Multi-Omics Approaches to Study Long Non-coding RNA Function in Atherosclerosis

Adam W. Turner, Doris Wong, Mohammad Daud Khan, Caitlin N. Dreisbach, Meredith Palmore and Clint L. Miller

Atherosclerosis is a complex inflammatory disease of the vessel wall involving the interplay of multiple cell types including vascular smooth muscle cells, endothelial cells, and macrophages. Large-scale genome-wide association studies (GWAS) and the advancement of next generation sequencing technologies have rapidly expanded the number of long non-coding RNA (lncRNA) transcripts predicted to play critical roles in the pathogenesis of the disease. In this review, we highlight several lncRNAs whose functional role in atherosclerosis is well-documented through traditional biochemical approaches as well as those identified through RNA-sequencing and other high-throughput assays. We describe novel genomics approaches to study both evolutionarily conserved and divergent lncRNA functions and interactions with DNA, RNA, and proteins. We also highlight assays to resolve the complex spatial and temporal regulation of lncRNAs. Finally, we summarize the latest suite of computational tools designed to improve genomic and functional annotation of these transcripts in the human genome. Deep characterization of lncRNAs is fundamental to unravel coronary atherosclerosis and other cardiovascular diseases, as these regulatory molecules represent a new class of potential therapeutic targets and/or diagnostic markers to mitigate both genetic and environmental risk factors.

Keywords: long noncoding (Inc) RNAs, genomics, cardiovascular disease, atherosclerosis, gene regulation

INTRODUCTION

Despite intensive research into the underlying pathogenesis of atherosclerosis/coronary artery disease (CAD), this disease still remains a significant public health burden. Atherosclerosis is a complex disease involving both environmental and genetic risk factors resulting in plaque formation and inflammation in the vessel wall. A number of cellular responses have been proposed to contribute to disease progression, including endothelial cell dysfunction, vascular smooth muscle cell phenotypic switching, macrophage/foam cell activation/ necrosis/defective efferocytosis, and defective lipid/lipoprotein metabolism (1, 2). Much of the research to date has focused on the role of protein coding genes in atherosclerosis leading to the identification of a number of proteins described as key drivers. However, our understanding of the causal mechanisms of this disease remains limited, likely due to our incomplete functional knowledge of the non-coding genome.
It is now well-established that more than 90% of disease-associated variants reside in non-coding regions, once considered “junk sequences” (3). An increasing number of studies, especially those utilizing high-throughput sequencing technologies, have shown that a number of non-coding RNAs (ncRNAs) are differentially regulated during disease, supporting potential functional roles for these molecules (4).

The formation of the GENCODE project (part of the ENCODE project) and other large-scale initiatives such as NONCODE have revealed 75% of the genome is transcribed, yet only 2% encodes for protein, suggesting alternative functional roles for ncRNA transcripts (5, 6). Since the development of RNA-seq and other high-throughput sequencing assays, ncRNAs are now appreciated as key regulators of gene expression (7). A number of the essential types of non-coding RNAs are already well-characterized such as transfer, spliceosomal, and ribosomal RNAs. Aside from these housekeeping RNA elements, the remaining types of non-coding RNAs are subdivided into two classes, small and long non-coding RNAs, based on their size. Examples of small non-coding RNAs include microRNAs (miRNA), small nucleolar RNAs ( snoRNA), and PIWI-interacting RNAs (piRNA) (8). These elements can act as positive or negative regulators of gene expression and generally exert their influence through complementation base pairing to their target transcript 3′ or 5′ untranslated regions (9).

Long non-coding RNAs (lncRNAs) represent a heterogeneous class of non-coding RNAs that includes transcripts >200 nucleotides, which lack functional protein coding ability (10). Within this lncRNA class, they are also classified based on their genomic location and broadly encompass enhancer-related RNAs (eRNAs), transcribed ultraconserved RNAs, intronic RNAs, long intergenic RNAs (lincRNAs), and natural antisense transcripts (NATs) (10). Contrary to canonical linear lncRNAs, a distinct group of lncRNAs are known as circular RNAs (circRNAs) due to their circular structure, which often results from backsplicing (11). By participating in both transcriptional and post-transcriptional stages, lncRNAs modulate gene expression through multiple distinct mechanisms. Further insight into these regulatory mechanisms will facilitate a better understanding of disease biology and identify additional viable targets for therapeutic intervention or diagnostics. Here, we present an overview of various lncRNAs relevant to atherosclerosis and highlight next-generation sequencing approaches to systematically investigate lncRNA function, as well as the ongoing challenges in this exciting field.

Mechanisms of Long Non-coding RNA Function

LncRNAs are a heterogeneous class of ncRNAs (>200 nucleotides in length) that do not contain a functional open reading frame (12). LncRNAs can be encoded on either the sense or antisense DNA strand and may be located within a protein coding gene or in the intergenic regions (12). Similar to mRNAs, lncRNAs are transcribed by RNA polymerase II. Many transcripts are polyadenylated, multi-exonic, undergo RNA splicing, and contain a 5′ cap. Often, their active promoters are marked with H3K4me3 and gene bodies have H3K36me3 histone modifications (13). Unlike protein coding genes, lncRNAs are not translated into protein and thereby lack functional initiation and termination codons (8). They are expressed at a much lower levels relative to their protein coding counterparts and lack robust evolutionary conservation. Despite this low level of conservation, the expression pattern of lncRNAs has been shown to be relatively cell/tissue specific (14–16).

The mechanism of action of these regulatory factors were categorized into four broad groups proposed by Wang and Chang (17). Signaling lncRNAs represent a class which exhibit a high degree of spatial and temporal specificity that serve a role in signal transduction. Once transcribed, these signaling lncRNAs have effector functions in activating appropriate downstream pathways in response to a stimulus. Additionally, their presence may indicate a particular developmental cell state, condition, or overall transcriptional activity (17). Another mechanism in which lncRNAs exert their regulatory function is by acting as decoy molecules to limit the availability of RNA binding factors to interact with their partners. By impairing the ability of chromatin remodelers, transcription factors, and miRNA from binding to their target genes, decoy lncRNAs can inhibit downstream effector functions (17). It is noteworthy that miRNAs also have the ability to target lncRNAs directly and thereby influence transcriptional regulation and vascular functions (18–20). Aided by their ability to bind protein as well as base pair with target sequences, guide lncRNAs are responsible for localizing transcriptional regulators to specific regions. Similar to the role of guide lncRNAs, scaffold lncRNAs use their protein binding ability to provide a surface to mediate protein-protein interactions (17). In these various ways, lncRNAs represent a distinct class of regulatory elements to modulate transcriptional activities.

OVERVIEW OF LncRNAs INVOLVED INATHEROSCLEROSIS

Functional Studies of LncRNAs inAtherosclerosis

Depending on the cell types involved, lncRNAs play myriad roles in diverse atherosclerotic processes in the vessel wall including cell proliferation, migration, differentiation, apoptosis, and inflammation. They also play important roles in the regulation of cholesterol and lipid metabolism. Pertinent cell types include smooth muscle cells, endothelial cells, macrophages, and hepatocytes (Figure 1). A more comprehensive overview of key lncRNAs in atherogenic processes is given in Table 1. Perhaps the most well-studied in atherosclerosis is CDKN2B-AS1 or ANRIL (Antisense Non-coding RNA in the INK4 Locus), which acts in several cell types relevant to CAD (21, 22, 81). ANRIL acts as a guide lncRNA to localize polycomb repressive complex (PRC) at target promoters through a direct interaction with its subunits, CBX7 or SUZ12. PRC then adds H3K27me3 modifications to this region to repress transcription (22). Loss of function studies have suggested ANRIL acts in cis in order to regulate transcription of the nearby tumor suppressors, CDKN2A and CDKN2B (82, 83). Consistent with
Turner et al. Multi-Omics and lncRNAs

FIGURE 1 | Schematic of atherosclerotic processes and specific lncRNA functions. Top, lncRNAs are shown with described smooth muscle cell (SMC) functions, such as proliferation, apoptosis, autophagy, phenotypic switching, and differentiation. lncRNAs are also shown with endothelial cell (EC) functions such as differentiation, regulation of endothelial nitric oxide synthase (eNOS) mediated signaling, growth and angiogenesis. lncRNAs are shown with macrophage functions, such as macrophage polarization, cholesterol efflux, and inflammation. Also, lncRNAs are listed with functions in regulating cholesterol and triglyceride metabolism in hepatocytes and/or macrophages. Bottom, schematic showing example of atherosclerotic lesion after invasion of vascular endothelium by activated monocytes, which become macrophages upon chronic inflammatory stimulation. Exposure to oxidized LDL (oxLDL) particles promote macrophage transformation to lipid-laden foam cells. Also depicted is the transformation of contractile SMCs to de-differentiated or modulated SMCs, as well as the transition of modulated SMCs to macrophage-like cells in the lesion. ECM, Extracellular matrix.

1. **SMC apoptosis, autophagy**

   - ANRIL
   - GAS5
   - HIF1α-AS1
   - LncRNA-p21

2. **SMC proliferation, phenotypic switching**

   - ANRIL
   - MEG3
   - AK098656
   - MIR122HG
   - CASC15
   - NEAT1
   - H19
   - PACER
   - LncRNA-p21
   - SENCPR
   - MALAT1
   - SMILR

3. **EC function, signaling**

   - CLDN10-AS1
   - CTC-4596.1
   - GATA6-AS
   - LEENE
   - sONE/ATG9B
   - EC growth, angiogenesis
   - NRON
   - TIE1-AS
   - STEEL

4. **Macrophage function, inflammation**

   - ANRIL
   - Dnm3os
   - GAS5
   - LncRNA-21
   - LncRNA-Mirt2
   - MeXIS
   - NEAT1
   - NEXN-AS1

Legend:

- EC
- SMC
- modulated SMC
- monocyte
- macrophage
- foam cell
- oxLDL
- ECM

---

this mechanism of action, ANRIL expression correlates with a more proliferative phenotype in endothelial cells and vascular smooth muscle cells (VSMC) (22, 84, 85). In addition to acting in *cis*, ANRIL also acts in *trans* (via Alu elements) to regulate other genes that participate in proatherogenic pathways (22). Since ANRIL is not well-conserved in mice, *in vivo* functional studies have been challenging (86). A more complete overview of the atherogenic roles of ANRIL RNA species has been recently documented in a review by Holdt and Teupser (87).

