Title

The how and why of lncRNA function: An innate immune perspective

Journal

Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 1863(4)

ISSN

1874-9399

Authors

Robinson, Elektra K
Covarrubias, Sergio
Carpenter, Susan

Publication Date

2020-04-01

DOI

10.1016/j.bbagrm.2019.194419

Peer reviewed

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

BBA - Gene Regulatory Mechanisms xxx (xxxx) xxxx

Contents lists available at ScienceDirect

BBA - Gene Regulatory Mechanisms

journal homepage: www.elsevier.com/locate/bbagrm

Review

The how and why of lncRNA function: An innate immune perspective☆

Elektra K. Robinson[1], Sergio Covarrubias[1], Susan Carpenter*

Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz, Santa Cruz, CA, United States of America

ARTICLE INFO

Keywords:
Long noncoding RNAs (lncRNAs)
Toll-like receptors (TLRs)
Interferon (IFN)
Autoimmunity
Macrophages
Inflammation
RNA-sequencing
ChIP-sequencing
ATAC-sequencing

ABSTRACT

Next-generation sequencing has provided a more complete picture of the composition of the human transcriptome indicating that much of the "blueprint" is a vastness of poorly understood non-protein-coding transcripts. This includes a newly identified class of genes called long noncoding RNAs (lncRNAs). The lack of sequence conservation for lncRNAs across species meant that their biological importance was initially met with some skepticism. LncRNAs mediate their functions through interactions with proteins, RNA, DNA, or a combination of these. Their functions can often be dictated by their localization, sequence, and/or secondary structure. Here we provide a review of the approaches typically adopted to study the complexity of these genes with an emphasis on recent discoveries within the innate immune field. Finally, we discuss the challenges, as well as the emergence of new technologies that will continue to move this field forward and provide greater insight into the biological importance of this class of genes. This article is part of a Special Issue entitled: ncRNA in control of gene expression edited by Kotb Abdelmohsen.

1. Introduction

One of the most profound discoveries from the sequencing of the human genome is that over 85% of the genome is transcribed, yet < 2% encodes protein-coding genes [1]. Large consortiums such as ENCODE and FANTOM have embarked on attempting to characterize all functional coding and noncoding elements in the genome and have compiled important regulatory data for these elements [2–4]. Long noncoding RNAs (lncRNAs) represent the largest group of non-coding RNAs produced from the genome. LncRNAs are defined as transcripts > 200 nucleotides in length, lacking protein-coding potential. In the most recent GENCODE V30 release, there are 16,193 annotated lncRNAs in the human genome [4]. Additionally, there are over 14,000 pseudogenes, that could fall under the description of long noncoding RNAs which is simply based on them being 200 nucleotides or greater in length. Less than ~3% of annotated lncRNAs have ascribed functions. Hence this class of RNAs is greatly in need of further investigation [4]. From those that have been characterized, it is clear that lncRNAs can function through a variety of mechanisms to regulate gene expression both at the transcriptional and post-transcriptional levels [5].

As we will discuss through this review lncRNAs can mediate their functions through interactions with proteins, RNA, DNA, or a combination of these. Furthermore, the function of lncRNAs can often be dictated by their localization, sequence and/or secondary structure. There are many categories and sub-categories of lncRNAs, but some of the major classifications include: antisense [6], bi-directional [7], enhancer-associated [8], intergenic lncRNAs (lincRNAs) [9], pseudogenes [10], while a full review of all classifications can be obtained from the recent review by Jarroux et al. [11]. LncRNA function cannot be determined simply based on the lncRNA classification. However, the classification can sometimes provide insight into its mechanism of action, such as antisense lncRNAs impacting their neighboring genes. However, this same classification can also lead to erroneous assumptions about how the lncRNA regulates gene expression. Recently, some lncRNAs have been demonstrated to actually encode small peptides indicating that these genes are misclassified as noncoding, although it is possible that they could also have functions as a noncoding RNA in addition to their peptide coding capacity [12–17]. It is therefore, important to have a logical methodology to study the biological importance of these genes.

The innate immune system functions as a rapid initial response against specific pathogens, while also promoting the activation and development of the adaptive immune system [18]. Macrophages and dendritic cells are important innate immune cells that initiate the immune response through recognition of specific pathogen-associated molecular patterns (PAMPs) through their germline-encoded pattern

☆ This article is part of a Special Issue entitled: ncRNA in control of gene expression edited by Kotb Abdelmohsen.

* Corresponding author.
E-mail address: sucarpen@ucsc.edu (S. Carpenter).

[1] Authors contributed equally.

https://doi.org/10.1016/j.bbagrm.2019.194419
Received 31 May 2019; Accepted 21 August 2019
1874-9399/

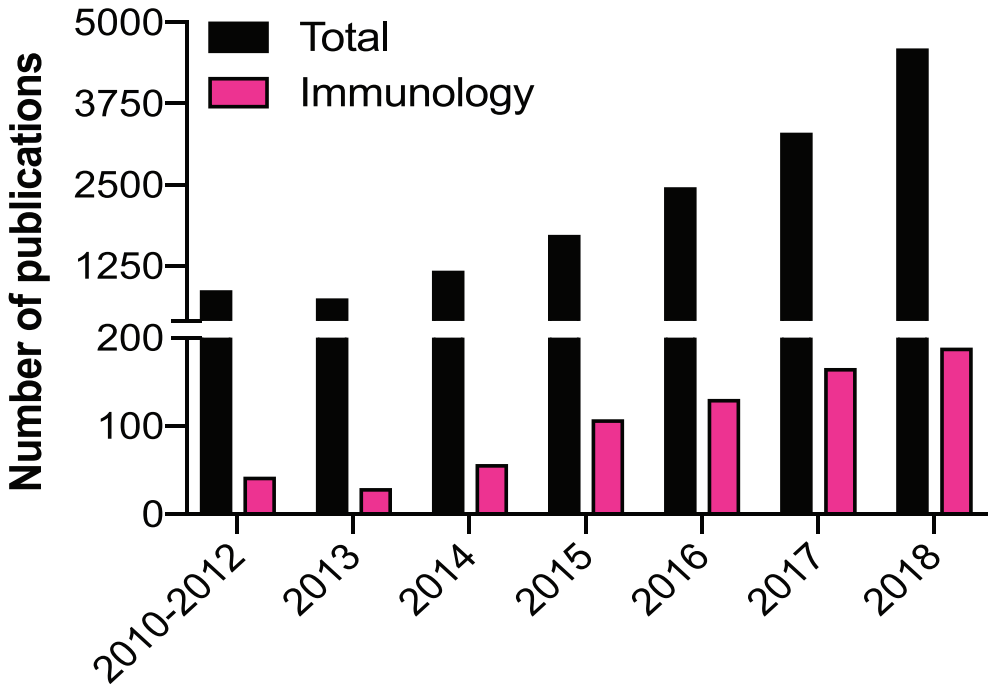


Fig. 1. LncRNA publications.
Graph representing the total number (*black bar*) or immunology related (*pink bar*) publications in the lncRNA field since 2010 to 2018.

recognition receptors (PRRs) [18]. These receptors couple pathogen-sensing to activation of downstream signaling cascades resulting in up-regulation of numerous inflammatory pathways [19]. While a robust immune response is crucial for eliminating pathogens, prolonged activation can be detrimental to the host [20]. Not surprisingly, many aspects of the inflammatory response are tightly regulated at both transcriptional and post-transcriptional levels allowing for a transient antimicrobial response while subsequently promoting a return to homeostasis [21]. Perturbations in this regulation can have significant consequences that can manifest in diseases, such as arthritis [22], multiple sclerosis [23] and cancer [20,24]. While the role of coding genes in immune cell function has been well characterized, the role of lncRNAs in these processes is just beginning to emerge [25] (Fig. 1). Here, we use the biological model system of macrophage activation as a framework to demonstrate how we approach the study of lncRNA biology. We provide a step-by-step guide to consider when studying lncRNAs. Furthermore, we discuss the challenges, as well as the emergence of new technologies that are helping evolve the ways we study these genes.

