305  | Human          | N/A                                         | HUVEC     | Guides EZH2 [the catalytic subunit of polycomb repressive complex 2 (PRC2)] to the promoter region of the VCAM1 gene to suppress VCAM1. | Zhang B. Y. et al., 2017  |
| LINC00341  | Human          | N/A                                         | HUVEC     | Binds miR-130b to promote cell apoptosis and induce cell cycle progression.                      | Huang et al., 2017       |
| LINC00657  | Human          | N/A                                         | HUVEC     | Binds miR-590-3p to attenuate the suppression of miR-590-3p on HIF-1α, and to promote angiogenesis through VEGF, MMP-2, and MMP-9. | Bao et al., 2018         |
| lincRNA-p21| Mouse          | N/A                                         | mouse lymphoid endothelial cell line SVEC4       | Binds miR-130b to promote cell apoptosis and induce cell cycle progression.                      | He et al., 2015          |
| LISPR1     | Human          | N/A                                         | HUVEC; HAoEC; HMEC | Acts as a novel regulatory unit important for S1PR1 expression and EC function.                  | Josipovic et al., 2018   |
| lncRNA          | Organism          | Tissue                                      | Cell type                      | Function                                                                                                                                   | References |
|----------------|-------------------|---------------------------------------------|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| LOC100129973   | Human             | N/A                                         | HUVEC                         | Binds miR-4707-5p and -4767, which promote apoptosis by targeting and downregulating two apoptosis inhibitors, APIS and BCL2L12, respectively. | Lu et al., 2016 |
| MALAT1         | Human; mouse      | Mouse retinal angiogenesis model            | HUVEC                         | Protects the endothelium from ox-LDL-induced endothelial dysfunction partly through competing with miR-22-3p for endogenous RNA.          | Michalik et al., 2014 |
| MALAT1         | Human; mouse      | Retina of diabetic rats                     | Monochrome retina cell line RF/6A | Regulates EC function via p38 MAPK signaling pathway.                                                                                       | Liu et al., 2014 |
| MALAT1         | Human; mouse      | Kidneys of diabetic mice                    | HUVEC                         | Regulates glucose-induced up-regulation of inflammatory mediators IL-6 and TNF-α through activation of SAA3.                                | Puthanveetil et al., 2015 |
| MANTIS         | Human             | Brain microvessel isolation from glioblastoma patients | HUVEC; HAoEC; HDLEC; PAEC      | Regulates EC function and vessel growth by binding to the chromatin modifying enzyme BRG1.                                                  | Leisegang et al., 2017 |
| Meg3           | Mouse; monkey     | Retina of diabetic mice                     | Monochrome retina cell line RF/6A | Activates PI3K/Akt signaling.                                                                                                               | Qiu et al., 2016 |
| Meg3           | Rat               | Brain                                       | RSMVEC                        | Physically interacts with p63, which binds to the promoter of Atax to regulate cell growth and the blood vessel growth factor expression. | Zhan et al., 2017 |
| MEG3           | Human             | N/A                                         | HUVEC                         | Is regulated by HIF-1α to maintain VEGFR2 expression in ECs and plays a vital role for VEGFA-mediated endothelial angiogenesis.           | Ruan et al., 2018 |
| MEG3           | Human; mouse      | Hind-limb ischemia in aged mice             | HUVEC                         | Its silencing prevents aging-mediated inhibition of sprouting activity.                                                                      | Boon et al., 2016 |
| MEG3           | Human; mouse      | Circulating ECs from metabolic syndrome (MetS) patients | EPC                           | Protects ECs via decreasing miR-140-5p expression and increasing HDAC7 expression in MetS.                                                   | Liu H. Z. et al., 2016 |
| MEG3           | Human             | N/A                                         | HUVEC                         | Binds miR-9 to control the proliferation and angiogenesis of ECs.                                                                             | He et al., 2017 |
| MiAT           | Human; rat        | Diabetes mellitus                           | HUVEC; HMVEC                   | Binds miR-150-5p to regulate EC function by forming a feedback loop with VEGF.                                                              | Yan et al., 2015 |
| PUNISHER       | Human; zebrafish  | Heart                                       | HUVEC                         | Its inhibition results in severe vascular defects in zebrafish embryos and reduced cell proliferation in HUVEC.                             | Kurian et al., 2015 |
| PVT1           | Human             | N/A                                         | Human cerebral microvascular endothelial cell line hCMEC/D3 | Binds miR-186, which targets Atg7 and Beclin1 mRNAs.                                                                                     | Ma Y. et al., 2017 |
| RNOR3/LINC005999| Human; mouse      | Aortic atherosclerotic lesions              | HUVEC                         | Forms a feedback loop with KLF2 and miR-185-5p to regulate EC function.                                                                      | Shan et al., 2016 |
| SENCOR         | Human             | N/A                                         | HUVEC                         | Induces proliferation, migration, and angiogenesis.                                                                                         | Boulberdaa et al., 2016 |
| SIRT1 AS       | Mouse             | N/A                                         | EPC                           | Relieves miR-22-induced SIRT1 downregulation by competitively sponging miR-22.                                                              | Ming et al., 2016 |
TABLE 1 | Continued