The ubiquitously expressed and evolutionarily conserved lncRNA, MALAT1, is decreased in atherosclerotic plaques (88, 89), and reduced MALAT1 levels in hematopoietic cells promotes atherosclerosis and inflammation in mice *in vivo* (89). In contrast, MALAT1 knockdown in VSMC and EC results in cell cycle arrest and reduced proliferation (35, 36). Other lncRNAs (e.g., lncRNA-p21, HIF1α-AS, circular ANRIL, and GAS5) have been implicated in cell death/apoptotic pathways through various mechanisms (23, 29, 33, 90), as described elsewhere (91, 92). In particular, lncRNA-p21 was shown to be reduced in CAD patients and mouse models of atherosclerosis, and regulates p53-dependent smooth muscle cell proliferation and apoptosis (33). Many lncRNAs have established immune and inflammatory roles. For example, heterozygous MALAT1-deficient ApoE/- mice have increased inflammation and atherosclerosis (93). LncRNA Dnm3os (dynamin 3 opposite strand) is upregulated in diabetic induced macrophages and regulates nucleolin and epigenetic mediated inflammatory responses (73). Finally, in human macrophages treated with oxidized LDL, HOTAIR regulates oxidative stress and inflammation (94).

There are several lncRNAs with regulatory roles in lipid and cholesterol metabolism. CHROME (cholesterol homeostasis regulator of miRNA expression) is a lncRNA upregulated in carotid plaques, which regulates cholesterol homeostasis in primates in liver and macrophages by inhibiting miRNAs, such as miR-33 (49). NEAT1 promotes pro-atherogenic functions in THP-1 human macrophage cells such as increased ox-LDL lipid accumulation and inflammation by serving as a sponge of miR-342-3p target (44). Finally, differential expression of TRIBAL, APOA1-AS, and LncLSTR is linked to defects in lipid metabolic pathways, mainly in the liver. TRIBAL (TRIB1 associated locus) regulates Trib1 mRNA stability through mitogen activated kinase, consistent with Trib1 regulation (80, 95, 96). Increased TRIBAL expression stabilizes Trib1 expression and upregulates fatty acid oxidative pathways (80). Likewise, LncLSTR (liver-specific triglyceride regulator) regulates plasma triglyceride clearance by modulating apolipoprotein C2 (APOC2) levels and lipoprotein lipase activity (79). APOA1-AS regulates cholesterol levels through epigenetic modulation of APOA1, a protein involved in the cholesterol efflux pathway (77).
| Name | Full name | Initial detection method(s) | Cell types/tissues relevant to CAD | Function relevant to CAD | References |
|------|-----------|-----------------------------|-----------------------------------|--------------------------|------------|
| ANRIL | A new large antisense non-coding RNA | Long range PCR and nucleotide sequencing | SMC, EC, Mac | Cell cycle regulation, Regulation of apoptosis | (21–24) |
| GAS5  | Growth arrest specific 5 | Characterization of cDNA library | SMC, EC, Mac | Regulates apoptosis, Regulates autophagy, Regulates smooth muscle differentiation | (25–28) |
| HIF1α-AS1 | Hypoxia-inducible factor 1-alpha Antisense RNA 1 | RIP, qRT-PCR | SMC | Regulation of VSMC apoptosis | (29–32) |
| LincRNA-p21 | Long intergenic non-coding RNA at p21 locus | Nimblegen lincRNA tiling microarray platform | SMC, Mac | Regulation of cell proliferation, Regulation of apoptosis | (33, 34) |
| MALAT1 | Metastasis associated lung adenocarcinoma transcript 1 | Identified in numerous physiological processes | SMC, EC, Mac | Regulation of cell proliferation, Regulation of inflammation, Regulation of autophagy | (35–40) |
| MEG3  | Maternally expressed 3 | Characterization of cDNA library | EC, SMC | Regulates endothelial cell proliferation, Inhibits angiogenesis, Regulates smooth muscle proliferation/migration | (41–43) |
| NEAT1 | Nuclear paraspeckle assembly transcript 1 | Affymetrix expression array | SMC, Mac | Regulation of VSMC phenotypic switching, Macrophage polarization | (44–47) |
| NEXN-AS1 | Nexilin antisense transcript 1 | lncRNA array | EC, SMC | Regulation of adhesion molecules/monocyte recruitment, Regulation of inflammatory cytokines | (48) |
| CHROME | Cholesterol homeostasis regulator of miRNA expression | Characterization of transcript proximal to CAD and plasma HDL-C associated locus | Mac, Liver | Regulation of cholesterol homeostasis | (49) |
| AK098656 | AK098656 | lncRNA array | SMC | Regulation of VSMC phenotypic switching | (50) |
| CASC15 | Cancer susceptibility 15 | Quantification of copy number gains in metastatic melanoma | SMC | VSMC stiffness | (51, 52) |
| H19   | H19 | cDNA characterization | SMC | Regulation of VSMC differentiation, Let 7 miRNA sponge | (53–55) |
| MIR122HG/Lnc-Ang362 | MicroRNA 122 Host Gene | RNA-seq | SMC | Regulation of VSMC proliferation | (56) |
| MYOSLID | MYOcardin-induced Smooth muscle LncRNA, Inducer of Differentiation | RNA-seq | SMC | Regulation of SMC differentiation | (57) |
| PACER | P50-associated COX-2 extragenic RNA | ANalysis of ChiP data, RT-qPCR | SMC | Regulation of COX-2 expression, VSMC stiffness | (51, 58) |
| SENCR | Smooth muscle and Endothelial cell-enriched migration/differentiation-associated long Non-coding RNA | RNA-seq | SMC | Regulation of myocardin, SMC contractile gene program | (59) |
| SMILR | Smooth muscle-induced lncRNA enhances replication | RNA-seq | SMC | Regulation of VSMC proliferation | (60) |
| CLDN10-AS1 | Claudin 10 antisense transcript 1 | lncRNA microarray | EC | Regulation of endothelial signaling | (61) |

(Continued)
| Name          | Full name                                      | Initial detection method(s) | Cell types/tissues relevant to CAD | Function relevant to CAD                                                                 | References |
|---------------|-----------------------------------------------|-----------------------------|-----------------------------------|-----------------------------------------------------------------------------------------|------------|
| CTC-4596.1    | RASGRF2 antisense RNA 1                       | IncRNA microarray           | EC                                | Regulation of endothelial signaling                                                     | (61)       |
| GATA6-AS      | GATA6 antisense                               | RNA-seq                     | EC                                | Regulation of endothelial signaling                                                     | (62)       |
|               | IncRNA that enhances eNOS expression          | RNA-seq, chromosome         | EC                                | Regulation of eNOS and endothelial function                                              | (63)       |
| MIAT          | Myocardial infarction associated transcript    | Isolation of specific cDNA, RACE | EC                                | Regulation of angiogenesis                                                               | (64, 65)   |
| NRON          | Non-protein coding RNA, repressor of NFAT     | Screen of cDNA library to identify conserved IncRNAs shRNA screen to identify NFAT regulators | EC                                | Regulation of angiogenesis and vascular development                                       | (66-68)    |
| sONE/ATG96B   | Autophagy related 9B                          | cDNA characterization        | EC                                | Regulation of eNOS                                                                      | (69)       |
| STEEL         | Spliced transcript endothelial-enriched IncRNA | Custom IncRNA microarray    | EC                                | Angiogenesis                                                                            | (70)       |
| TIE1-AS       | Endothelial-specific receptor tyrosine kinase 1 antisense | Detection in EST libraries, RACE | EC                                | Regulation of vascular development                                                        | (71, 72)   |
| Dnm3os        | Dynamin-3 opposite strand/antisense RNA       | Isolation of clones, RACE   | Mac                               | Regulation of macrophage inflammation                                                   | (73, 74)   |
| IncRNA-MirT2  | IncRNA Myocardial infarction associated transcript | IncRNA expression microarray | Mac                               | Negative regulator of inflammation                                                       | (75)       |
| MeXIS         | Macrophage-expressed LXR-induced sequence     | RNA-seq                     | Mac                               | Regulation of cholesterol metabolism                                                     | (76)       |
| APOA1-AS      | Apolipoprotein A1 antisense transcript         | Identification in EST library, RACE | Liver                             | Regulation of APOA1 expression                                                          | (77)       |
| LeXis         | Liver-expressed LXR-induced sequence           | RNA-seq                     | Liver                             | Regulation of cholesterol metabolism                                                     | (78)       |
| IncLSTR       | IncRNA Liver-Specific Triglyceride Regulator   | Characterization of imprinted genes | Liver                             | Regulation of triglyceride metabolism                                                   | (79)       |
| TRIBAL        | Tribbles homolog 1-associated locus            | Identification in EST library, RACE | Liver                             | Regulation triglyceride metabolism                                                      | (80)       |

Color scheme: Gray, lncRNAs associated with multiple cell types/tissues. Purple, lncRNAs associated with smooth muscle cells (SMC). Green, lncRNAs associated with endothelial cells (EC). Yellow, lncRNAs associated with macrophages (Mac). Blue, lncRNAs associated with cholesterol metabolism in liver. RACE: Rapid amplification of cDNA ends, EST: Expressed sequence tag.