2. Getting started

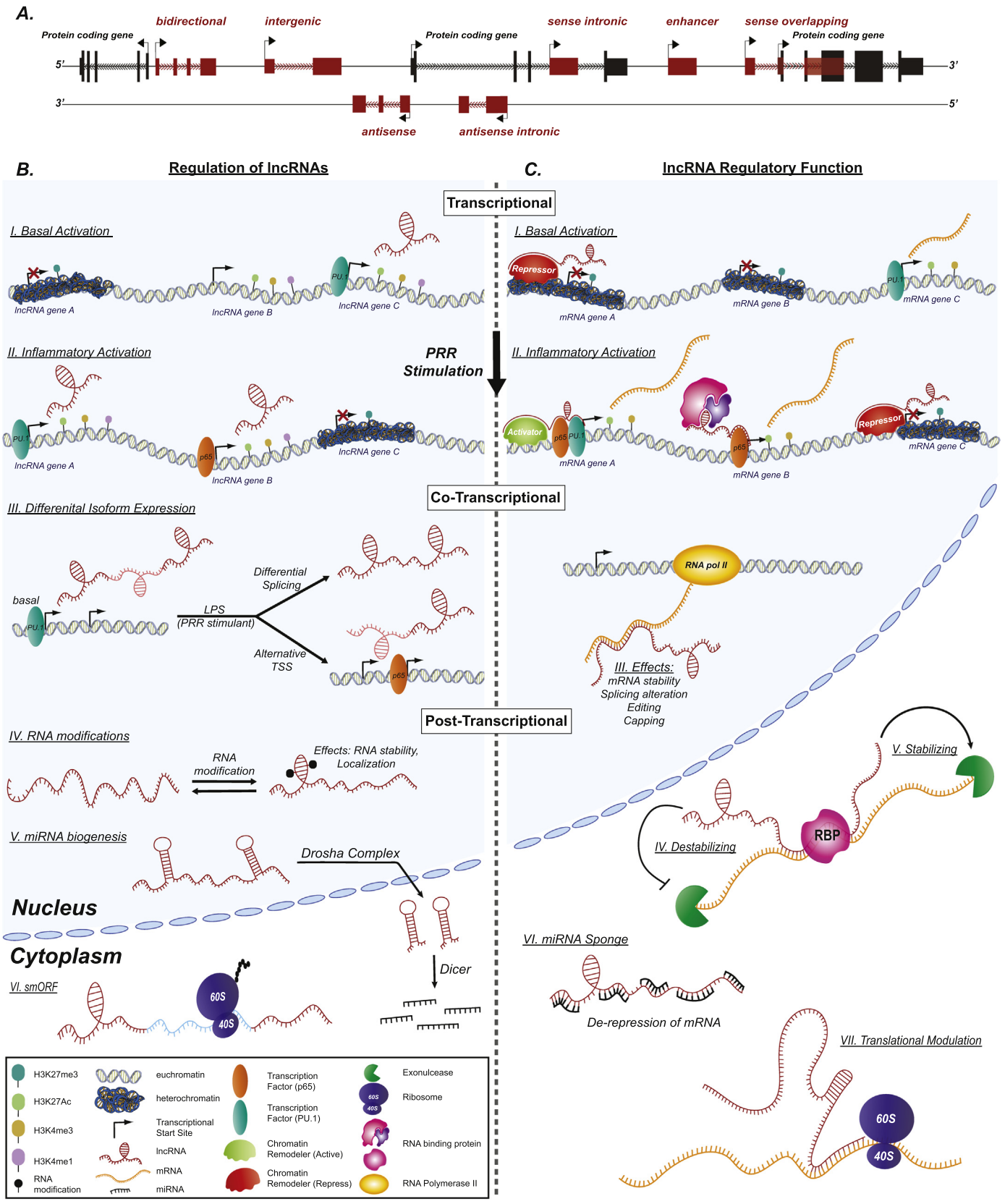
2.1. The biological question

The lncRNA field is in its infancy yet from what we do know we find that lncRNAs play critical roles in a wide variety of biological processes and diseases from cell differentiation, tissue organ development, flowering in plants, to cancer metastasis to name just a few [26–30]. We believe that lncRNAs play regulatory roles in many biological processes and diseases. Therefore, no matter what your research area is, there is a rich source of information to be obtained from the study of lncRNAs in your field of interest.

The bulk of lncRNA studies to date have focused on the cancer field [31–33]. Meanwhile, studies of lncRNAs in the context of innate immunity have lagged, making up only ~4% of all lncRNA papers to date (Fig. 1). The innate immune system provides one of the first lines of defense against infection through the induction of inflammation [34,35]. The inflammatory response of murine macrophages offers a powerful system for applying genomic approaches for studying novel lncRNAs within the framework of a pathway that has been studied for decades. Macrophages are important mediators of inflammation and initiate this response through recognition of specific pathogen-associated molecular patterns (PAMPs) through their germline-encoded pattern recognition receptors (PRRs). These receptors couple pathogen-sensing to activation of downstream signaling cascades resulting in activation of numerous transcription factors, including NF-kappaB (NF-κB) and interferon regulatory factors (IRFs) that can act in combination to both positively and negatively regulate the expression of thousands of genes [36,37]. There are 10 different TLR genes in the human genome and 13 TLR genes in mice [38–40], each binding a different PAMP [41]. Using this extensively studied biological system, we identified the first example of a TLR-stimulated lncRNA, lincRNA-Cox2, which was capable of positively and negatively regulating distinct types of innate immune genes [42–46]. Knockdown of lincRNA-Cox2 resulted in impaired production of proinflammatory genes (*i.e.*, IL-6), while IFN-related genes were hyperactivated in the absence of lincRNA-Cox2 [42–46]. Numerous other studies have made use of the TLR-signaling biological system, uncovering and characterizing dozens of novel lncRNAs that act in a wide range of mechanisms to either positively or negatively regulate this pathway as reviewed in Carpenter et al. and Hadjicharalambous et al. [47,48].

2.2. LncRNA candidate selection

As mentioned lncRNAs are categorized into five main classes of long noncoding RNAs based on their genomic location: antisense, bidirectional, intronic, enhancer-associated, and intergenic. Intergenic and enhancer lncRNAs contain their own promoters and are distinct from protein-coding genes. Bidirectional lncRNAs share a promoter and are transcribed from the opposite strand of a protein-coding gene, while intronic lncRNAs are transcribed within an intronic region of a protein-coding gene (Fig. 2A) [49,50]. The specific class of lncRNAs can often provide significant insight into how it may regulate gene expression. For example, lncRNAs antisense to a coding gene have been



(caption on next page)