| IncRNA     | Organism          | Tissue                  | Cell type       | Function                                                                 | References          |
|------------|-------------------|-------------------------|-----------------|--------------------------------------------------------------------------|---------------------|
| STEEL      | Human             | N/A                     | HUVEC; HMVEC    | Binds the chromatin-associated enzyme PARP1 to assist its binding to the KLF2 and eNOS promoters. | Man et al., 2018    |
| TGFB2-OT1  | Human             | N/A                     | HUVEC           | Binds miR-3960, -4488, and -44459, which target CERS1, NAT8L, and LARP1, respectively, the key proteins involved in autophagy and inflammation. | Huang et al., 2015  |
| tie-1AS lncRNA | Human; mouse; Zebrafish | Tg(tx:EGFP)           | HUVEC           | Binds tie-1 miRNA and regulates tie-1 transcript levels, resulting in specific defects in endothelial cell contact junctions. | Li et al., 2010     |
| uc001pwg.1 | Human             | Stenosed and nonstenotic uremic veins | HUVEC; EC derived from human-induced pluripotent stem cells | Its overexpression increases eNOS phosphorylation and NO production by affecting the expression level of nearby protein-coding gene. | Lv et al., 2017     |

Each IncRNA is listed with organism(s), tissue(s), cell type(s), and function(s) along with the corresponding reference. The abbreviations used are as follows: “coronary artery disease (CVD)”; “endothelial progenitor cells (EPCs)”; “human aortic endothelial cell (HAEC)”; “human aortic endothelial cells (HAoEC)”; “human brain microvascular endothelial cells (HMVEC)”; “human dermal lymphatic endothelial cells (HDLEC)”; “human microvascular endothelial cells (HMVEC)”; “human pulmonary artery endothelial cells (PAEC)”; “rat brain microvascular endothelial cells (RBMEC)”; and “not applicable (N/A)”.

(Boeckel et al., 2015; Weirick et al., 2016b) indicate that circRNAs functioning as miRNA sponges are extremely rare. Along with lncRNAs, more studies are necessary to uncover the functions of circRNAs in ECs.

RNA-seq DATA ANALYSIS USING BIOINFORMATICS

There are two major methods of generating libraries for RNA-seq, which are based on poly-A selection and ribosomal RNA (rRNA)-depletion. Both methods are aimed at removing rRNAs, which constitute ~80% of total RNA followed by 15% transfer RNAs (tRNAs) and only 5% for all other RNAs, including protein-coding genes and lncRNAs (Lodish et al., 2000). The poly-A selection will result in the identification of protein-coding genes and lncRNAs with poly A tails (~60% of total lncRNAs; Cheng et al., 2005), while the rRNA-depletion can identify the rest of lncRNAs and circRNAs—in addition to those identified in the former method. The presence of circRNAs is detected only with the latter method as circRNAs arise from exons and/or introns that are spliced out, which are devoid of poly A tails.

Analysis of RNA-seq data usually involves a number of common computational steps to obtain the expression profiles of the RNA in a set of samples. At the start of a typical analysis pipeline, reads are trimmed to remove primers and low-quality regions of reads. Next, the reads are aligned to a genome in a “guided alignment.” In the case of the organism with no reference genome, a “de novo assembly” of the transcriptome is performed. However, de novo assembly is more error-prone and difficult to operate, thus we will simply focus on guided alignments. Traditionally, Tophat (Trapnell et al., 2012) has been the most popular aligner, but it is now being supplanted by newer programs (e.g., STAR, HISAT2), which offer greater speed and alignment accuracy (Engström et al., 2013; Conesa et al., 2016; Costa-Silva et al., 2017; Zhang C. et al., 2017).