LncRNAs With Genetic Associations in Atherosclerosis

Genome wide association studies (GWAS) have linked genetic variation at the ANRIL locus (9p21.3) to many complex phenotypes including CAD, stroke, type 2 diabetes and multiple cancers (97). In addition to ANRIL, the 9p21.3 locus encodes three tumor suppressor proteins: CDKN2A, CDKN2B, and MTAP. Despite each being attractive candidates underlying the locus association with various diseases, several studies report CAD risk polymorphisms associated with ANRIL expression (21, 98). However, association studies of 9p21.3 genotype with ANRIL expression remain complex due to the numerous linear and circular ANRIL forms (23).

Other less studied lncRNAs have been identified from genetic studies of CAD or related traits and may play critical roles in atherosclerosis. For instance, genetic variation in the imprinted IncRNA H19, involved in embryonic development (99) and oncogenesis (100), was associated with CAD and ischemic stroke in Chinese populations (101, 102). H19 was initially shown to be re-expressed in smooth muscle cells in human
and rodent atherosclerotic plaques (103), and promotes VSMC proliferation by acting as a let-7a miRNA sponge to upregulate cyclin D (104). However, a recent study revealed endothelial cell restricted expression in human atherosclerotic plaques and a role in endothelial cell aging by suppressing STAT3 signaling (105), similar to lncRNA MEG3 (106). Another endothelial cell lncRNA, MIAT (Myocardial Infarction Associated Transcript), was previously associated with myocardial infarction in a large genetic study of a Japanese population (64). MIAT is upregulated in atherosclerosis plaques (88), and regulates microvascular dysfunction by acting as a competing endogenous RNA (65). Another lncRNA associated with CAD through large-scale GWAS is known as TARID (TCF21 antisense RNA inducing promoter demethylation) (107). TARID was identified as an eQTL target gene in human coronary artery smooth muscle cells (108), and molecular studies suggest this lncRNA guides GADD45A mediated DNA demethylation and inactivation of TCF21 (109), a known tumor suppressor and vascular wall transcription factor associated with CAD (110–113). Yet, functional studies of TARID both in VSMCs and in vivo are needed to elucidate its potential role in atherosclerosis. With larger GWAS sample sizes, and complementary eQTL colocalization (114), and transcriptome-wide association studies (TWAS) (115), it is anticipated that even more lncRNAs will be identified with genetic association evidence.

APPLICATION OF TRANSCRIPTOMICS TO IDENTIFY AND STUDY LncRNAs RELEVANT TO CAD

General Considerations for Transcriptomics Studies of LncRNAs

While traditional methods to profile lncRNA transcriptomes have relied on microarrays or serial analysis of gene expression (SAGE), these approaches have largely been replaced with the decreasing costs and greater output achieved by RNA-seq (116). In general, RNA-seq provides greater sensitivity and specificity to detect a broad range of ncRNA transcripts, novel isoforms, and interactions between ncRNAs (117). Nonetheless there are some important considerations when designing and conducting RNA-seq based lncRNA screening experiments. For instance, given that lncRNAs are approximately 10X less abundant than mRNAs on average, the basal expression of a typical lncRNA is < 5 fragments per kilobase of transcript per million mapped reads (FPKM) (118). Thus, it is highly recommended to obtain deeper sequencing per sample (~100X read depth) than a typical RNA-seq experiment. Also, while up to 50% of lncRNAs appear to be poly-adenylated (119) and would be detected with mRNA library preparation kits, a more comprehensive landscape of lncRNAs, other ncRNAs, including eRNAs, would require total RNA [poly(A) and non-poly(A)], ribosomal RNA depletion methods of purification. Distinguishing lncRNA transcripts from mRNA transcripts from short-read sequencing data is a challenge, however deeper, paired-end and stranded sequencing should improve identification of lncRNAs (120). Also, careful study design is needed to ensure sufficient power to detect differentially expressed and transcript-specific lncRNAs, when using standard count-based tools (120). Since many lncRNAs are tissue and cell-specific (121–123), it is also worth considering the effects of diluting weak signals from bulk populations of cells, as well as specific environmental contexts that may regulate lncRNA transcript levels. Below, we summarize recent findings of RNA-seq based lncRNA discoveries in specific cell types relevant to CAD/atherosclerosis.

Transcriptomics of Vascular Smooth Muscle Cell Function

The first VSMC lncRNA discovered via RNA-seq was Lnc-Ang362 (HG-MIR222), which is upregulated in rat aortic smooth muscle cells upon stimulation with angiotensin II (56). Lnc-Ang362 promotes VSMC proliferation and is the host-transcript for both mir-221 and mir-222. Bell et al. conducted RNA-seq in human coronary artery smooth muscle cells and identified 31 previously unidentified lncRNAs (59). Notably, one of these was Smooth muscle and Endothelial cell-enriched migration/differentiation-associated long Non-coding RNA (SENCR), which is located antisense to the FLI1 gene. SENCr functionally promotes a contractile smooth muscle phenotype and inhibits migration (59). In a follow-up study, RNA-seq was performed in human coronary artery smooth muscle cells to examine the effect of myocardin (MYOC) overexpression (57). MYOD is a potent co-factor that binds with serum response factor (SRF) to activate an array of smooth muscle-specific genes that maintain smooth muscle cell differentiation (124–128). Over 100 lncRNAs were differentially expressed, one of which was identified as MYOCardin-induced Smooth muscle LncRNA, Inducer of Differentiation (MYOSLID). Functional studies demonstrated that MYOSLID, a direct transcriptional target of MYOCDF/SRF, promotes smooth muscle differentiation and inhibits proliferation (57).

Yu et al. used RNA-seq to compare transcriptomes of coronary and aortic smooth muscle cells subjected to both normal and pathological aortic stiffness, a subclinical risk factor for CAD and various aortic diseases (51). Only two of the top 20 ranked differentially expressed lncRNAs have been studied to date: CASC15 and PACER (RP5-973M2.2). These lncRNAs regulate expression of protein-coding genes in cis and PACER activates COX2 expression (52, 58, 129). Analysis of RNA-seq data highlighted the lncRNA MALAT1 as a key regulator of VSMC stiffness-induced proliferation and migration. Although MALAT1 was originally described as an endothelial lncRNA, MALAT1 regulates the phenotyping switching of VSMCs via activation of the autophagy pathway (36). Using RNA-seq in human smooth muscle cells Ballantyne et al. identified over 300 differentially expressed lncRNAs upon platelet-derived growth factor and interleukin-1 alpha stimulation. The novel lncRNA, Smooth Muscle-Induced LncRNA enhances Replication (SMILR) identified from this study enhances smooth muscle cell proliferation and has increased expression in unstable atherosclerotic plaques (60). The lncRNA NEAT1 (nuclear paraspeckle assembly transcript 1) has recently been implicated in promoting the phenotypic switching of VSMCs (45). RNA-seq
demonstrated NEAT1 silencing increases the mRNA levels of numerous critical smooth muscle cell marker genes. Finally, to identify IncRNAs key in smooth muscle cell differentiation, Lim et al. combined and queried diverse RNA-seq datasets from Gene Expression Omnibus (GEO). Dozens of IncRNAs with no previous evidence for roles in VSMC differentiation were identified in this analysis that warrant further investigation, either as cis transcriptional regulators or suppressing miRNA function (130).

The development of custom IncRNA arrays has been applied to identify IncRNAs involved in various processes critical in atherosclerosis. One example is a microarray analysis which identified 580 IncRNAs differentially expressed upon exposure of human aortic smooth muscle cells to cyclic mechanical stretch (131). Another example is identification of AK098656, predominantly expressed in VSMCs, also upregulated in hypertensive patients and involved in promoting a synthetic smooth muscle cell phenotype (50).

Transcriptomics of Endothelial Cell Function

Although not all IncRNAs have a poly(A) tail, Michalik et al. performed deep sequencing of poly(A)-selected RNA in human umbilical vein endothelial cells (HUVECs) and found over half of total RNA composed of non-coding RNA, many of which are IncRNAs (35). This study focused on five IncRNAs with high endothelial expression and strong conservation between mice and humans: MALAT1, linc00493, maternally expressed 3 (MEG3), taurine upregulated gene 1 (TUG1), and linc00657. MALAT1 and MEG3 are strongly upregulated in response to hypoxia while linc006757 are TUG1 are moderately upregulated. In regards to angiogenesis, MALAT1 promotes angiogenesis and induces a switch of endothelial cells from a migratory cell phenotype to a proliferative cell phenotype (132). Huang et al. postulated exosomal MALAT1 from oxidized LDL (oxLDL) treated endothelial cells (HUVECs) promotes macrophage polarization toward the M2 phenotype (133). MEG3 was shown to interact with epigenetic modifiers, to inhibit angiogenesis and contribute to age-related endothelial dysfunction (106, 134, 135).

In another study Miao et al. conducted RNA-seq profiling of endothelial cells subjected to both physiological and pathological flow for various time points (63). They identified and characterized LEENE (IncRNA that enhances eNOS expression) as a lncRNA highly correlated with endothelial nitric oxide synthase (eNOS) expression levels, which is downregulated upon pathological flow (63). Several lncRNAs characterized in smooth muscle cells also have functional significance in endothelial cells. For instance, the SMC IncRNA SENCR regulates the differentiation of pluripotent cells into endothelial cells and promotes angiogenesis in HUVECs (136).