Fig. 2. Classification and regulation of LncRNAs.
(A) Positional classifications of lncRNAs based on their genomic location in respect to nearby protein coding genes: bidirectional, intergenic, antisense, antisense intronic, sense intronic, enhancer, sense-overlapping. (B) The lower left panel of the figure represents how lncRNA activity can be regulated transcriptionally. Transcriptional activation is depicted basally 'Part I', as well as during inflammatory activation 'Part II' following stimulation of a pattern recognition receptor (PRR). Three lncRNA examples genes are shown A, B and C. Part III depicts how a lncRNA can undergo differential isoform expression which is regulated co-transcriptionally. During active transcription a lncRNA can either undergo differential splicing or utilize a new transcriptional start site post inflammatory stimulation, such as lipopolysaccharides (LPS). Post-transcriptional regulation of a lncRNA is broken up into three parts: IV, V and VI. After transcription is completed there is several processes that a lncRNA can undergo. Part IV depicts RNA modifications, which can change the structure of a lncRNA molecule. These modifications can be added or removed depending on the inflammatory state of a cell. Part V depicts the process of miRNA biogenesis, where a lncRNA can be processed to a mature miRNA. Part VI shows that a lncRNA can also be translated if it has a small open reading frame (smORF). (C) The lower right panel of the figure illustrates the regulatory function of a lncRNA in the nucleus and the cytoplasm. During active transcription (basally 'Part I' or inflammatory 'Part II'), a lncRNA can function to repress genes (mRNA gene A and C) or activate genes (mRNA gene A and B). A lncRNA can either be a scaffold for transcription factors to enhance activation, or a scaffold for chromatin remodeler proteins to open or close chromatin. A lncRNA can also regulate a mRNA transcript co-transcriptionally 'Part III' by affecting either the stability, change the splicing activity, editing of modifications, or even the capping of the mature mRNA. Finally, post-transcriptionally a lncRNA can function to regulate the mRNA in several ways. A lncRNA can affect the stability of a mRNA transcript: 'Part IV and Part V.' Alternatively, a lncRNA can function as a miRNA sponge which indirectly de-represses the expression of a mRNA that would be targeted by the miRNAs. Lastly, a lncRNA can modulate the translation of a mRNA by binding to ribosomes or mRNA transcripts during translation.

demonstrated to be involved in transcriptional interference, negatively affecting the expression of their coding gene [51]. While all categories of lncRNAs will no doubt be important in various biological processes, the most targetable lncRNAs are intergenic lncRNAs. A benefit to studying an intergenic lncRNA (lincRNA) is the immense variety of molecular techniques that would not apply to other types of lncRNAs, such as antisense, bidirectional and intronic lncRNAs, which often overlap coding genes and their targeting could lead to possible unwanted interference of that coding gene.

2.2.1. Cis regulators

Numerous studies have established that lncRNAs can regulate the expression of their neighboring coding genes (*cis* regulation) [51]. A recent study by Engreitz et al. demonstrated that a significant portion of lncRNAs had *cis* effects on their neighboring genes [52]. Interestingly, in most cases, the *cis* effect did not require the production of the lncRNA transcripts themselves, but instead required the processes associated with their production, such as transcription and splicing [52]. There are many examples of lncRNAs that regulate their neighboring coding genes, for example: lnc-MARCKS, lnc-TNFAIP3, AS-IL1α, lnc-IL7R, and IL-1β-eRNA [52–56]. Recently we used genetic mouse models to show that lincRNA-Cox2 can function as an enhancer RNA *in cis* to regulate its neighboring gene Ptgs2 [46]. For these reasons, one aspect to consider when selecting a candidate, no matter the class, is to investigate the effect of the transcriptional expression on neighboring coding genes. This candidate selection approach is sometimes referred to as "guilt by association" [57,58]. This bioinformatic approach drives an initial hypothesis that the lncRNA could be involved in the similar biological pathway as their neighboring protein-coding gene due to their co-expression.

2.2.2. Trans regulators

A large number of studies to date have also shown that lncRNAs can regulate gene expression on different chromosomes (*trans* regulation) [59] (Fig. 2C). The majority of lncRNAs studied in immunity were initially identified following RNA-sequencing to examine their expression profiles in specific cell lines or tissues during inflammatory activation. For example, lincRNA-Cox2 was initially identified as an up-regulated lncRNA in murine dendritic cells following TLR4 stimulation [60], as well as murine bone marrow derived macrophages (BMDMs) following TLR2-depedent stimulation [42]. Studies have also highlighted the functions of lncRNAs that are highly downregulated post inflammatory activation, such as lincRNA-EPS [61] and lnc13 [62]. Both lincRNA-Cox2 and lincRNA-EPS were discovered in BMDMs post TLR inflammatory activation and were chosen for further characterization based on their extreme expression profile. LincRNA-Cox2 is rapidly upregulated and regulates a large number of interferon stimulated genes (ISGs) and NF-κB regulated genes [42]. Meanwhile, lincRNA-EPS is rapidly down-regulated during inflammation and acts as an inflammatory brake on all ISGs during periods of homeostasis [61]. These are just two examples of lncRNAs that have provided critical insights into the roles of lncRNAs in immunity. For more information on specific lncRNAs that are involved in innate immunity we direct you to the following recent reviews on this topic [42,47,63]. In addition to these bulk RNA sequencing studies, a small number of single cell RNA sequencing studies have been performed in both human and mouse that can be utilized to examine differential expression of lncRNAs in basal *versus* treatment conditions or between cell types [64–68]. Numerous RNA-seq (both bulk and single cell) datasets are available for a variety of primary cells, cell lines or tissues of interest either basally or under a multitude of inflammatory or cellular differentiation treatments. These datasets outlined in Table 1 [42,46,74–77, 57, 61, 64, 69–73] and Table 2 [53,56, 82–90, 65–68, 78–81] provide a rich source of lncRNAs for further investigation.

EVLncRNAs [91], NONCODE [92], or LNCipedia [93] are databases that categorize published information on all annotated lncRNAs. These databases can be utilized to determine if a lncRNA is experimentally validated within one or more studies. Additionally, these databases can provide information on whether a lncRNA possesses multiple isoforms, secondary structure, cross-species conservation and/or disease-association *via* presence of single nucleotide polymorphisms (SNPs).

2.3. LncRNAs expression and specificity

Multiple studies have demonstrated that lncRNA expression is more cell type specific compared to protein-coding genes [94–96]. Such specific expression patterns can often provide important clues into the specific biology that the gene could be involved in [46] (Tables 1 and 2). A variety of consortiums exist for both human and mouse and can be utilized to determine cell type specificity of a lncRNA candidate further. For instance, GTEx [97] and XENA [98] are two websites that include RNA sequencing on healthy primary human tissue samples, in addition to samples from patients with diagnosed cancers. This will further assist the initial understanding of the expression of the lncRNA in specific tissues as well as obtaining information on whether a lncRNA is involved in cancer. In contrast, if a researcher is studying a mouse candidate lncRNA, the Mouse Cell Atlas (MCA) [99], as well as Tabula Muris [100] are excellent tools to assess specificity in cellular expression, as well as differential splicing isoforms amongst differentiated cell types. Additional websites for mouse and human expression datasets can be found at the ENCODE project [101], the European Bioinformatics Institute Expression and the FANTOM projects [102]. These sites are filled with raw and analyzed data sets from either single cell or bulk RNA sequencing from primary cells, tissues or immortalized cells ready to use to determine the statistical significance of expression for any annotated lncRNA candidate.

Table 1
Murine sequencing datasets.