Similar to protein-coding genes, lncRNAs undergo alternative splicing (AS) to produce isoforms (Deveson et al., 2017; White et al., 2017). The current understanding of AS is mainly based on EST-cDNA sequencing and short-read RNA-seq data. In the second-generation sequencing (e.g., Illumina-based short RNA-seq), long strands of cDNA must be broken into small segments to infer nucleotide sequences by amplification and synthesis (Metzker, 2010), which fall short of detecting intact full-length transcripts. To address this shortcoming, third-generation sequencing (also known as “long-read sequencing”) may be a solution. PacBio RS II (Pacific Biosciences, CA, U.S.A.) is the first commercialized third-generation sequencer, which utilizes a novel single molecule real-time (SMRT) technology (Schadt et al., 2010). Compared to second-generation sequencing, SMRT technology offers long read lengths (up to 92 kb), high consensus accuracy (free of systematic sequencing errors), and low degree of bias (even coverage across G+C content) (Nakano et al., 2017). When this technology is applied to any transcriptome (cDNA) sequencing (e.g., RNA-seq), it is called “Iso-Seq,” which can monitor AS (Abdel-Ghany et al., 2016). With Iso-Seq, the need for transcriptome assembly is eliminated as one read = one transcript” with each transcript can be read from its 5’-end to poly A tail. Iso-Seq has been applied to various species and tissues (Singh et al., 2016; Cheng et al., 2017; Hoang et al., 2017a,b; Jiang et al., 2017; Jo et al., 2017; Kim et al., 2017; Kuo et al., 2017; Wang et al., 2017a,b; Xue et al., 2017; Zhang S. J. et al., 2017; Zulkapli et al., 2017; Filichkin et al., 2018) but not yet to ECs.

The largely unbiased manner in which RNA-seq captures information is another interesting aspect of the technology, which enables new findings via re-analysis of published data. For example, most of the RNA-seq studies have been focused on analyzing expression of protein-coding genes. As lncRNA are
also present in the data sets, these data offer a rich resource for studying lncRNA expression patterns. We have developed a number of bioinformatics tools to exploit these resources (Gellert et al., 2013; Weirick et al., 2015, 2016b, 2017), including some specifically designed to identify lncRNAs and to associate their expressions in various tissues and cell types, including ECs (e.g., our database ANGIOGENES; Müller et al., 2016). Although ECs can be found throughout the human body, there are only a few databases available that contain the expression profiles for genes expressed in ECs (e.g., Causal Biological Network database Boué et al., 2015, dBASENA Savas, 2012, and PubAngioGen Li et al., 2015). Our ANGIOGENE is one of the few that contain the expression profiles of both protein-coding genes and lncRNAs in various ECs based on RNA-seq data. Furthermore, ANGIOGENES covers humans, mice, and zebrafish to allow for the screening of lncRNAs in the positional conserved regions (not necessary sequence-conserved) (Weirick et al., 2015).

There are many transcripts whose sequencing reads are present in RNA-seq data but are not annotated in the public databases, including NONCODE (Zhao et al., 2016), which is one of the hallmark databases for lncRNAs. Our previous study (Weirick et al., 2016a) shows that 77,656 novel isoforms of annotated reference transcripts and 102,848 intergenic transcripts are identified with 58,789 (75.70%) and 101,993 (99.17%) being predicted as non-coding, respectively, from 12 human tissues (Nielsen et al., 2014), while there are 181,434 annotated transcripts (87.13% out of 208,244 transcripts in Ensembl version 77) are expressed in at least one of 12 tissues analyzed. Although we could validate the presence of novel lncRNAs by RT-PCR experiments, many novel lncRNAs contain repetitive elements, such as microsatellites (Bidichandani et al., 1998) and short interspersed nuclear elements (SINE), including ALU elements (Häsler and Strub, 2006). Thus, it is highly recommended to consult the available methods to characterize lncRNAs (Li et al., 2014; Liu et al., 2017), including CAGE-seq to annotate the 5′-end of lncRNAs (Hon et al., 2017) and ribo-seq/ribosomal footprinting RNA-seq technology to understand the coding potential (Ruíz-Rorera et al., 2014; Ji et al., 2015; Alvarez-Dominguez and Lodish, 2017) before proceeding to more functional experiments.