Custom IncRNA microarrays have identified endothelial cell enriched IncRNAs (70) and endothelial IncRNAs differentially expressed in response to specific treatments (e.g., oxidized LDL) (61). These arrays have revealed new candidate lncRNAs in atherosclerosis including spliced-transcript endothelial-enriched IncRNA (STEEL) (70), NEXN-AS1 (48), CLDN10-AS1, and CTC-45916.1 (61). STEEL facilitates the transcriptional stimulation of both eNOS and Kruppel-like factor (KLF2) and in vivo promotes angiogenesis (70). STEEL expression is decreased in HUVECs exposed to “atheroprotective” flow and expression is increased in HUVECs exposed to disturbed “atheroprone” flow.

Transcriptomics of Macrophage Function and Inflammation

In macrophages, LXR activation promotes cholesterol efflux through activation of target genes such as Abca1 during the formation of HDL. To investigate the regulation of LXR-dependent transcription in macrophages, a recent study conducted large-scale transcriptional profiling of mouse peritoneal macrophages in response to the LXR agonist GW3965. LXR activation stimulated transcription of an array of IncRNAs, of which MeXis was among the strongest induced (76). MeXis is well-conserved in mice and was shown to amplify the LXR-dependent expression of Abca1 in vivo and promote cholesterol efflux in macrophages (76). Loss of MeXis in Ldlr−/− mice was shown to accelerate atherosclerosis through impaired Abca1 expression in macrophages and resulted in decreased cholesterol efflux (76). ATAC-seq in peritoneal macrophages demonstrated decreased chromatin accessibility across the Abca1 locus in response to loss of MeXis. Querying the MeXis interactome through mass spectrometry revealed protein interactions with the nuclear receptor coactivator DDX17. Either directly or indirectly through one of its interacting targets, MeXis represents a potential therapeutic target to regulate macrophage cholesterol efflux.

RNA-seq and IncRNA arrays have identified a number of other macrophage IncRNAs that could represent novel CAD targets. Zhang et al. performed deep RNA sequencing of human monocyte-derived macrophages as well as M1 activated (via interferon gamma and lipopolysaccharide stimulation) and M2 activated (via interleukin 4 stimulation) macrophages (137). This study identified 861 previously unannotated lncRNAs, most of which are not syntenic in mouse. Furthermore, the lncRNA expression profile is dramatically shifted upon M1 activation, supporting the inflammatory nature of atherosclerosis. Similarly, 109 unannotated CD14+ monocyte lncRNAs were highlighted upon exposure to inflammatory stress in vivo (138). Other recent array studies highlighted the macrophage IncRNAs Dnm3os and Mirt2. Dnm3os is upregulated in bone marrow derived macrophages in diabetic mice compared to controls and is higher in monocytes in human type 2 diabetic patients compared to controls (73). Dnm3os alters global histone modifications in macrophages and upregulates various immune-response and inflammatory genes. LncRNA-Mirt2 is strongly induced by LPS, a toll-like receptor 4 (TLR4) ligand where it acts as a negative feedback inflammatory regulator (75).

Transcriptomics of Cholesterol Metabolism and Hepatocyte Function

Liver X receptors (LXRs) are nuclear factor transcription factors that are important mediators of lipid and cholesterol metabolism. LXR targets include the ABC family of transporters, ApoE, LPL, and SREBP (139, 140). Liver-specific LXR alpha knockout mice develop increased cholesterol levels and atherosclerosis (141). Sallam et al. performed genome-wide transcriptional profiling...
of primary mouse hepatocytes upon stimulation with an LXR agonist (78). The strongest induced gene was a non-coding RNA termed LeXis (liver-expressed LXR-induced sequence) that lies adjacent to the Abca1 gene. LeXis regulates several genes with roles in cholesterol biosynthesis, subsequently altering both liver and plasma cholesterol levels. Mass spectrometry was used to characterize the LeXis interactome and revealed binding to RALY, a ribonucleoprotein that acts as a transcriptional cofactor in regulation of cholesterol biosynthetic genes (78). In the context of atherosclerosis, adenoviral overexpression of LeXis in the liver reduces atherosclerosis in a familial hypercholesterolemia mouse model (142). As discussed above, CHROME is another LXR-regulated lncRNA involved in cholesterol homeostasis (49). CHROME was first identified through a combination of genetic association studies for premature CAD and HDL-C and microarray based expression profiling in human atherosclerotic plaques (49). RNA-seq of control and CHROME shRNA treated HepG2 hepatocytes revealed downstream pathways affected, including the LXR pathway, bile acid metabolism, cholesterol excretion and fatty-acid β-oxidation pathways (49).

APPLICATION OF NOVEL GENOMIC TECHNOLOGIES TO DETECT AND STUDY LncRNA FUNCTIONS

Novel Sequencing Technologies to Discover and Annotate Long Non-coding RNAs

Although next-generation sequencing has resulted in the identification of thousands of lncRNAs in the genome, many of these lncRNAs remain poorly characterized and annotated. It is often unclear where transcription begins and which exons are present in a particular isoform. Since lncRNAs are often expressed at lower levels compared to protein-coding genes, current transcriptomic data is unable to provide comprehensive mapping/characterization of isoforms. However, new sequencing technologies allow for better characterization due to longer read lengths, higher sensitivity, and higher accuracy. Techniques such as Iso-Seq (Pacific Biosystems) offer long-read sequencing using single-molecule, real-time (SMRT) sequencing, in which the sequence of a full-length transcript is captured in a single read (143). Despite these benefits, these single-molecule sequencers yield higher error rates compared with short read sequencing technologies (e.g., Illumina). Nanopore technologies such as the MinION instrument (Oxford Nanopore Technologies) also allow single cDNA molecules to be sequenced without the need for amplification, providing sufficient read lengths to cover the full-length non-coding RNA, and results in less bias than other long-read approaches (144). This technique passes nucleic acids through an orifice $10^{-9}$ m in diameter, where instrumental electric current changes are utilized to decipher the identity of each nucleotide (145).

Since lncRNAs are typically less abundant than protein coding genes (usually one order of magnitude less), they remain a challenge to study in bulk transcriptomic datasets. To improve the detection and annotation of lncRNAs, a method known as RACE (Rapid Amplification of cDNA Ends)-Seq was developed (146), however this approach was limited by its low-throughput. Later a technique called RNA CaptureSeq was developed to enrich for long non-coding RNAs (147). RNA CaptureSeq employs an array of oligonucleotide probes to capture select genes of interest, which can be applied to pull-down lncRNAs of interest (148, 149). More recently the GENCODE consortium improved upon RNA CaptureSeq by developing RNA Capture Long Seq (RNA CLS) with the goal of annotating lncRNAs with much higher confidence (150). RNA CLS overcomes the short-read length hurdle of RNA CaptureSeq by first capturing lncRNAs and then integrating with long-read sequencing.

DNA-Based LncRNA Interactions

Despite their low abundance, lncRNAs are known to function through specific molecular interactions with other RNA species and RNA binding proteins. Several high-throughput methods are now available to uncover the genomic DNA sequences that lncRNAs interact with and likely regulate (Figure 2A). Chromatin Isolation by RNA Purification (ChIRP-Seq) is a well-established technique to study lncRNA-chromatin interactions through RNA/chromatin crosslinking, purification using biotinylated antisense oligonucleotides, followed by high-throughput sequencing (151, 152). Domain-specific ChIRP (dChIRP) is a variation of ChIRP that can characterize lncRNA function and architecture at the RNA domain level (153). dChIRP can not only investigate lncRNA-chromatin interactions but also pairwise lncRNA-RNA and lncRNA-protein interactions.

Capture hybridization analysis of RNA targets (CHART) (154, 155) is a similar method to experimentally determine where lncRNAs target and localize in the genome (Figure 2A). In the CHART protocol chromatin is crosslinked and lncRNAs subsequently hybridized to biotinylated C-oligos. After bead immobilization of lncRNA/DNA complexes, sequencing is conducted to identify lncRNA binding DNA regions.

GRID-seq (global RNA interactions with DNA by deep sequencing) is a new unbiased method to capture global RNA-interactions (Figure 2A) that can be applied to investigate lncRNA-DNA interactions in cell lines relevant to atherosclerosis (156). This GRID-seq technique uses a bivalent linker consisting of double-stranded DNA and single-stranded RNA to link RNAs with DNA in nuclei that have been fixed. Finally, MARGI (mapping RNA-genome interactions) is a high-throughput method that can be performed in vivo or on cells and reveal the genomic target sites of lncRNAs (157).

Protein-Based LncRNA Interactions

ChIRP-MS is an adaptation of the ChIRP protocol and used to characterize the interacting proteome for a lncRNA (158). ChIRP-MS has identified protein interactors for lncRNAs such as LeXis, MeXis, and AK098656 (50, 76, 78). lncRNA pull-down followed by mass spectrometry has been conducted for several lncRNAs with potential roles in CAD such as circANRIL (23), STEEL (70), MALAT1
Turner et al. Multi-Omics and lncRNAs

FIGURE 2 | Genomic approaches to capture lncRNA interactions. (A) DNA-based lncRNA interactions include Chromatin Isolation by RNA Purification (ChIRP) and Capture Hybridization of RNA Targets (CHART). An in situ based method to capture Global RNA Interactions with DNA (GRID) followed by deep sequencing uses a biotinylated bivalent linker to ligate RNA and dsDNA. (B) Protein-based lncRNA interactions include RNA Immunoprecipitation (RIP) which uses an antibody against RNA binding protein (RBP) to capture RNA-protein interactions. Cross-linking Immunoprecipitation (CLIP) combines UV cross-linking with immunoprecipitation to capture RNA-protein interactions. Targets of RNA-binding proteins Identified By Editing (TRIBE) couples an RBP to an RNA editing enzyme (ADAR). Targets of RBP are marked by adenosine to inositol RNA editing events and identified by sequencing. (C) RNA-based lncRNA interactions include RNA Antisense Purification, which uses a biotinylated probe to capture interacting RNAs that could be followed with sequencing or mass spectrometry. Ligation of interacting RNA (LIGR) followed by sequencing is a powerful approach to capture lncRNA-RNA interactions by in vivo crosslinking of RNA duplexes using the psoralen derivative 4'-aminomethyltrioxalen (AMT) and UV irradiation at 365 nm.