	Type	Treatment	Duration (h)	Sequencing	PMID
Spleen	Tissue	TLR4	0, 6	RNA	30404006
Lung	Tissue	TLR4	0, 6	RNA	30404006
		SARS-CoV	0, 24	RNA	27462873
		PR8	0, 24, 48, 120	RNA	20978541
Monocyte	FACs sorted	–	–	RNA	24586061
				ATAC	
				ChIP - H3K27Ac	
				ChIP - H3K4me1	
				ChIP - H3K4me2	
				ChIP - H3K4me3	
Dendritic cells	CD103+	PR8	0, 24, 48, 120	RNA	20978541
	Bone marrow derived	None	–	RNA	19182780
		TLR1/2	0, 6		
		TLR4	0, 6		
		TLR9	0, 6		
Macrophage	Alveolar macrophage depletion	NS1-GFP	0, 12	RNA	30886410
	Bone marrow (monocyte)	–	–	RNA	25480296
	Lung tissue resident			ATAC	
	Spleen tissue resident			ChIP-H3K4me1	
	Kupffer cells			ChIP-H3K4me3	
	Peritoneal macrophage			ChIP-H3K27Ac	
	Colonic macrophage				
	Microglia				
	FACs Sorted	–	–	RNA	25103404
				ATAC	
				ChIP - H3K27Ac	
				ChIP - H3K4me1	
				ChIP - H3K4me2	
				ChIP - H3K4me3	
	Bone marrow derived	TLR4	0, 0.25, 0.5, 1, 2	RNA	22817891
				Chromatin, nucleus, cytoplasm	
		Mtb	0, 4, 12, 24, 48	RNA	29712924
		IFNγ Mtb	0, 4, 12, 24, 48		
		IL4/IL13 Mtb	0, 4, 12, 24, 48		
		TLR1/2	0, 5	RNA	23907535
		TLR4	0, 2, 6	RNA	27315481
		TLR4	0, 2, 6	ATAC	
		TLR4	0, 1	ATAC	26924576
		TLR1/2	0, 0.25, 0.5, 1, 2	RNA	
		TLR4	0, 0.25, 0.5, 1, 2	RNA	
				ChIP - SRF	
				ChIP - IRF3	
				ChIP - p65	
		TLR4	0, 1	ChIP - H3K27Ac	29779944
				ChIP- HeK4me2	27462873
				ChIP - PU.1	
				ChIP - Nrf2	
				ChIP - p65	
				ChIP - Smad3	
				ChIP - Fos	
				RNA	

2.4. Determining disease association

Genome-wide association studies (GWAS) have revolutionized the study of complex diseases by allowing quantitative disease-association of thousands of genetic loci [103]. These studies include evaluation of single-nucleotide polymorphisms (SNPs) or deletions and determination of their association with a disease phenotype. Diseases studied range from Inflammatory Bowel Disease (IDB) to schizophrenia [104,105]. Until recently, most GWAS studies focused on protein-coding genes, even though 90% of disease-associated SNPs lie in non-coding regions of the genome [106]. There are several databases that summarize the plethora of published human sequencing studies, including UK Biobank [107] and GWAS Catalog - EMBL-EBI [108]. Other databases specifically focus on SNPs within lncRNAs, such as lncRNASNP2 [109] and Lnc2Catlas [110].

To date, a couple of studies have clearly shown how SNPs from GWAS studies can be used to identify clinically relevant lncRNAs. Castellanos-Rubio et al. identified a SNP rs917997 associated with Celiac Disease and showed it was located within a novel lncRNA, Lnc13 [62]. Lnc13 regulates inflammatory genes and mediates its function *via* an interaction with an hnRNP protein [62]. Furthermore, they showed that the SNP disrupted the RNA-protein interaction, thus making the lncRNA dysfunctional [62]. Another fascinating study began by mining GWAS for atherosclerosis disease-associated SNPs, which led to the discovery of a novel lncRNA, LIN00305, that had 5 SNPs that were associated with atherosclerosis all located within an intronic region [111]. The group went on to characterize LINC00305 in human primary and immortalized monocytes as a promoter of inflammation through activating the aryl-hydrocarbon receptor repressor (AHRR) – NF-κB pathway by directly binding to lipocalin-1 interacting membrane receptor (LIMR), acting as a scaffold to promote the interaction between LIMR and AHRR [111]. Both of these studies started by investigating clinically relevant SNPs that led to the discovery of disease-associated lncRNAs, providing the groundwork for future studies on potential biomarkers or the development of novel therapeutic targets for a variety of inflammatory diseases.

Table 2
Human sequencing datasets.

	Type	Treatment	Duration (h)	Sequencing	Reference
Spleen	Tissue	–	–	Bulk RNA ChIP - H3K36me3 ChIP - H3K4me1 ChIP - H3K4me3	*26030523*
Lung	Tissue	–	–	Bulk RNA ChIP - H3K36me3 ChIP - H3K4me1 ChIP - H3K4me3	*26030523*
Liver	Tissue	–	–	Single cell RNA	*30348985*
B cell	Lymphoblastoid cell line	TNF	6	ChIP - NFkB	*30361341* **GSE31477**
PBMCs	Healthy IRF7 −/− UNC93 −/−	Influenza	0, 8, 16	Bulk RNA	*25814066*
	Blood	Healthy SLE	Longitudinal study	Bulk RNA	*22675550, 30962246*
	Macrophages	Salmonella/Listeria	0, 2, 24	Bulk RNA	*27690314*
	Monocytes	TLR4	0, 1, 4	Bulk RNA	*25855049*
	(Blood) CD14+	–	–	ChIP - H3K4me3	*25258085*
		–	–	ChIP - H3K27me3	
		–	–	Bulk RNA	
	(Blood) CD14+ CD16+	–	–		
	Monocytes (PBMCs)	–	0, 4	ChIP-H3K4me3	*22675550, 28900427*
		–	0, 4	ChIP-H3K27Ac	
		Uninfected	–	Bulk RNA	
		Influenza, S. aureus, S. pneumoniae, E. coli	–	Bulk RNA	
		LPS	0, 4, 24	Bulk RNA	
	Monocyte bead purification	LPS	0, 4, 24, 144	Bulk RNA ChIP-H3K4me1 ChIP-H3K27Ac	*27863248*
		β-Glucan	0, 4, 24, 144	Bulk RNA ChIP-H3K4me1 ChIP-H3K27Ac	
Dendritic cells	(Blood) CD141+ (Blood) CD1c+	–	–	Bulk RNA	*28475900*
	Plasmacytoid	CpG-C	0, 1, 2	Single cell RNA	*30127440*
	Monocyte derived	Differentiation	–	RNA bulk	*24744378*
		Differentiation	–	ChIP - H3K27Ac	
		Differentiation	–	ChIP - STAT1	
		Differentiation	–	ChIP - IRF1	
		A. fumigatus	0, 6	Bulk RNA	*28280489*
		C. albicans	0, 6	Bulk RNA	
		TLR4 (LPS)	0, 6	Bulk RNA	
Macrophages	THP-1 (PMA)	Mtb	0, 4, 18, 48	Bulk RNA	*22675550*
		TLR1/2	0, 8	Bulk RNA	*24371310*
		TLR1/2	0, 8	Bulk RNA	*30918008*
		TLR2/6	0, 8		
		TLR4	0, 8		
	Monocyte derived	–	–	Bulk RNA	*29475453*
		Mtb	0, 24	Single cell RNA	*28192419*
		H1N1	0, 1, 3, 6	Bulk RNA	*29475453*
		N5N1	0, 1, 3, 6	Bulk RNA	

Additionally, as previously mentioned, XENA or Lnc2Catlas are databases that combine RNA sequencing of tissue samples from healthy and cancer patients, these tools can also be adapted to assess age, gender, tissue or even disease association of splice variants [98,110]. These disease-associated SNPs or splice variants will help guide further studies to determine therapeutic or biomarker potential. In summary, assessing disease association by a variety of databases can provide insight into the function of a lncRNA [112].

2.5. Conservation of lncRNAs

Unlike lncRNAs, coding genes are highly conserved across distal and related species. The requirement of coding genes to encode functional peptides likely constrains the variation within the open reading frame (ORF) sequence [113]. LncRNAs by definition are not translated and often have poor sequence conservation across related species. Many lncRNAs display species-specific expression. Thus, inferring function based on sequence similarity is a challenge (reviewed in [114]). A more useful conservation metric is to assess whether the lncRNAs have conservation of synteny (location relative to flanking coding genes), along with expression conservation [115]. The databases LNCipedia and NONCODE have user-friendly interfaces, allowing assessment of both conservation of synteny and sequence for lncRNAs. Additionally, expression conservation can be useful to assess whether the specific lncRNA has the same biological role in similar and divergent species [116]. Conservation of lncRNA expression can also be indicative of conservation of regulatory regions, such as transcription factor binding sites within promoters [117]. However, conservation of expression does not necessarily mean the RNA product is important for lncRNA function. Enhancer RNAs (eRNAs), for example, are thought to predominantly function by creating a localized, active transcriptional state, which can activate neighboring genes [118–120]. It is unclear to what extent the specific RNA sequence of eRNAs is important for their function [118]. Nevertheless, a couple of studies have provided

examples of eRNAs for which the transcript sequence was necessary for their function [121,122].