It is well-known that ECs are heterogeneous populations of cells as their activities and functions differ based on their physiological locations (Aird, 2012; Regan and Aird, 2012; Yuan et al., 2016). In order to understand such heterogeneity of ECs, it is important to perform single-cell RNA-seq (scRNA-seq) instead of bulk RNA-seq by using a piece of tissue or those in a culture dish. As the technique for scRNA-seq matures, the immediate problem is the data analysis, especially positioning each cell to a particular cell type in order to organize their molecular signatures matching to the anatomical location in which each cell was isolated from. For example, hearts contain multiple cell types (e.g., cardiomyocytes, ECs, fibroblasts, pericytes, and smooth muscle cells). In regards to ECs, their expression profiles may differ for those contained in the artery and vein. When such profiles are compared to ECs from other tissues (e.g., kidneys, lungs), there are some genes that are expressed at the similar level in all tissues while others are expressed specifically in ECs isolated from a particular tissue. In order to understand such hierarchical organization of cells, their corresponding cell types, and tissues, it is utmost importance that the ontology of each cell must be organized in relation to its corresponding cell type and tissue. To achieve this hierarchical and ontological organization, we recently introduced the usage of logic programming (Weirick et al., 2016b), which was applied to kidneys. Logic programming is a programming paradigm based on formal logic, using a set of logical sentences consisting of facts, rules, and queries (Eklund and Klawonn, 1992). For example, consider a transcript expressed in the renal cortex. The renal cortex is located within kidneys. When sequencing whole kidney under the same condition, the same transcript should be expressed. One could even descend to the level of cell types (e.g., ECs isolated from interlobular arteries, which are located within the kidney cortex). Similarly, all sequences expressed within these ECs are expressed in the kidney. Furthermore, it is well-known that high abundance sequences can overwhelm lower abundance sequences. Thus, logic programming can be useful for integrating RNA-seq data at different hierarchical levels and beyond. This can be accomplished by: (1) modeling the anatomical and experimental relationships; (2) creating rules to define various types of expression characteristics; and (3) using queries to determine expression characteristics of a given RNA. The analysis of RNA-seq data of ECs in the heart for lncRNAs, coupled with logic programming, should help to facilitate the further usage of the available RNA-seq data (e.g., single cell RNA-seq data from the heart) to test various hypotheses that were not originally intended when the data were generated. Such an approach should yield the identification of lncRNAs in a variety of conditions (e.g., expressed in atherosclerotic plaques but not in the healthy artery), which can be further validated in functional studies.

DETECTION OF RNA EDITING PATTERNS FROM RNA-seq DATA

In addition to studying lncRNAs, re-analysis of publicly-available RNA-seq data is also useful for studying RNA editing. RNA editing is a post-transcriptional modification to alter the sequence of RNA molecules (Keegan et al., 2001; Hideyama and Kwak, 2011). The full extent and reasons for RNA editing is largely unknown. However, recent studies show that the editing in exons leads to an amino acid substitutions from altered codons (Alon et al., 2015; Liscovitch-Brauer et al., 2017), whereas editing in 3′-untranslated regions (UTRs) may affect binding of RNA binding proteins (RBPs) or microRNAs (miRNAs) thereby modulating RNA stability and/or translation (Keegan et al., 2001). There are two types of RNA editing: adenosine to inosine (A-to-I) and cytidine to uridine (C-to-U). A-to-I is the most common form and occurs through RNA editing enzymes called “adenosine deaminases acting on RNA (ADARs),” which convert adenosine in double-stranded RNA into inosine (Savva et al., 2012). When reverse transcribed to complementary DNA (cDNA), an inosine is converted to guanine (“G”), which can be identified by comparison to the reference genome. A number
of studies have been conducted to detect RNA editing events from RNA-seq data (Bahn et al., 2012; Park et al., 2012; Peng et al., 2012; Ramaswami et al., 2012, 2013; Solomon et al., 2013), including our recent study in ECs (Stellos et al., 2016). Because of the detection from RNA-seq data, several databases for RNA editing events have been constructed to provide evidence for the frequency of RNA editing in various conditions (Kiran and Baranov, 2010; Picardi et al., 2011, 2017; Laganà et al., 2012; Ramaswami and Li, 2014; Solomon et al., 2016; Gong et al., 2017). We recently reported that cathepsin S (CTSS), which encodes a cysteine protease associated with angiogenesis and atherosclerosis, is highly edited (Stellos et al., 2016). Such RNA editing enables the recruitment of stabilizing RBP human antigen R (HuR) to the 3′-UTR of CTSS transcript, thereby controlling CTSS mRNA stability and expression. The RNA editing enzyme ADAR1 levels and the extent of CTSS RNA editing are associated with changes in CTSS levels in patients with coronary artery diseases. Our study highlights the involvement of RNA editing in cardiovascular diseases, which has not yet been investigated (Uchida and Jones, 2018). Our finding was further supported by the recent large-scale, multi-center study analyzing RNA-seq data from the NIH Common Fund’s Genotype-Tissue Expression (GTEx) program, which reported that the aorta, coronary, and tibial arteries were the most highly edited tissue type among 53 body sites from 552 individuals analyzed (Tan et al., 2017).