RNA-Based LncRNA Interactions

RNA-centric RNA antisense purification (RAP) is a general approach to identify and study lncRNA functions (Figure 2C). This method uses long capture probes (120 nucleotides) tiled across an entire RNA sequence to pull down lncRNAs, followed by stringent wash conditions to reduce non-specific binding (169). There are now next-generation sequencing derived methodologies that have been established to better define RNA-RNA interactions. LIGR-seq (LIGATION of interacting RNA followed by high-throughput sequencing) can capture base-paired RNA-RNA interactions (Figure 2C) (170). In LIGR-seq, RNA duplexes are cross linked with the psoralen derivative 4'-aminomethyltrioxalen (AMT) along with UV irradiation at 365 nm, and RNase R is added to digest linear and structural RNAs. This step enriches for AMT-crosslinked RNA-RNA duplexes that are subsequently subjected to next-generation sequencing. Though LIGR-seq does not work well for small RNAs such as microRNA (miRNA), it should be able to uncover novel dynamic and long range interactions between lncRNAs and other RNA molecules. Various other methods have been developed to study the RNA interactome for lncRNAs with functional relevance in atherosclerosis including PARIS (Psoralen Analysis of RNA Interactions and Structures) (171), SPLASH (172), and MARIO (173). These techniques all can provide valuable information because many
IncRNA sequence and structural motifs act as functional scaffolds in the assembly of RNA-protein complexes (17). However, it should be noted that many of these assays could be biased toward capturing stable interactions, while more transient and stimulation specific interactions may require some enrichment steps.

**In situ Hybridization-Based Methods**

A critical consideration when interrogating a given IncRNA function, is identifying its endogenous tissue and cellular localization. While many IncRNAs are expected to be cytosolic and contribute to post-transcriptional, translational or post-translational gene regulation, nuclear IncRNAs could participate in transcriptional regulation, chromatin structure or mRNA export mechanisms (17). RNA Fluorescence in situ Hybridization (FISH) has been a traditional method to identify the subcellular localization of RNA within cells, however it lacks sensitivity for lowly expressed IncRNAs. Single-molecule RNA FISH (smFISH) is a quantitative technique that provides the sensitivity to detect these IncRNAs and measures absolute transcript levels by using multiple short probes per target RNA (174). However, given that smFISH relies heavily on the optical detection of a limited number of fluorophores, it is restricted in its multiplexing capacity. Attempts to overcome this issue include implementation of combinatorial labeling by spectral barcodes and the incorporation of sequential hybridizations (seqFISH) using different colored probes in each hybridization round (175, 176). In seqFISH individual transcripts are imaged as different colored dots and quantified by counting the number of dots. Multiplex error-robust combinatorial labeling (merFISH) is an in situ targeted approach that utilizes two-step labeling and the detection of binary barcodes assigned to specific targets. This is accomplished by several rounds of hybridization, imaging, and cleavage of fluorophores from probes conjugated to readout sequences that interchange each cycle. Hybridization to readout sequences by the merFISH technique is much less time consuming than methods that utilize hybridization directly to target RNAs (177, 178). RNA SPOTS (sequential probing of targets) follows the same rationale as merFISH, except that it is used in vitro instead of in situ (179).

While still in the nascent stage, the emergence of spatial transcriptomics facilitates integration of RNA-seq expression data with spatial locations of RNA molecules in individual tissue sections (180). In this procedure, fixed tissue samples are annealed to regionally barcoded reverse transcription primers. Following reverse transcription, RNA-seq followed by computational reconstruction allows the two-dimensional localization and quantification of RNA molecules (180). This barcoded method has already been applied to spatially resolve gene expression in the human adult heart (181). While this procedure was originally developed to study mRNAs, it shows promise for the spatial resolution of IncRNAs, given the increased sensitivity and ability to identify context-specific expression profiles. One consideration for atherosclerosis FISH experiments, is that heterogeneous cell types in lesions may be impacted differently by various fixation and hybridization conditions, so careful titration of reagents is recommended.

**Other LncRNA Functions**

Finally, there are various omics methods that can define the dynamics of IncRNA transcription, stability and RNA modifications. Nascent RNA sequencing analysis, including global nuclear run-on sequencing (GRO-seq) and precision run-on sequencing (PRO-seq) assays, could enable comprehensive detection of transient RNA transcriptional events for multiple RNA species, including mRNA, IncRNA, and eRNA (182). While most transcriptomic datasets capture steady-state levels of IncRNA transcripts, they do not provide direct insights into the stability of IncRNAs. BRIC-seq (5’-bromo-uridine immunoprecipitation chase-deep sequencing analysis) is a method that pulse-labels endogenous RNAs and employs next-generation sequencing to measure RNA decay over time (183). Total RNAs (including IncRNAs) can be isolated from cells at desired time points under various cell-specific perturbations to facilitate functional analysis of IncRNA stability (e.g., IncRNA related to CAD). For example, direct measurements of IncRNA stability in response to CRISPR based loss/gain of gene function or drug treatments could be examined.

Another technique, ICE-seq (inosine chemical erasing coupled with sequencing) (184) represents a promising approach to globally identify IncRNA adenosine to inosine modifications (e.g., in the context of atherosclerosis). Adenosine to inosine (A-to-I) RNA editing is the most abundant form of RNA editing in humans and results from adenosine deaminase acting on RNA (ADAR). A-to-I editing is common to all IncRNAs and affects IncRNA function through altered stability and target recognition (185, 186). A-to-I RNA editing of mRNA has already been demonstrated to have important functional consequences in atherosclerosis. For example, A-to-I editing of cathepsin S mRNA (CTSS) is associated with cathepsin S levels in patients with atherosclerosis. Treatment of endothelial cells with inflammatory cytokines or exposure to hypoxia was shown to induce cathepsin S RNA editing and gene/protein expression (187).

**NOVEL COMPUTATIONAL TOOLS FOR LncRNA ANNOTATION AND FUNCTIONAL PREDICTION**

Genomic annotation of IncRNA sequences requires defining the precise genomic coordinates of IncRNA exons and their respective transcription start sites. LncRNA annotation also involves functional annotation with respect to predicted biological mechanisms, subcellular localization, and affected cell types/tissues. While lncRNAs share some similarities with mRNAs such as transcript length and splicing structure (188), proper identification and characterization of specific long non-coding transcripts still remains a challenge. Unlike mRNAs, lncRNAs often exhibit lower stability, lower abundance, less splicing and greater nuclear localization (189).

With the widespread application of high-throughput sequencing technologies, both automated and manual methods
have been adopted to properly define lncRNA sequences from RNA-seq data. Automated annotation generates a larger catalog of lncRNAs and harnesses a transcriptome assembly consisting of two distinct strategies. In one automated approach, reads are first aligned to the reference genome to reveal all the possible splicing events which are subsequently assembled into transcripts (190, 191). In another automated approach, transcripts are built de novo from experimental reads and later aligned to a particular reference genome. Fu et al. (192) used both short and long sequencing reads to demonstrate superior sensitivity of transcript assembly and isoform annotation accuracy with the de novo approach. Automated assembly is fast as it does not require wet-lab based characterization, and it is considerable cheaper than the manual approach (144).

Although it produces a smaller catalog of lncRNAs compared to the automated method, manual annotation produces higher quality lncRNA transcript sequences and thus improves functional characterization. The widely adopted GENCODE project annotation of lncRNAs utilizes a manual curation approach (193), and integrates different sources of data together with computational analyses to generate a transcript model. cDNA and expressed sequence tag (EST) sequences deposited in publicly available databases are typically the starting point for manually annotating lncRNA transcripts. These are integrated with Cap Analysis of Gene Expression (5' CAGE) and poly(A) position profiling by sequencing (3P-seq) to characterize 5' and 3' ends, respectively. These manually annotated transcripts are then mapped to reference genomes and assigned exon and splice site locations (119). The RefSeq (Reference Sequence) project also implements manual annotation of long non-coding RNAs that are integrated with automated methods (194). Manually annotated lncRNAs can be further divided into subclasses such as intergenic lncRNAs (lincRNAs), antisense lncRNAs, and intronic lncRNAs. As cDNA annotation depends on the availability of full length transcripts, manual annotation focuses primarily on genomic annotation. As a result the manual approach produces a more comprehensive set of pseudogenes and alternatively spliced transcripts (193).

Another comprehensive database established in 2016 is NONCODE that dedicates itself to collecting lncRNAs through integration with other databases (e.g., RefSeq and Ensembl) and exhaustive annotation. Compared to these other databases, NONCODE has collected more lncRNA transcripts (excluding tRNAs and rRNAs) and provides unique annotations of lncRNAs (e.g., RNA secondary structure, expression in exosomes, associations between lncRNA and disease) (6, 195). NONCODE also provides lncRNAs for over 15 species including mouse, zebrafish, and C-elegans. The latest version of NONCODE (v5), which also captures lncRNAs from the literature, consists of nearly 550,000 annotated lncRNAs (195).