To further investigate conservation of a candidate, BLAST (basic local alignment search tool) on NCBI (national center for biotechnology information) can be utilized to explore conservation across species [123] or lncRNAdb v2.0 [124], by inputting the entire sequence of a lncRNA of interest. If a lncRNA has a known structure, inputting the shorter structured RNA sequence can enhance conservation results. One can also, view the conservation track of the UCSC genome browser [125] to assess if this specific sequence within the lncRNA transcript is conserved across species. In summary, conservation of a lncRNA is complicated to assess using current bioinformatic methods. However, understanding if there is a functional motif (therefore a shorter starting input sequence) within a lncRNA can allow for an increased assessment of functional conservation across species.

2.6. Transcriptional regulation of lncRNAs

Activation of inflammatory pathways result in both up and down regulation of specific lncRNAs, which in turn can have either positive or negative regulatory effects on the pathway, such as activation of sequestered transcription factors [45,126] or enhanced or repressed expression of specific inflammatory cytokines [61,84,127]. Genes that are immediate regulators of immunity are poised for transcriptional activation, which can be assessed by defining the openness of promoter regions of a lncRNA pre- and post- inflammatory stimulation. Common methods for assessing chromatin accessibility of promoters include DNase-hyper-sensitivity (DNAseHS) [128] and the Assay for Transposase-Accessible Chromatin (ATAC) [129]. DNAse HS-seq and ATAC-seq datasets are available from a variety of tissues on ENCODE for both mouse and humans (Tables 1 and 2). If the promoter is open (accessible), this is indicative of either a poised or actively transcribed gene. Accessibility of promoter regions in the hematopoietic cell lineage was assessed by Lara-Astiaso et al. through performing ATAC-seq for all cells in the hematopoietic lineage [76]. This dataset provides insight into a gene's promoter accessibility, as well as cell type specificity. For instance, if the promoter is open in all cell types, it shows that it is ubiquitously accessible and possibly expressed, while if a promoter is only accessible in myeloid cells or terminally differentiated macrophages this provides insight into the cell type that could be most biologically relevant for a particular lncRNA. Another interesting data set from Tong et al. have provided ATAC sequencing from bone marrow derived macrophages (BMDMs) pre- and post- inflammatory time course stimulation [72]. This dataset assesses both poised genes and genes that undergo promoter remodeling during inflammatory activation. If a promoter of a lncRNA is inaccessible or accessible during inflammatory stimulation, this could provide insight to its regulation and biological significance during an immune response.

Once the accessibility of a promoter region is determined, defining post-translational modifications of histones on promoter regions will assess promoter activity in specific cell types or inflammatory states [130–132]. A histone modification is a covalent post-translational modification (PTM) to histone proteins which includes methylation, phosphorylation, acetylation, ubiquitylation, and sumoylation (reviewed in [133]). Posttranslational histone modifications do not affect DNA nucleotide sequence but can modify chromatin availability to the transcriptional machinery [134]. Identifying the types of epigenetic histone marks will add additional layers of understanding if the candidate is active, poised, silenced or an enhancer. Many publicly available datasets provide this information outlined in Tables 1 and 2. While the examples we outlined here are immune focused there is a vast array of additional primary and immortalized cell line for every histone mark available through the easily accessible ENCODE database [101].

Finally, the transcriptional regulation of a lncRNA can be defined by analyzing the transcription factor (TF) motif's that lie within the predicted promoter region [135–137]. RNAReg2.0 [138] or HOMER [139,140] are useful tools to predict TF binding sites by motif analysis. TF motifs can be indicative of biological pathway regulation indicating when a gene is expressed. These findings of predicted TF can then be put into a gene ontology tool (PANTHER or DAVID) to assess how a candidate lncRNA is transcriptionally regulated. For instance, the presence of pioneering transcription factors could provide information on the cell type specificity of a lncRNA. On the other hand, if the promoter motifs are enriched for p65, Interferon Response Factors (IRFs) or Activating Transcription Factors (ATFs) this would be indicative of inflammatory specific expression. This finding can be further supported by Chromatin immunoprecipitation (ChIP) sequencing data sets from a multitude of labs, as well as data from the ENCODE project (outlined in Tables 1 and 2).

3. Functional characterization of lncRNAs

3.1. Assessing whether lncRNAs are translated

The categorization of lncRNAs as “noncoding” is initially determined bioinformatically using the arbitrary cut-off of < 100 codons [141]. This leaves the possibility that some lncRNAs may be mRNAs. Therefore, one of the first steps in characterizing a candidate lncRNA is to confirm that it is noncoding. For example, in Drosophila, a gene annotated as a lncRNA FBgn0087003, was shown to encode multiple small ~ ten amino acid peptides critical for development [142]. The discovery of these germline-encoded biologically active peptides has opened the door into new and exciting levels of regulation. However, this discovery also shows that we should be cautious when characterizing a lncRNA to confirm that indeed they are noncoding.

Several bioinformatic tools exist for predicting small ORFs (smORF) but have noted that the predictive ability for smaller ORF size is in general very poor [143]. PhyloCSF uses codon substitution frequency, together with conservation across multiple species, to provide a score metric that can be used to determine the presence of a conserved ORF [113]. Other approaches have sought to identify novel small peptides using mass spectrometry. However, a major challenge is determining whether the peptides identified correspond to novel smORFs or represent degraded intermediates of larger proteins [144].

The development of ribosomal foot-printing coupled to next-generation sequencing (ribo-seq) has provided a powerful quantitative method for assessing global translation [145,146]. While ribosome profiling allows for ribosome nuclease-protected RNA fragments to be mapped to transcripts to enable quantitative measurement of the translation efficiency [146]. In addition, to mapping ORFs, ribo-seq can be performed with a drug that stalls ribosomes at the start codon to globally map translation start sites—this revealed significant numbers of non-canonical translation initiation from CTG codons [146]. Ribo-seq has also found that many lncRNAs appear to be translated, raising the possibility that some of these transcripts could be producing small peptides [145]. Guttman et al. developed the ribosome release score metric as part of the ribosome profiling analysis pipeline to more accurately predict translational efficiency. Their findings show that ribosome occupancy of lncRNAs and 5′UTRs does not always equate to translation [147]. Furthermore, Guttman et al. concluded that most noncoding transcripts are not translated into peptides. Several additional studies have examined lncRNAs binding to ribosomes and have concluded that ribosome binding may not be functional and may serve as a quality control process to degrade transcripts with low coding potential *via* the nonsense-mediated decay (NMD) pathway [148,149]. A recent study by Jackson et al. used ribo-tagging in LPS-stimulated mouse macrophages to identify ribosome footprints within hundreds of annotated lncRNAs, raising the possibility that they may be producing functional peptides [150]. They characterized an 83aa peptide located within a previously annotated lncRNA Aw112010 and showed it was produced by non-canonical “CTG” translation initiation [150]. They demonstrated that this small peptide had a critical role in mucosal

Gene knock-down	Target	Reversible loss-of function?	Type of phenotype	Time to phenotype	Limitation(s)?
A. RNAi (RISC pathway)	Transcript	Yes	hypomorphic	Short	Minimal knockdown of nuclear lncs
B. GAPMERs/ASOs (RNase-H pathway)	Transcript	Yes	hypomorphic	Short	None
C. CRISPRi (KRAB)	gDNA (TSS)	Yes	hypomorphic	Medium	> 1000 bp from nearest coding gene TSS, can only target lincRNAs
D. CRISPR-splice-sites	gDNA	No	hypomorphic	Medium	Splice site cannot overlap coding gene
E. CRISPR-chomp	gDNA	No	Null (complete)	Medium/Long	Deleted region cannot overlap coding gene
F. poly(A) terminator	gDNA	Yes (loxP)	hypomorphic	Medium/Long	Region cannot overlap coding gene