In humans, RNA editing occurs mostly in repetitive Alu regions (Levanon et al., 2004; Peng et al., 2012), which can be found in lncRNAs as lncRNAs can also be edited (Picardi et al., 2014; Szczesniak and Makalowska, 2016; Gong et al., 2017). Although proposed but not tested extensively, the functions of lncRNAs may depend on their conformation (e.g., 3D structures), which can be affected by their primary sequences. This folding process can be influenced by a variety of factors, including (but not limited to) RNA modifications on lncRNAs, such as RNA editing. Given that RNA editing can be readily detected from RNA-seq data, more systematic analysis of RNA editing patterns is necessary, especially targeting lncRNAs in the heart (Uchida and Jones, 2018). For this purpose, several bioinformatics tools are available to detect editing within RNA-seq data, including GIREMI (Zhang and Xiao, 2015), JACUSA (Piechotta et al., 2017), RED-ML (Xiong et al., 2017), REDTools (Picardi and Pesole, 2013), RES-Scanner (Wang et al., 2016), and our RNAEditor (John et al., 2017).

**HOW COULD WE TRANSLATE THE CONCEPT OF lncRNAs INTO RNA THERAPEUTICS**

The one obvious usage of lncRNAs in medicine is using lncRNAs as diagnostic biomarkers as lncRNAs are more cell-type specifically expressed than protein-coding genes (Thurman et al., 2012; Gellert et al., 2013; Necsulea et al., 2014; Weirick et al., 2015). Although some progresses have been made, most of RNA-seq data analyzed so far does not consider lncRNAs due to the reasons mentioned above. Thus, without performing further RNA-seq experiments, it should be feasible to discover lncRNAs that capable of differentiating between diseased and healthy individuals by re-analyzing publicly-available RNA-seq data. For this purpose, bioinformatics tools mentioned above should be useful.

**AUTHOR CONTRIBUTIONS**

All authors made contributions to survey the current status of lncRNA research. All authors approved the final version of this manuscript.

**FUNDING**

This study was supported by the V. V. Cooke Foundation (Kentucky, U.S.A.); grant from the University of Louisville School of Medicine; an EVPRI Internal Research Grant from the Office of the Executive Vice President for Research and Innovation at the University of Louisville; University of Louisville 21st Century University Initiative on Big Data in Medicine; and the startup funding from the Mansbach Family, the Gheens Foundation and other generous supporters at the University of Louisville.

**REFERENCES**

Abdel-Ghany, S. E., Hamilton, M., Jacobi, J. L., Ngam, P., Devitt, N., Schliskey, F., et al. (2016). A survey of the sorghum transcriptome using single-molecule long reads. Nat. Commun. 7:11706. doi: 10.1038/ncomms11706

Aird, W. C. (2012). Endothelial cell heterogeneity. Cold Spring Harb. Perspect. Med. 2:a006429. doi: 10.1101/cshperspect.a006429

Alon, S., Garrett, S. C., Levanon, E. Y., Olson, S., Graveley, B. R., Rosenthal, J. J., et al. (2015). The majority of transcripts in the squid nervous system are extensively recoded by A-to-I RNA editing. Elife 4:e05198. doi: 10.7554/eLife.05198

Alvarez-Dominguez, J. R., and Lodish, H. F. (2017). Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. Blood 130, 1965–1975. doi: 10.1182/blood-2017-06-788695

Archer, K., Broskova, Z., Bayoumi, A. S., Teoh, J. P., Davila, A., Tang, Y., et al. (2015). Long non-coding RNAs as master regulators in cardiovascular diseases. Int. J. Mol. Sci. 16, 23651–23667. doi:10.3390/ijms161023651

Asahara, T., Kawamoto, A., and Masuda, H. (2011). Concise review: circulating endothelial progenitor cells for vascular medicine. Stem Cells 29, 1650–1655. doi:10.1002/stem.745

Bahn, J. H., Lee, J. H., Li, G., Greer, C., Peng, G., and Xiao, X. (2012). Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. Genome Res. 22, 142–150. doi:10.1101/gr.124107.111

Ballantyne, M. D., McDonald, R. A., and Baker, A. H. (2016). lncRNA/MicroRNA interactions in the vasculature. Clin. Pharmacol. Ther. 99, 494–501. doi:10.1002/cpt.355

Bao, M. H., Li, G. Y., Huang, X. S., Tang, L., Dong, L. P., and Li, J. M. (2018). Long non-coding RNA LINC00657 acting as miR-590-3p sponge to facilitate low concentration oxidized low-density lipoprotein-induced angiogenesis. Mol. Pharmacol. 93, 368–375. doi: 10.1124/mol.117.110650

Bianchessi, V., Badi, I., Bertolotti, M., Nigro, P., D’alessandra, Y., Capogrossi, M. C., et al. (2015). The mitochondrial lncRNA ASncmtRNA-2 is induced in aging and replicative senescence in Endothelial Cells. J. Mol. Cell. Cardiol. 81, 62–70. doi:10.1016/j.yjmcc.2015.01.012
Bidichandani, S. I., Ashizawa, T., and Patel, P. I. (1998). The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. Am. J. Hum. Genet. 62, 111–121. doi: 10.1086/301680