There are now an array of computational tools to annotate the sequences and functions of the expanding catalog of lncRNAs, as described in Table 2. Existing computational methods for lncRNA identification include those that require a reference genome and those that are reference-free. Examples of methods requiring a reference genome include UChncR, IncScore (205), COME (206), and IncRScan-SVM (207). Reference-free methods to identify lncRNAs from RNA-seq data include LncADeep (198), lncRNAAnet (197, 208), FEELnc (197), longdist (204), lncRNA-MFDL (209), and CPC2 (210). Many of these tools employ artificial intelligence algorithms (e.g., machine learning, deep learning) in order to distinguish lncRNAs from their protein-coding transcript counterparts.

Unlike protein functions that can be inferred from protein-coding sequences, it is more difficult to infer lncRNA function from RNA sequences. Zhou et al. developed a tool, lncFuniTK that calculates a Functional Information Score (FIS) to quantitatively measure the functional importance of a lncRNA (199), based on the top Gene Ontology and inferred regulatory networks for lncRNAs and their neighboring genes. Another tool, FEELnc, annotates lncRNA function by evaluating neighboring genes to predict both lncRNA function and mRNA partners (197). Given that lncRNA function often depends on subcellular localization, the lncLocator tool predicts five lncRNA categories: nucleus, cytoplasm, cytosol, exosome, and ribosome (200). LncADeep provides enriched pathways and functional modules for lncRNA functional annotation by integrating KEGG and Reactome Pathway databases in a deep learning framework (211). A novel method for lncRNA classification is SEEKR, which counts lncRNA k-mer frequencies from nucleotide sequences, which may be correlated with lncRNA localization or protein binding (202).

Like other ncRNAs, lncRNA functional annotation also depends on accurate secondary structure prediction. There are several computational tools to predict lncRNA structures such as pysster (212), RNAfold (213), RNAstructure (214), and UNAFold (215). Other tools are available to study the evolutionary conservation of lncRNAs, including EvoFold (216), Evolinc (217), RNAz (218), and SISSiz (219). Since lncRNAs are generally poorly conserved, there are strategies to examine remnants of protein coding sequences in these lncRNAs (220). Identifying evolutionarily conserved lncRNA structures and binding domains may provide clues to predict lncRNA function for follow-up experimentation.

**CONCLUSION**

The emergence of RNA-seq and other omics technologies in the past decade have catalyzed the identification of a plethora of novel lncRNAs. To date, more than 30 lncRNAs with functional relevance to CAD have been characterized (Table 1), yet numerous lncRNAs remain to be studied in greater detail that are linked to endothelial, smooth muscle, macrophage, and lipid traits. With the growing number of CAD GWAS candidate loci harboring lncRNAs, and improved fine-mapping and annotation approaches, there is an opportunity to functionally dissect these regions to develop novel strategies to target non-coding genomic risk factors. As outlined in this review, a multi-faceted approach is likely required to successfully prioritize and study these lncRNAs, which may include implementation of long-read and high-depth sequencing, improved computational tools, coupled
with orthogonal high-throughput experimental validation assays. Careful consideration of the lower abundance, context-specific expression of IncRNAs, and thoughtful study designs may improve chances of success in these multi-omics assays. However, it should also be noted that in many cases, more traditional and lower throughput approaches would be equally appropriate to characterize a given IncRNA, thus reducing the overall costs and required expertise.

For conserved IncRNAs with predicted roles in altering CAD pathogenesis, loss of function studies can be performed in animal models, such as the mouse (ApoE<sup>−/−</sup> or LDLR<sup>−/−</sup> backgrounds) or zebrafish. However, with the majority of human IncRNAs being poorly conserved across species, they may be better suited to studies in primary human cells or induced pluripotent stem cell (iPSC) derived vascular cells. In the context of CAD and other cardiometabolic disorders, genetic manipulation of IncRNAs via antisense oligonucleotides (221) or CRISPR/Cas9 to either delete (23, 222, 223) or activate/repress IncRNA expression, may lead to the identification of specific IncRNA binding partners, subcellular localization and functional insights relevant to CAD. IncRNA discovery/annotation can be further improved by integrating these genetic perturbations with high-dimensional transcriptomic and epigenomic assays (e.g., RNA-seq, ATAC-seq and ChIP-seq) to mark IncRNA promoters, decipher RNA polymerase and transcription factor binding, and reveal the dynamics of IncRNA regulatory activities. Single-cell based assays may also shed light on cell-specific markers and dynamics of IncRNAs across lineages (224, 225). Unraveling the complexity of IncRNA function in the setting of atherosclerosis may hold the key to delineate causal disease-associated pathways. In this regard it will also be important to determine whether IncRNAs operate

| Tool name and reference* | Model<sup>8</sup> | Code available | Application | Performance results |
|--------------------------|-----------------|----------------|-------------|---------------------|
| AnnoLnc (196)            | Statistical approach | http://annolnc.cbi.pku.edu.cn | Annotate long non-coding RNA (IncRNA) | Not reported. |
| FEELnc (197)             | Random forest    | https://github.com/tderrien/FEELnc | Annotate long non-coding RNA (IncRNA) | High classification power (AUC = 0.97). |
| LncADeep (198)           | Deep belief network built as a stack of restricted Boltzmann machines | https://github.com/cyang235/LncADeep | Identification and functional annotation for IncRNAs | With 10-fold cross validation, average sensitivity of 98.1% and specificity of 97.2% and an average harmonic mean of 97.7% |
| LncFunTK (199)           | Statistical approach | https://github.com/zhoujj2013/lncfuntk | To integrate ChIP-seq, CLIP-seq and RNA-seq data to predict, prioritize and annotate IncRNA functions. | Calculates a Functional Information Score (FIS) to quantitatively predict functional importance. |
| IncLocator (200)         | Ensemble of support vector machine and random forest classifiers. | http://www.csbio.sjtu.edu.cn/bioinf/IncLocator/ | To predict IncRNA subcellular localization. | Accuracy of 59% for prediction. |
| PennDiff (201)           | Regression-based statistical approach | https://github.com/tigerhu15/PennDiff | To detect differential transcript isoforms from RNA-seq data | Based on both annotations (RefSeq and Ensemble), estimates from PennDiff have Spearman correlation coefficients of 0.87 and 0.76, respectively. |
| SEEKR (202)              | Statistical approach | https://github.com/CalabreseLab | Prediction of IncRNA subcellular localization, protein interactors | LncRNAs of related function have similar k-mer profiles, despite linear sequence similarity |
| UClncR (203)             | Statistical approach | http://bioinformaticstools.mayo.edu/research/UClncR | Performs transcript assembly, prediction of IncRNA candidates in bulk RNA-seq data, quantification and annotation both known and novel IncRNA candidates. | For IncRNA prediction, UClncR reported 66 “novel” IncRNA transcripts and 12 IncRNAs overlapping with nearby genes (the recall rate of 90.7%). |
| A support vector machine based method to distinguish long non-coding RNAs from protein transcripts (204) | Support vector machine | https://github.com/hugowschneider/longdist.py | To distinguish IncRNAs from protein coding transcripts. | 98.21% accuracy in classifying long non-coding RNAs from protein coding transcripts. |

* Three of the publications have not been constructed into available tools but rather represent a framework for analysis. § Model type does not include preprocessing which may or may not include alignment of protein-coding regions. The link is provided if the code is available otherwise the column is marked with an “X”. AUC, area under the curve.
synergistically, serve redundant and/or compensatory roles with other dysregulated lncRNAs and/or mRNAs associated with CAD.

**AUTHOR CONTRIBUTIONS**

AT and CM conceived of the manuscript. AT, DW, MK, CD, MP, and CM wrote the manuscript.

**REFERENCES**

1. Weber C, Noels H. Atherosclerosis: current pathogenesis and therapeutic options. *Nat Med.* (2011) 17:1410–22. doi: 10.1038/nm.2538
2. Tabas I, García-Cardena G, Owens GK. Recent insights into the cellular biology of atherosclerosis. *J Cell Biol.* (2015) 209:13–22. doi: 10.1083/jcb.201412052
3. St. Laurent G, Vyatkin Y, Kapranov P. Dark matter RNA illuminates the puzzle of genome-wide association studies. *BMC Med.* (2014) 12:97. doi: 10.1186/1741-7015-12-97
4. Liu Y, Zheng L, Wang Q, Hu Y-W. Emerging roles and mechanisms of long noncoding RNAs in atherosclerosis. *Int J Cardiol.* (2017) 228:570–82. doi: 10.1016/j.ijcard.2016.11.182
5. Djebari S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. *Nature* (2012) 489:101–8. doi: 10.1038/nature11233
6. 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:D2203–8. doi: 10.1093/nkar/gkv1252
7. Palazzo AF, Lee ES. Non-coding RNA: what is functional and what is junk? *Front Genet.* (2015) 6:2. doi: 10.3389/fgene.2015.00002
8. Bhartiya D, Scaria V. Genomic variations in non-coding RNAs: Structure, function and regulation. *Genomics* (2016) 107:59–68. doi: 10.1016/j.ygeno.2016.01.005
9. Costa MC, Leitão AL, Enqupta F. Biogenesis and mechanism of action of small non-coding RNAs: insights from the point of view of structural biology. *Int J Mol Sci.* (2012) 13:10268–95. doi: 10.3390/ijms130810268
10. Bhat SA, Ahmad SM, Mumtaz PT, Malik AA, Dar MA, Urwat U, et al. Long non-coding RNAs: mechanism of action and functional utility. *Noncoding RNA Res.* (2016) 1:43–50. doi: 10.1007/s41467-016-0002
11. Bartsch D, Zirkel A, Kuriyan J. Characterization of Circular RNAs (circRNA) Associated with the Translation Machinery. *Methods Mol Biol.* (2018). 1724:159–66. doi: 10.1007/978-1-4939-7564-4_13
12. Shi X, Sun M, Liu H, Yao Y, Song Y. Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* (2013) 339:159–66. doi: 10.1016/j.canlet.2013.06.013
13. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* (2009) 458:223–7. doi: 10.1038/nature07672
14. Pang KC, Fritch MC, Mattick JS. Rapid evolution of noncoding RNA: lack of conservation does not mean lack of function. *Trends Genet.* (2006) 22:1–5. doi: 10.1016/j.tig.2005.10.003
15. Qu Z, Adelson DL. Evolutionary conservation and functional roles of ncRNA. *Front Genet.* (2012) 3:205. doi: 10.3389/fgene.2012.00205
16. Diederichs S. The four dimensions of noncoding RNA conservation. *Trends Genet.* (2014) 30:121–3. doi: 10.1016/j.tig.2014.01.004
17. Wang KC, Chang HY. Molecular mechanisms of long non-coding RNAs. *Mol Cell.* (2011) 43:904–14. doi: 10.1016/j.molcel.2011.08.018
18. Natrajeel L, Geißler C, Csaba G, Wei Y, Zhu M, di Francesco A, et al. miR-103 promotes endothelial maladaptation by targeting IncWDR59. *Nat Commun.* (2018) 9:2645. doi: 10.1038/s41467-018-05065-z
19. Santovito D, Egea V, Weber C. Small but smart: MicroRNAs orchestrate atherosclerosis development and progression. *Biochim Biophys Acta* (2016) 1861(12 Pt B):2075–86. doi: 10.1016/j.bbalip.2015.12.013
20. Feinberg MW, Moore KJ. MicroRNA regulation of atherosclerosis. *Circ Res.* (2016) 118:703–20. doi: 10.1161/CIRCRESAHA.115.306300