Gene Activation	Target	Reversible gain-of function?	Type of phenotype	Time to phenotype	Limitation(s)?
G. CRISPRa	gDNA (TSS)	Yes	over-expression	Medium	> 1000 bp from nearest coding gene TSS, can only target lincRNAs
H. Ecotopic	cDNA	Yes (tet-on)	over-expression	Medium	cDNA < 5-6kb size (lenti constucts)

Fig. 3. LncRNA expression manipulation.
(A) RNA interference acts post-translationally, degrading transcripts in the cytoplasm. (B) Gapmers, are composed of RNA and core stretch of ~10 nt of DNA sequence that is critical for engaging the RNase H pathway. (C) CRISPRi uses a catalytic inactive version of Cas9 fused to a KRAB domain. The CRISPRi system can target transcription start sites (TSS) to induce heterochromatin-based transcriptional silencing. CRISPR/Cas9 can be used to target (D) splice sites or (E) to delete specific regions of lncRNA loci using two flanking gRNAs. (F) Polyadenylation signals (3–5) can be inserted just downstream of the TSS, which result in premature termination of transcription. These poly(A) signals can be flanked by loxP sites to allow their removal. (G) The CRISPRa system can target transcription start sites (TSS) to induce transcriptional activation. (H) LncRNA sequences (cDNA) can be cloned into plasmid either with native promoter or a constitutively active promoter (*i.e.* EF1a, CMV). The different tools are compared base on criteria described in the subheadings of each column. 'Target' describes what aspects of the lncRNA's gene expression are targeted, for example transcriptional (CRISPRi, C) or posttranscriptional silencing (RNAi, A). 'Reversible loss-of-function?' is the silencing effect of this tool reversible in a cell? In general, only the nuclease active Cas9 activity (D) is not reversible since the regions of the loci are actually deleted. 'Type of phenotype' refers to whether the tool completely shuts off expression of the lncRNAs. Hypomorphic refers to the fact that some level of expression remains despite the level of knock-down. Only compete removal of the locus completely abolishes its expression. Knock-down of nuclear lncRNAs? Touches on the observation made by many groups, which have demonstrated that RNAi is not very active in the nuclear (in contrast to the other technology). 'Time to phenotype' refers to the amount of time from planning experiment to obtaining phenotype. *Short:* transfection of RNAi and ASO (A and B) remain the quickest route to phenotype since they don't require cloning or modification to the cell. *Medium:* CRISPRi/a (C, G) requires design and cloning guide RNAs to a specific locus in addition to generation of a functional CRISPRi/a cell line. Medium/Long: Both CRISPR mediated deletions and adding poly(A) signal (D and E) require screening for cells (making single cell clones) to identify those with successful modifications.

immunity in mice and was specifically required for the expression of Il12 mRNA [150]. The exact mechanism of how this peptide drives Il12 expression is yet to be elucidated.

4. LncRNA expression manipulation

4.1. RNA interference (RNAi)

The laborious process of functionally characterizing lncRNAs has remained a major limitation to lncRNA functional discovery. For example, while ~16,000 long noncoding RNAs (lncRNAs) have been identified in the human genome, only 3% of all validated lncRNAs have an ascribed function [47,91,151]. For more than two decades since its discovery, RNA interference (RNAi) has been the method of choice for loss of function studies. The ease and versatility of RNAi make it appealing for use, requiring short complementary small RNAs transfected into cells that can then utilize the endogenous cellular machinery to target specific transcripts [152,153] (Fig. 3A). Additionally, short hairpin RNAs (shRNAs) can be expressed in a lentiviral context to accomplish stable RNAi in cells [154]. RNAi is most active in the cytoplasm, which makes it most useful for targeting mRNAs and lncRNAs that reside in the cytoplasm [155]. Success in knocking down some lncRNAs, for example lincRNA-Cox2, has been attributed to the fact that this lncRNA is expressed both in the cytoplasm and the nucleus (perhaps cycling between both compartments) and hence is susceptible to RNAi [42]. However, many lncRNAs are thought to be nuclear-restricted where RNAi has been demonstrated to have limited efficiency [156] (Fig. 3).

4.2. Antisense oligonucleotides (ASOs)

Antisense oligonucleotides (ASOs) can be either RNA or DNA-based and can be used to target complementary sequences within a transcript. Unlike RNAi, ASOs do not engage the cellular RNAi machinery [157]. Instead, ASOs function by hybridizing to the target RNA and inhibiting its function by either inducing the RNase H pathway or by steric inhibition. The RNase H enzyme is part of a cellular pathway that normally functions to resolve unwanted DNA: RNA interactions that can

occur during replication and/or transcription [158]. One of the most widely used types of ASOs for targeting lncRNAs are gapmers which contain a "hybrid" modified/unmodified configuration consisting of ~10 nt DNA core flanked by 2′-*O*-Methyl or LNA-modified synthetic nucleotides (Fig. 3B). Gapmers offer the benefit of modifications for stability and reduced toxicity, while still allowing engagement of the RNase H pathway (Fig. 3B) [159].

Efficient depletion of nuclear lncRNAs has been demonstrated using modified DNA anti-sense oligonucleotides [155]. Ilott et al. used gapmers to successfully knock-down TLR-induced enhancer RNA, as they were unable to knock-down these same eRNAs with 4 different siRNAs [56]. Recent *in vivo* applications for DNA oligonucleotides have also proven successful as a possible treatment of a neural degenerative disease called Angelman's syndrome. Angelman's syndrome is a monogenic disorder caused by mutations in the E3 ubiquitin ligase 3A (UBE3A) [160,161]. As UBE3A is a paternally-imprinted gene, a mutation in the maternal allele is sufficient to lead to disease [160,161]. Paternal imprinting of UBE3A requires the expression of an antisense lncRNA called UBE3A-ATS. Therefore, targeting the paternal lncRNAs UBE3A-ATS using ASOs leads to de-repression of expression of paternal UBE3A, allowing the rescue of the defective maternal copy [161]. In a mouse model, they showed that even partial restoration of the UBE3A protein expression ameliorated some cognitive deficits associated with the disease [161].

Mechanistically, it is important to dissect what features of a lncRNA are important for its function — for example, determining whether the lncRNA transcript is required or whether the mere act of transcription is the important feature needed to mediate its function (discussed further below). siRNAs and ASOs are thought to inhibit gene expression at the posttranscriptional level, albeit by different mechanisms [162,163]. Nevertheless, there are examples of siRNAs (reviewed in [164]) mediating transcriptional inhibition. Targeting sequences near the 5′ end of the transcript can induce the "torpedo-effect," resulting in pre-mature termination of transcription [165,166]. Therefore, when targeting lncRNAs with ASOs or siRNAs, it's important to consider where along with the transcript you are targeting to ensure biological interpretation of the ablation is correct.