Boeckel, J. N., Jat, N., Heumuller, A. W., Chen, W., Boon, R. A., Stellos, K., et al. (2015). Identification and characterization of hypoxia-regulated endothelial circular RNA. Circ. Res. 117, 884–890. doi: 10.1161/CIRCRESAHA.115.306319

Boon, R. A., Hofmann, P., Michalk, K. M., Lozano-Vidal, N., Berghäuser, D., Fischer, A., et al. (2016). Long noncoding RNA Meg3 controls endothelial cell aging and function: implications for regenerative angiogenesis. J. Am. Coll. Cardiol. 68, 2589–2591. doi: 10.1016/j.jacc.2016.09.049

Boué, S., Talikka, M., Westra, J. W., Hayes, W., Di Fabio, A., Park, J., et al. (2015). Causal biological network database: a comprehensive platform of causal biological network models focused on the pulmonary and vascular systems. Database (Oxford) 2015:bav030. doi: 10.1093/database/bav030

Boulberdaa, M., Scott, E., Ballantyne, M., Garcia, R., Descamps, B., Angeline, G. D., et al. (2016). A role for the long noncoding RNA SENCER in commitment and function of endothelial cells. Mol. Ther. 24, 978–990. doi: 10.1038/mt.2016.41

Busch, A., Eken, S. M., and Maegdefessel, L. (2016). Prospective and therapeutic screening value of non-coding RNA as biomarkers in cardiovascular disease. Ann. Transl. Med. 4:236. doi: 10.21037/atm.2016.06.06

Chan, S. S., Chan, H. H. W., and Kyba, M. (2016). Heterogeneity of Mesp1 functions in diverse cellular contexts. Nat. Rev. Mol. Cell Biol. 17, 69–78. doi: 10.1038/nrm.2016.117

Chen, L., Yang, W., Guo, Y., Chen, W., Zheng, P., Jeng, J., et al. (2017). Exosomal IncRNA GAS5 regulates the apoptosis of macrophages and vascular endothelial cells in atherosclerosis. PLoS ONE 12:e0185406. doi: 10.1371/journal.pone.0185406

Cheng, B., Furtado, A., and Henry, R. J. (2017). Long-read sequencing of the coffee bean transcriptome reveals the diversity of full-length transcripts. Gigascience 6, 1–13. doi: 10.1093/gigascience/gix086

Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., et al. (2005). Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science 308, 1149–1154. doi: 10.1126/science.1108625

Cines, D. B., Pollak, E. S., Buck, C. A., Loscalzo, J., Zimmerman, G. A., McEver, R. P., et al. (1998). Endothelial cells in physiology and in the pathophysiology of vascular disorders. Blood 91, 3527–3561.

Cocquerelle, C., Mascrè, B., Hétuin, D., and Baleul, B. (1993). Miss-splicing yields circular RNA molecules. FASEB J. 7, 155–160. doi: 10.1096/fasebj.7.1.7678559

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrerío, D., Cervera, A., McPherson, A., et al. (2016). A survey of best practices for RNA-seq data analysis. Genome Biol. 17:13. doi: 10.1186/s13059-016-0881-8

Coniglio, A., Costa, V., Lo Dico, A., Saieva, L., Buccheri, S., Dieli, F., et al. (2015). CD90+ liver cancer cells modulate endothelial cell phenotype through the release of exosomes containing H19 IncRNA. Mol. Cancer 14:135. doi: 10.1186/s12943-015-0426-x

Costa-Silva, J., Domingues, D., and Lopes, F. M. (2017). RNA-Seq differential expression analysis: an extended review and a software tool. PLoS ONE 12:e0190152. doi: 10.1371/journal.pone.0190152

Delauerta, D. M., Bick, A. G., Wakimoto, H., McLean, D., Gorham, J. M., Kathiriya, I. S., et al. (2016). Single-cell resolution of temporal gene expression during heart development. Dev. Cell 39, 480–490. doi: 10.1016/j.devcel.2016.10.001

Devaux, Y., Zangrando, J., Schroen, B., Creemers, E. E., Pedrazzini, T., Chang, C. C., Agarwal, V., Guo, H., and Bartel, D. P. (2014). Expanded identification and characterization of mammalian circular RNAs. Genome Biol. 15:409. doi: 10.1186/s13059-014-0409-z

Hoang, N. V., Furtado, A., Chenaus, E., Sikk, A., Bick, A. G., Docastro, D., Sahu, A., Hosono, Y., et al. (2013). Natural RNA circles function as efficient microRNA sponges. Nature 495, 384–388. doi: 10.1038/nature12193

Häser, F., and Strub, K. (2006). Alu elements as regulators of gene expression. Nucleic Acids Res. 34, 5491–5497. doi: 10.1093/nar/gkl2706