**FUNDING**

This work was supported by National Institutes of Health (NIH) grants R00 HL125912 (CM) and F31 NR017821 (CD) and a Leducq Foundation Transatlantic Network of Excellence award (CM). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
35. Michalik KM, You X, Manavski V, Doddballapur A, Zornig M, Braun T, et al. Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circ. Res.* (2014) 114:1389–97. doi: 10.1161/CIRCRESAHA.114.303265.

36. Song T-F, Huang L-W, Yuan Y, Wang H-Q, He H-P, Ma W-L, et al. LncRNA MALAT1 regulates smooth muscle cell phenotype switch via activation of autophagy. *Oncotarget.* (2018) 9:4411–26. doi: 10.18632/oncotarget.23230.

37. Luo J-H, Ren B, Keryanov S, Tseng GC, Rao UNM, Monga SP, et al. Identification and initial functional characterization of a human vascular cell-dominant long noncoding RNA. *BMC Genomics.* (2014) 53:393–406. doi: 10.1016/j.molcel.2014.06.013.

38. Guo F, Li Y, Liu Y, Wang J, Li Y, Gi L. Inhibition of metastasis-associated lung adenocarcinoma transcript 1 in CaSkii human cervical cancer cells suppresses cell proliferation and invasion. *Acta Biochim Biophys Sin (Shanghai).* (2010) 42:224–9. doi: 10.1039/ab090008f.

39. Koshimizu T, Fujiwara Y, Sakai N, Shibata K, Tsuchiya H. Oxytocin stimulates expression of a noncoding RNA tumor marker in a human neuroblastoma cell line. *Life Sci.* (2010) 86(11–12):455–60. doi: 10.1016/j.lfs.2010.02.001.

40. Guffanti A, Iacono M, Pelucchi P, Kim N, Soldà G, Croft LJ, et al. A novel non-coding RNA, CHROME, regulates cholesterol homeostasis. *Nat Metab.* (2018) 9:237. doi: 10.1038/s41467-017-02113-y.

41. Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, et al. Long noncoding RNA NEAT1-dependent SFPQ relocation from NFAT. *Proc Natl Acad Sci USA.* (2017) 114:133–9. doi: 10.1073/pnas.1617096114.

42. Ahmed ASI, Dong K, Liu J, Wen T, Yu L, Xu F, et al. Long noncoding RNA NEAT1 is critical for phenotypic switching of vascular smooth muscle cells. *Circ Res.* (2015) 116:1143–56. doi: 10.1161/CIRCRESAHA.116.307387.

43. Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, et al. The long noncoding RNA NEAT1 regulates smooth muscle cell proliferation and migration via repressing PTEN by activating NFAT. *J Clin Invest.* (2019) 133:2050–65. doi: 10.1172/JCI96339.

44. Bell RD, Long X, Lin M, Bergmann JH, Nanda V, Cowan SL, et al. Identification and initial functional characterization of a human vascular cell-dominant long noncoding RNA. *Arterioscler Thromb Vasc Biol.* (2014) 34:1249–59. doi: 10.1161/ATVBAHA.113.303240.

45. Menon NK, Kaur R, Makharia G, Malhotra M, Singh AK, et al. The long non-coding RNA NEXN-AS1 mitigates atherosclerosis by regulating the actin-binding protein NEXN. *Nat Commun.* (2018) 9:292. doi: 10.1038/s41467-017-02431-1.

46. McVey DG, Tuller AS, Kim N, Nusinoff-Levin T, Croft L, et al. Long noncoding RNA MEG3 suppressed endothelial cell proliferation and migration through regulating miR-21. *Am J Transl Res.* (2016) 8:39. doi: 10.3892/ajtr.2015.25310.

47. Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, et al. Long noncoding RNA NEAT1-dependent SFPQ relocation from NFAT. *Proc Natl Acad Sci USA.* (2017) 114:133–9. doi: 10.1073/pnas.1617096114.
mRNA by long noncoding RNA during brain vascular assembly. *Arterioscler Thromb Vasc Biol.* (2018) 38:1562–75. doi: 10.1161/ATVBAHA.118.310848

73. Das S, Reddy MA, Senapati P, Stapleton K, Lanting L, Wang M, et al. Diabetes mellitus-induced long noncoding RNA Dnm3os regulates macrophage functions and inflammation via nuclear mechanisms. *Arterioscler Thromb Vasc Biol.* (2018) 38:1806–20. doi: 10.1161/ATVBAHA.117.310663

74. Loebel DAE, Tsoi B, Wong N, Tam PPL. A conserved noncoding intrinsic transcript at the mouse Dnm3 locus. *Genomics* (2005) 85:782–9. doi: 10.1016/j.ygeno.2005.02.001

75. Du M, Yuan L, Tan X, Huang D, Wang X, Zheng Z, et al. The LPS-inducible long non-coding RNA, lncLSTR, regulates systemic lipid metabolism in mice. *Nat Med.* (2018) 24:304–12. doi: 10.1038/nm.4479

76. Halley P, Kadakkuzha BM, Faghihi MA, Magistri M, Zeier Z, Khorova O, et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell.** (2010) 38:662–74. doi: 10.1016/j.molcel.2010.03.021

77. Matterle A, Pu X, Wood H, Xiao Q, Gor S, Ng FL, et al. Functional analyses of coronary artery disease associated variation on chromosome 9p21 in vascular smooth muscle cells. *Hum Mol Genet.* (2012) 21:4201–9. doi: 10.1093/hmg/ddc224

78. Zhou X, Han X, Wittfeldt A, Sun J, Liu C, Wang X, et al. Long non-coding RNA ANRIL regulates inflammatory responses as a novel component of NF-κB pathway. *RNA Biol.* (2016) 13:98–108. doi: 10.1080/15476286.2015.1122164

79. Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* (2010) 464:409–12. doi: 10.1038/nature08801

80. Holdt LM, Teupser D. Long Noncoding RNA ANRIL: Lnc-ing genetic variation at the chromosome 9p21 locus to molecular mechanisms of atherosclerosis. *Front Cardiovasc Med.* (2018) 5:145. doi: 10.3389/fcvm.2018.00125

81. Ahrabi S, Berkant Ö, Lalem T, Özbekli N, Göksel S, Korkmaz Ö, et al. Long non-coding RNAs in the atherosclerotic plaque. *Atherosclerosis* (2017) 266:176–81. doi: 10.1016/j.atherosclerosis.2017.10.012

82. Cremer S, Michalik KM, Fischer A, Pfisterer L, Jäe N, Winter C, et al. Hematopoietic deficiency of the long non-coding RNA MALAT1 promotes atherosclerosis and plaque inflammation. *Circulation* (2018). doi: 10.1161/CIRCULATIONAHA.117.029015. [Epub ahead of print]

83. Koldemir O, Özgür E, Gezer U. Accumulation of GAS5 in exosomes is a marker of apoptosis induction. *Biomed Rep.* (2017) 6:358–62. doi: 10.3892/br.2017.848

84. Aryal B, Roalln N, Fernández-Hernando C. Noncoding RNAs and atherosclerosis. *Curr Atheroscler Rep.* (2014) 16:407. doi: 10.1007/s11883-014-0407-3

85. Zhou T, Ding J, Wang X, Zheng X. Long noncoding RNAs and atherosclerosis. *Atherosclerosis* (2016) 248:51–61. doi: 10.1016/j.atherosclerosis.2016.02.025

86. Visel A, Zhu R, Liu X, He Z. Long non-coding RNA H19 and MALAT1 gene variants and role of long non-coding RNA H19 in carcinogenesis. *Front Biosci.** (2018) 23:614–25. doi: 10.2741/4608

87. Gao W, Zhu M, Wang H, Zhao S, Zhao D, Yang Y, et al. Association and atherosclerosis. *Am J Hum Genet.* (2010) 86:658–70. doi: 10.1016/j.ajhg.2010.06.031