4.3. CRISPR/Cas technology

The introduction of CRISPR/Cas9 technology has revolutionized the field of functional genomics by providing a novel tool for interrogating gene function. CRISPR/Cas9 is a deoxyribose nuclease (DNase) that can be specifically targeted to genomic regions *via* a guide RNA (gRNA) [167,168]. Targeting of Cas9 to a region results in a blunt double-stranded DNA break that engages the cellular Non-Homologous End-Joining (NHEJ) DNA repair pathway, which promotes imprecise repair, yielding small deletions in the repaired sequence. These small deletions result in a frame-shift that disrupts the ORFs within coding genes, thereby disrupting protein synthesis. LncRNAs do not contain ORFs and span tens of kilobases of sequence in size. As such, targeting them with a single gRNA is thought to be insufficient to disrupt their function [169]. An alternative application of Cas9 that has proven effective for targeting lncRNAs involves using two gRNAs flanking the lncRNA region of interest to induce deletion of the entire locus (Fig. 3E). The main advantage of deleting lncRNAs is obtaining complete loss-of-function, as demonstrated in a recent study that performed a CRISPR-mediated deletion-screen to identify lncRNAs that positively and negatively regulate cancer growth [170]. On the other hand, deletion of the DNA sequence may result in an inability to resolve whether a phenotype is due to loss of lncRNA production or loss of DNA sequence (discussed below) [171]. In an alternative approach, Liu and colleagues recently performed a CRISPR screen targeting the splice sites of over 10-thousand lncRNAs and identified 230 lncRNAs that proved essential for viability (Fig. 3D) [172]. Targeting of splice sites has also been shown to induce exon skipping within coding genes [173]. LncRNAs have many of the same regulatory elements as coding genes. Hence future studies may opt to target TF binding sites, secondary structure and/or polyadenylation sites as a way to more finely dissect the functional portions of lncRNAs.

The ease in which Cas9 can be targeted to specific genomic regions sparked the development of a modified (catalytically inactivated) version of the protein fused to the KRAB (Krüppel associated box) chromatin-silencing domain termed CRISPRi [174,175] (Fig. 3C). CRISPRi can be used to target both coding and noncoding genes (such as lncRNAs), triggering localized heterochromatin-silencing at the transcription start site (TSS) [176]. Unlike RNAi, the transcription-based inhibition by CRISPRi offers the ability to efficiently target lncRNAs regardless of their localization in the cell. Gilbert et al. have shown maximum knock-down efficiency when targeting regions +50 to +500 nucleotides relative to the transcription start site (TSS) [174]. However, CRISPRi is limited by the availability of an accurately annotated TSS, particularly for lncRNAs. Incorporating CAGE-seq, Gro-seq and/or ChIP-seq data into gRNA design may improve efficient targeting of a lncRNA of interest [177]. Nevertheless, a recent study from the Weissman and Lim groups utilized CRISPRi to target 16,401 lncRNA loci in 7 diverse cell lines and identified hundreds of lncRNAs required for cell growth [96]. Therefore, despite its caveats, it appears that CRISPRi can be a useful and powerful tool for interrogating lncRNA biology.

An alternative approach to gene ablation is overexpression, gain-of-function. Plasmid-based overexpression systems have been used for decades to overexpress specific coding genes [178]. Plasmid-based over-expression of lncRNAs is possible but is limited by the size (max size: 6-8 kb) of the specific lncRNA (unpublished observations). However, because lncRNAs can function in *cis*, it's important that there is the significant mechanistic characterization of the candidate lncRNA to justify its cloning into a plasmid.

The same catalytically inactivated Cas9 has also been fused to a transcriptional activation domain, for example, the VP16, a strong transcriptional activator derived from herpesviruses [179,180]. This strategy allows CRISPR-based transcriptional gene activation that can be used to study gain-of-function phenotypes. The Zhang group recently performed a CRISPR activation screen targeting over 10-thousand lncRNAs [181]. They identified 11 lncRNAs that upon activation, mediated BRAF inhibitor resistance in melanoma cells [181]. One of the advantages of using a CRISPRa library is that the same library can be used across different cell types, which makes it more cost effective. Nevertheless, several caveats to the CRISPRa system, include the possibility that high expression of lncRNAs may create non-physiological conditions leading to incorrect conclusions of specific biology.

LncRNAs have been shown to mediate functions *via* binding to specific proteins. Hence over-expression of lncRNAs, without its protein partner, may result in the inability to identify important biology. In conclusion, it's important to understand the advantages and disadvantages of the loss- or gain-of-function methods used to modulate lncRNA expression. All methods have their caveats, and these are important to consider when deciding which method to use.

4.4. Dissecting the complex functions of lncRNAs in vivo

LncRNAs can span large stretches of DNA sequence and can contain important regulatory regions, such as enhancers, that are functionally independent of the lncRNA product [182]. LncRNA promoters have been proposed to also function as enhancers, promoting the recruitment of transcriptional-activating factors that can affect the local nucleosome environment and ultimately the expression of neighboring genes [118,183,184]. Nevertheless, *in vivo* assessment of lncRNA function has predominantly relied on assessing the consequence of deleting entire lncRNA loci [171,185]. Numerous studies involving deletion of lncRNA loci have been unable to rescue the deletion phenotype using a transgene approach, making it difficult to attribute the phenotype to the

lncRNA product itself [186,187]. While deleting the entire lncRNA is a useful first step to establish a phenotype, this approach can make it difficult to identify which component of the lncRNA is important for the observed phenotype [171,185,188–190].

There are many approaches to generating a lncRNA knockout/knockdown mouse. For instance, the complete knockout is the easiest first step to take, which can be designed to remove the entire gene. After this approach, one can use a more fine-tuned approach to reveal exactly how the gene is working, including deleting specific regions of the gene or inserting a poly-adenylation cassette before the exon1 of the gene. A well-documented example of this phenomena is *Fendrr*, which was shown to have a lethal phenotype in two independent studies, but the importance of the lncRNA in development differed due to the mouse ablation strategy. One *in vivo* study generated a knockout mouse by removing the genomic loci completely and replacing it with LacZ while leaving the native promoter intact [185]. This study identified *Fendrr* as a key regulator of lung development and mesenchymal differentiation. However this study did not attempt to rescue the ablated *Fendrr* allele *via* a transgene [185]. An additional group generated a *Fendrr* knockdown mouse using an alternative approach. Instead of disrupting the chromatin architecture or removing any possible DNA enhancer regions, they inhibited the transcription of *Fendrr* through the insertion of a poly-A cassette into exon1 [191]. Using this lncRNA knockdown mouse design, the scientists determined that loss of *Fendrr* led to heart and body wall defects, which is slightly different from the phenotype in the KO mouse study. Importantly, the heart and body wall defects caused by the terminator insertion were rescued by a transgene of *Fendrr in vivo*, further confirming that the phenotype is due to the RNA while the KO phenotype could be due elements within the genomic DNA [191]. *Fendrr* is located ~4 kb downstream of Foxf1 and ~12 kb upstream of Irf8. The observed phenotype of the KO mouse is possibly due to the deletion of an enhancer element that could impact the expression of protein-coding genes, which could have roles in lung development [192,193].

We recently used multiple genetic mouse models to dissect both *cis* and *trans* functions of lincRNA-Cox2 *in vivo* [46]. Working with a complete lincRNA-Cox2 knockout [185], in which the gene is replaced with a LacZ cassette, we observed a strong *cis* defect on the neighboring protein-coding gene Ptgs2. From these studies, we concluded that lincRNA-Cox2 functions *in cis* through an enhancer RNA mechanism to regulate its neighboring gene Ptgs2 [46]. In order to determine how lincRNA-Cox2 functions *in trans* to regulate genes independent of its cis effects, we generated a mutant/intronless mouse by targeting the splice sites of lincRNA-Cox2 using CRISPR/Cas9. This mouse represents a knockdown mouse, and because there is a low level of transcription of lincRNA-Cox2, Ptgs2 levels are the same as WT. LincRNA-Cox2 is not inducible in the mutant mouse, probably because of transcript instability due to the lack of splicing, enabling us to study its trans regulatory roles. Similar to our early *in vitro* work we observed using an LPS shock model that many genes are both up and downregulated in the serum of mutant mice indicating that lincRNA-Cox2 can indeed function *in trans* to regulate immune genes *in vivo* [46].