He, C., Ding, J. W., Li, S., Wu, H., Jiang, Y. R., Yang, W., et al. (2015). The ROLE of LONG INTERGENIC NONCODING RNA p21 in vascular endothelial cells. DNA Cell Biol. 34, 677–683. doi: 10.1089/dna.2015.2966

He, C., Yang, W., Yang, J., Ding, J., Li, S., Wu, H., et al. (2017). Long Noncoding RNA MEG3 negatively regulates proliferation and angiogenesis in vascular endothelial cells. DNA Cell Biol. 36, 475–481. doi: 10.1089/dna.2017.3682

Hidayama, T., and Kwak, S. (2011). When does males start! ADAR2-GluA2 hypothesis for the etiology of sporadic ALS. Front. Mol. Neurosci. 4:33. doi: 10.3389/fnmol.2011.00033

Hoang, N. V., Furtado, A., Mason, P. J., Marquardt, A., Kasirajan, L., Thirugnanasambandam, P. P., et al. (2017a). A survey of the complex transcriptome from the highly polyploid sugarcane genome using full-length isoform sequencing and de novo assembly from short read sequencing. BMC Genomics 18:395. doi: 10.1186/s12864-017-3757-8

Hoang, N. V., Furtado, A., O’keefe, A. J., Botha, F. C., and Henry, R. J. (2017b). Association of gene expression with biomass content and composition in sugarcane. PLoS ONE 12:e0183417. doi: 10.1371/journal.pone.0183417

Hon, C. C., Ramlowski, J. A., Harshberger, J., Bertin, N., Rackham, O. J., Gough, J., et al. (2017). An atlas of human long non-coding RNAs with accurate 5' ends. Nature 543, 199–204. doi: 10.1038/nature21374

Huang, S., Luo, W., Ge, D., Meng, N., Li, Y., Su, L., et al. (2015). A new microRNA signal pathway regulated by long noncoding RNA TGFBR2-OT1 in autophagy and inflammation of vascular endothelial cells. Autophagy 11, 2172–2183. doi: 10.1080/15548627.2015.1106663

Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., et al. (2015). The landscape of long noncoding RNAs in the human transcriptome. Nat. Genet. 47, 199–208. doi: 10.1038/ng.3192

Jaksobsson, L., Franco, C. A., Bentley, K., Collins, R. T., Ponsioen, B., Aspalter, I. M., et al. (2010). Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. Nat. Cell Biol. 12, 943–953. doi: 10.1038/ncb2103
endothelial nitric oxide synthase and endothelial function. Nat. Commun. 9:292. doi: 10.1038/s41467-017-02113-y

Michalik, K. M., You, X., Manavski, Y., Doddballapur, A., Zörnig, M., Braun, T., et al. (2014). Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. Circ. Res. 114, 1389–1397. doi: 10.1161/CIRCRESAHA.114.303265

Ming, G. F., Wu, K., Hu, K., Chen, Y., and Xiao, J. (2016). NAMPT regulates senescence, proliferation, and migration of endothelial progenitor cells through the SIRT1 AS IncRNA/miR-22/SIRT1 pathway. Biochem. Biophys. Res. Commun. 478, 1382–1388. doi: 10.1016/j.bbrc.2016.08.133

Muller, R., Weirick, T., John, D., Milletello, G., Chen, W., Dimmeler, S., et al. (2016). ANGIOGENES: knowledge database for protein-coding and noncoding RNA genes in endothelial cells. Sci. Rep. 6:32475. doi: 10.1038/srep32475

Nakano, K., Shrima, A., Shimoji, M., Tamotsu, H., Ashimine, N., Ohki, S., et al. (2017). Advantages of genome sequencing by long-read sequencer using SMRT technology in medical area. Hum. Cell 30, 149–161. doi: 10.1007/s13577-017-0168-8

Necesula, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U., Picardi, E., D’erchia, A. M., Lo Guidice, C., and Pesole, G. (2017). REDIportal: a comprehensive database of A-to-I RNA editing events from replicate sequencing data. BMC Bioinformatics 18:7. doi: 10.1186/s12859-016-1432-8

Peng, Y., Meng, K., Jiang, L., Zhong, Y., Yang, Y., Lan, Y., et al. (2017). REDItools high-throughput RNA editing detection made easy. Bioinformatics 29, 1813–1814. doi: 10.1093/bioinformatics/btt287

Peng, Z., Cheng, Y., Tan, B. C., Kang, L., Tian, Z., Zhu, Y., et al. (2012). Identification of expressed and conserved human noncoding RNAs. Nat. Methods 9, 579–581. doi: 10.1038/nmeth.1982