88. Hofmann P, Sommer J, Theodorou K, Kirchhof L, Fischer A, Li Y, et al. Hematopoietic deficiency of the long non-coding RNA MALAT1 and role of long non-coding RNA H19 in carcinogenesis. *Front Biosci.** (2018) 23:614–25. doi: 10.2741/4608

89. Zhou T, Ding J, Wang X, Zheng X. Long noncoding RNAs and atherosclerosis. *Atherosclerosis* (2016) 248:51–61. doi: 10.1016/j.atherosclerosis.2016.02.025
EMT of cardiac fibroblast progenitors. Development (2012) 139:2139–49. doi: 10.1242/dev.079970

Miller CL, Anderson DR, Kundu RK, Raisenda A, Nürnberg ST, Diaz R, et al. Disease-related growth factor and embryonic signaling pathways modulate an enhancer of TCF21 expression at the 6p22.2 coronary heart disease locus. PLoS Genet. (2013) 9:e1003652. doi: 10.1371/journal.pgen.1003652

Miller CL, Haas U, Diaz R, Leeper NJ, Kundu RK, Patolla B, et al. Coronary heart disease-associated variation in TCF21 disrupts a miR-224 binding site and miRNA-mediated regulation. PLoS Genet. (2014) 10:e1004263. doi: 10.1371/journal.pgen.1004263

Nürnberg ST, Cheng K, Raisenda A, Kundu R, Miller CL, Kim JB, et al. Coronary artery disease associated transcription factor TCF21 regulates smooth muscle precursor cells that contribute to the fibrous cap. PLoS Genet. (2015) 11:e1005155. doi: 10.1371/journal.pgen.1005155

Hormozdari F, van de Bunt M, Segrè AV, Li X, Joo JW, Bilow M, et al. Functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. Biochim Biophys Acta (2016) 1859:3–15. doi: 10.1016/j.bbagrm.2015.10.010

Guil S, Soler M, Portela A, Carrère J, Fonalleras E, Gómez A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genome Biol. (2011) 12:193. doi: 10.1186/gb-2011-12-11-193

Freedman JE, Miano JM, National Heart, Lung, and Blood Institute. Coronary artery disease associated transcription factor TCF21 regulates colocalization of GWAS and eQTL signals detects target genes. Am J Hum Genet. (2016) 99:1245–60. doi: 10.1016/j.ajhg.2016.10.003

Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. (2016) 48:245–52. doi: 10.1038/ng.3506

Kashi K, Henderson L, Bonetti A, Carninci P. Discovery and functional analysis of IncRNAs: Methodologies to investigate an uncharacterized transcriptome. Biosci Biotechnol Biochem. (2016) 80:159–63. doi: 10.1002/j.2164-0105.2015.01010

Kukurba KR, Montgomery SB. RNA sequencing and analysis. Cold Spring Harb Protoc. (2015) 2015:951–69. doi: 10.1101/pdb.top084970

Freedman JE, Miano JM, National Heart, Lung, and Blood Institute Participants’. Challenges and opportunities in linking long noncoding mRNAs to cardiovascular, lung, and blood diseases. Arterioscler Thromb Vasc Biol. (2017) 37:21–5. doi: 10.1161/ATVBAHA.116.308513

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgher H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. (2012) 22:1775–89. doi: 10.1101/gr.132159.111

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

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. (2011) 25:1915–27. doi: 10.1101/gad.1744611

Guil S, Soler M, Portela A, Carrère J, Fonalleras E, Gómez A, et al. Intronic RNAs mediate EZH2 regulation of epigenetic targets. Nat Struct Mol Biol. (2012) 19:664–70. doi: 10.1038/nsmb.2315

Tsoi LC, Iyer MK, Stuart PE, Swindell WR, Gudjonsson JE, Tejasvi T, et al. Long non-coding RNA meg3 promotes angiogenesis after ischemic brain injury by activating notch signaling. Mol Neurobiol. (2017) 54:8179–90. doi: 10.1007/s12035-016-0270-z

Boulberdais M, Scott E, Ballantyne M, Garcia R, Descamps R, Angellin GD, et al. A role for the long noncoding RNA SENCr in commitment and function of endothelial cells. Mol Ther. (2016) 24:978–80. doi: 10.1038/mt.2016.41

Zhang H, Xue C, Wang Y, Shi J, Zhang X, Li W, et al. Deep RNA sequencing uncovers a repertoire of human macrophage long intergenic noncoding RNAs modulated by macrophage activation and associated with cardiometabolic diseases. J Am Heart Assoc. (2017) 6:e007431. doi: 10.1161/JAHA.117.007431

Xue C, Zhang X, Zhang H, Ferguson JJ, Wang Y, Hinkle CC, et al. De novo RNA sequence assembly during in vivo inflammatory stress reveals hundreds of unannotated lncRNAs in human blood CD14+ monocytes and in adipose tissue. Physiol Genomics. (2017) 49:287–305. doi: 10.1152/physiogenomics.00001.2017

Hong C, Tontonoz P. Liver X receptors in lipid metabolism: opportunities for drug discovery. Nat Rev Drug Discov. (2014) 13:433–44. doi: 10.1038/nrd4280

Parikh M, Patel K, Soni S, Gandhi T. Liver X receptor: a cardinal target for atherosclerosis and beyond. J Atheroscler Thromb. (2014) 21:519–31. doi: 10.5551/jat.19778

Zhang X, Brevvort SR, Angidsen J, Fu M, Schmidt DR, Holmstrom SR, et al. Liver LXRα expression is crucial for whole body cholesterol homeostasis and reverse cholesterol transport in mice. J Clin Invest. (2012) 122:1688–99. doi: 10.1172/JCIC59187

Tontonoz P, Wu X, Jones M, Zhang Z, Salisbury D, Sallam T. Long noncoding RNA facilitated gene therapy reduces atherosclerosis in a murine model of familial hypercholesterolemia. Circulation. (2017) 136:776–8. doi: 10.1161/CIRCULATIONAHA.117.029902

Tilgner H, Jahanbani F, Blauwkamp T, Moshefi A, Jaeger E, Chen F, et al. Comprehensive transcriptome analysis using synthetic long-read sequencing reveals molecular co-association of distant splicing events. Genome Biol. (2016) 17:239. doi: 10.1186/s13059-016-1103-0

Uszczynska-Ratajczak B, Lagarde J, Frankish A, Guigó R, Johnson R. Towards a complete map of the human long non-coding RNA transcriptome. Nat Rev Genet. (2018) 19:535–48. doi: 10.1038/s41576-018-0017-y

Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. (2016) 17:239. doi: 10.1186/s13059-016-1103-0

Lagarde J, Uszczynska-Ratajczak B, Santoyo-Lopez J, Gonzalez JM, Tapaneri N, Mudge JM, et al. Extension of human lncRNA transcripts by RACE coupled with long-read high-throughput sequencing (RACE-Seq). Nat Commun. (2016) 7:12339. doi: 10.1038/ncomms12339
204. Schneider HW, Raiol T, Brigido MM, Walter MEMT, Stadler PF. A
210. Kang Y-J, Yang D-C, Kong L, Hou M, Meng Y-Q, Wei L, et al. CPC2: a fast
209. Fan X-N, Zhang S-W. lncRNA-MFDL: identification of human long
207. Sun L, Liu H, Zhang L, Meng J. lncRScan-SVM: a tool for predicting
196. Hou M, Tang X, Tian F, Shi F, Liu F, Gao G. AnnoLnc: a web server for
198. Yang C, Yang L, Zhou M, Xie H, Zhang C, Wang MD, et al. FEELnc:
197. Wucher V, Legeai F, Hédan B, Rizk G, Lagoutte L, Leeb T, et al. FEELnc:
199. Zhou J, Huang Y, Ding Y, Yuan J, Wang H, Sun H. lncFunTK: a tool kit
200. Cao Z, Pan X, Yang Y, Huang Y, Shen H-B. The lncLocator: a
201. You B-H, Yoon S-H, Nam J-W. High-confidence coding and
190. Steijger T, Abril JF, Engström PG, Kokocinski F, RGASP Consortium,
208. Baek J, Lee B, Kwon S, Yoon S. lncRNAnet: long Non-coding RNA
Frontiers in Cardiovascular Medicine | www.frontiersin.org
18 February 2019 | Volume 6 | Article 9
Turner et al.

215. Markham NR, Zuker M. UNAFold: software for nucleic acid
213. Lorenz R, Bernhart SH, Höner Zu Siederdissen C, Tafer H, Flamm C,
216. Pedersen JS, Bejerano G, Siepel A, Rosenbloom K, Lindblad-Toh K,
217. Nelson ADL, Devisetty UK, Palos K, Haug-Baltzell AK, Lyons E,
219. Gesell T, Washiel S. Dineucleotide controlled null models for
214. Kim DH, Marinov GK, Pepke S, Singer ZS, He P, Williams B, et al.
212. Aparicio-Prat E, Arnan C, Sala L, Bosch N, Guigo R, Johnson R.
211. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for
218. Kim DH, Marinov GK, Pepke S, Singer ZS, He P, Williams B, et al.
219. Markham NR, Zuker M. UNAFold: software for nucleic acid
216. Pedersen JS, Bejerano G, Siepel A, Rosenbloom K, Lindblad-Toh K,
217. Nelson ADL, Devisetty UK, Palos K, Haug-Baltzell AK, Lyons E,
219. Gesell T, Washiel S. Dineucleotide controlled null models for
214. Kim DH, Marinov GK, Pepke S, Singer ZS, He P, Williams B, et al.
212. Aparicio-Prat E, Arnan C, Sala L, Bosch N, Guigo R, Johnson R.