Picardi, E., Elia, M., Gatsuou, A., Stamateopoulos, K., Pesole, G., et al. (2013). Long non-coding RNA NRC3R in athersclerosis-related vascular dysfunction. Cell Death Dis. 4:e03523. doi: 10.1038/cddis.2013.111

Picardi, E., and Pesole, G. (2013). REDItools high-throughput RNA editing detection made easy. Bioinformatics 29, 1813–1814. doi: 10.1093/bioinformatics/btt287

Picardi, E., Regna, T. M., Verbitsky, D., Brennicke, A., and Quagliariello, C. (2011). REDItools: an upgraded bioinformatics resource for organellar RNA editing sites. BMC Genomics 12:252. doi: 10.1186/1471-2164-12-252

Picardi, E., Ruíz-Orera, J., Messeguer, X., Subirana, J. A., and Alba, M. M. (2014). A curated database of genetic markers from the human ENCODE RNA-seq data. Cell Rep. 8, 651–656. doi: 10.1016/j.celrep.2014.08.006

Puthanveetil, P., Chen, S., Feng, B., Gautam, A., and Chakrabarti, S. (2015). Long non-coding RNA MALAT1 regulates hyperglycaemia induced process in the human endothelial cells. J. Cell. Mol. Med. 19, 1418–1425. doi: 10.1111/jcmm.12576

Qu, G. Z., Tian, W., Fu, H. T., Li, C. P., and Liu, B. (2016). Long noncoding RNA-MEG3 is involved in diabetes mellitus-related microvascular dysfunction. Biochem. Biophys. Res. Commun. 471, 135–141. doi: 10.1016/j.bbrc.2016.01.164

Ramaswami, G., and Li, J. B. (2014). NAD+–a radically annotated database of A-to-I RNA editing. Nucleic Acids Res. 42, DI09–DI13. doi: 10.1093/nar/gkt996

Ramaswami, G., Lin, W., Piskol, R., Tan, M. H., Davis, C., and Li, J. B. (2012). Accurate identification of human Alu and non-Alu RNA editing sites. Nat. Methods 9, 579–581. doi: 10.1038/nmeth.1982

Ruan, W., Zhao, F., Zhao, S., Zeng, L., Shi, L., and Pang, T. (2018). Knockdown of long noncoding RNA MEG3 impairs VEGF-stimulated endothelial sprouting angiogenesis via modulating VEGFR2 expression in human umbilical vein endothelial cells. Gene 649, 32–39. doi: 10.1016/j.gene.2018.01.072

Ruiz-Oreja, J., Messequer, X., Subirana, J. A., and Alba, M. M. (2014). Long non-coding RNAs as a source of new peptides. Elife 3:e03523. doi: 10.7554/eLife.03523

Sallam, T., Sandhu, J., and Tontonoz, P. (2018). Long noncoding RNA discovery in cardiovascular disease: decoding form to function. Circ. Res. 122, 155–166. doi: 10.1161/CIRCRESAHA.117.311802

Savas, S. (2012). A curated database of genetic markers from the angio genesis/VEGF pathway and their relation to clinical outcome in human cancers. Acta Oncol. 51, 243–246. doi: 10.3109/02841075.2011.636758

Singh, N., Sahu, D. K., Chowdhry, R., Mishra, A., Goel, M. M., Faheem, M., et al. (2016). IsoSeq analysis and functional annotation of the infrarenal ependymoma tumor tissue on PacBio RSII platform. Meta Gene 7, 70–75. doi: 10.1016/j.mgene.2015.11.004

Skelly, D. A., Squiers, G. T., McLellan, M. A., Bolisetty, M. T., Robson, P., Rosenthal, N. A., et al. (2018). Single-cell transcriptional profiling reveals cellular diversity and intercommunication in the mouse heart. Cell Rep. 22, 600–610. doi: 10.1016/j.celrep.2017.12.072

Solomon, O., Eyal, E., Amariglio, N., Unger, R., and Rechavi, G. (2016). e23D: database and visualization of A-to-I RNA editing sites mapped to 3D protein structures. Bioinformatics 32, 2213–2215. doi: 10.1093/bioinformatics/btw204

Stellos, K., Gatsuou, A., Stamateopoulos, K., Pesole, G., et al. (2013). Global regulation of alternative splicing by adenosine deaminase acting on RNA (ADAR). RNA 19, 591–604. doi: 10.1261/rna.038042.112

Weirick, T., John, D., Milletello, G., Chen, W., Dimmeler, S., et al. (2016). Long non-coding RNA NRC3R in atherosclerosis-related vascular dysfunction. Cell Death Dis. 7:e2248. doi: 10.1038/cddis.2016.145

Weirick et al. lncRNAs in Endothelial Cells