To prove genetically that a lncRNA is functioning in *trans,* a *trans* rescue experiment can be performed using transgenic mice that constitutively express a lncRNA. This rescue strategy has been utilized in some studies including *Evf2* [194,195], *Jpx* [196], as well as *Pnky* [197] which both demonstrate successful rescue experiments where the phenotype from germline ablation of a lncRNA is rescued through generation and crossing with a transgenic animal.

5. Understanding the mechanism of action of a lncRNA

5.1. Subcellular localization and binding partners

LncRNAs are immensely adaptable molecules that are capable of working through RNA-RNA, RNA-DNA, or RNA-protein interactions. RNA-directed technologies such as Chromatin Isolation by RNA Purification (ChIRP) [198,199] or RNA antisense purification (RAP) [200,201], will help uncover lncRNA interactomes for RNA, genomic or protein partners for highly expressed candidates. If a candidate is lowly expressed, one can exogenously introduce a biotinylated form of the lncRNA using the RNA pull-down method, which we have successfully used to identify binding partners for lincRNA-Cox2. This has been performed for many lncRNAs [42,202,203].

Functions of lncRNAs are associated with their subcellular fates. Web servers can assist in quickly assessing experimentally determined or predictive RNA-RNA interactions [204] or even RNA-protein interactions [205,206]. Depending on the subcellular or extracellular compartmental localization of a lncRNA, this patterning will elude to the regulatory role of the gene, as well as how a lncRNA might execute its function [207–209]. Some LncRNAs that are localized to the nucleus or chromatin have been experimentally shown to function *in cis* to regulate the transcriptional expression of a neighboring gene, or *in trans* to regulate the transcriptional regulation of a subclass of genes through the interactions between heterogeneous nuclear ribonucleoproteins (hnRNPs) [42,54,60,210]. In the cytosol, some lncRNAs have been shown to interact with RNAs and proteins to carry out their molecular functions [42,79,211,212].

Cellular localization of lncRNAs can be predicted using a publicly available user-friendly web server established at http://lin-group.cn/server/iLoc-LncRNA [213]. This can be the first step by a researcher to attempt to predict the localization of your candidate based off of its sequence. iLoc-LncRNA predicts subcellular location of a lncRNA by utilizing the 8-tuple nucleotide features into the general PseKNC (Pseudo K-tuple Nucleotide Composition) and rigorous tests show the overall accuracy achieved by the new predictor is 86.72%, which is over 20% better than previous algorithms [213]. In addition to prediction methods, there are also publicly available sequencing datasets that have performed RNA sequencing on fractionated cells. Bhatt et al. utilized murine bone marrow derived macrophages, with and without inflammatory stimulation, and fractionated these cells into chromatin, nuclear, and cytoplasmic compartment to determine RNA localization [74]. This data set can now be utilized to investigate the localization of any murine candidate expressed in macrophages. A recent study took this question a step further by performing RNA sequencing on nine separate locations with a cell including: nucleus, nucleolus, nuclear lamina, nuclear pore, cytosol, endoplasmic reticulum membrane (ERM), outer mitochondrial membrane (OMM), mitochondrial matrix (MITO), and endoplasmic reticulum lumen [214]. This exciting study utilized an APEX-sequencing method, where the peroxidase enzyme APEX2 was localized to these nine separate locations in nine separate Human embryonic kidney (HEK) 293T cell lines. APEX2 can biotinylate nearby RNA molecules allowing for streptavidin-based immunoprecipitation and RNA sequencing. The APEX-seq datasets will provide a powerful resource for referencing localization of specific lncRNA candidates that are expressed in HEK293T cells [214].

There are a few commonly used experimental approaches that can be used to validate and determine the localization of a lncRNA. Subcellular chromatin, nuclear and cytoplasmic fractionations of any primary or immortalized cells can be prepared using previously published procedures [74,215], followed by RNA isolation and RT-qPCR to assess the localization. If additional compartmental fractionation is desired other compartments can be enriched, such as mitochondria [216,217]. Another standard gold technique is to visually determine cellular localization by RNA FISH [42]. RNA localization can also be directly visualized by microscopy [218]. SeqFISH techniques have recently been pioneered for imaging thousands of cellular RNAs at once using barcoded oligonucleotides [219]. The drawbacks of these *in-situ* fluorescence hybridization (FISH) based approaches, however, are the need for cell fixation and permeabilization, which can re-localize or extract cellular components [220]. In addition to the difficulty of assigning RNAs to specific organelles or cellular landmarks due to spatial

resolution limits, some of these difficulties can be overcome with the addition of stains for markers of specific organelles.

5.2. Determining structure or motifs within a lncRNA

Conventionally, lncRNAs display poor sequence conservation across species, with the exception of finite regions of conserved bases surrounded by large seemingly unconstrained sequences [221]. While sequence conservation does not constrain lncRNA genes, lncRNA function is found to be conserved across species when identifying motifs or structure. A great example of this phenomena is represented in the study of human maternally expressed gene 3 (MEG3), which utilized the computer program *mfold* and multiple *ex vivo* and *in vitro* chemical probing techniques to identify common motifs critical for retained function in orangutan, rat, mouse and pig [222,223]. Another study highlighting the marsupial Rsx lncRNA was initially found to have no linear sequence similarity with the lncRNA Xist. However, it shared substantial levels of non-linear conservation within k-mer repeats that share functionally analogous protein-binding domains [224]. Publicly available web servers can be utilized to determine the RNA structure of a lncRNA depending on the size. The caveats to these web servers are they do not work efficiently with large transcripts. To overcome this, one can attempt to identify the critical sequence within the lncRNA that could be functional using RIP-sequencing databases, which allow you to input a gene ID to identify possible RNA binding proteins (RBPs). Additionally, if a lncRNA has already been identified to bind to a protein (s), RIP-sequencing databases can elucidate the specific binding location(s). Knowing the location of RNA-protein interaction narrows down the sequence input, that can be used for many web servers, which could enhance the elucidation of a predictive structure. If there is no information about potential protein binding partners, then these RNA structure webservers will be of little use.

An alternative approach to further investigate structure is to utilize bioinformatic tools which now include parameters for covariation. Covariation analysis identifies the positions in an RNA molecule that have similar patterns of variation and the purpose of this covariance is due to structural constraints initially shown for ribosomal RNA [225] and now also for lncRNAs [123,226]. This study predicts structures for MALAT1, using over 130 vertebrate sequences, as well as lncRNAs RepA and HOTAIR [227]. These powerful tools allow scientists to predict structures in an RNA molecule based on covariance, which in turn drive the next steps of experimentally validating these findings.

Fortunately, there are several *in vitro* and *in vivo* experimental techniques used to assess RNA structure for any size, even up to 17,000 nt (XIST). Dimethyl sulfate (DMS) probing uses a base-specific reagent that can bind and alter the methylation state of unpaired adenosine and cytosine nucleotides [228,229]. DMS "footprinting" is optimized for structural analysis of RNA. Protein binding to RNA will generate a "footprint" that can be traced due to alterations in the RNA structure. The transcript size that can be evaluated is rather small (< 500 nt) but this method can be performed both *in vitro* and *in vivo* as DMS can easily penetrate the cell membrane, shown to work for lncRNAs from 590 nt, Braveheart (Brvht) [230]. Xue et al. utilized DMS and SHAPE to determine the multiple smaller order structures of Brvht, including an AGIL motif and a 90-degree turn [231]. This AGIL motif is critical for transcription factor binding, which specifies the cardiovascular lineage. Targeting Structure-Seq relies on RNA methylation by DMA being performed *in vivo*. Using this method, structural models of elements within Xist were developed [232]. SHAPE (selective 2′hydroxyl acylation by primer extension), as well as the modified SHAPE-MaP [233], In-cell SHAPE-Seq [234] and icSHAPE-seq [235], can interrogate the RNA structure both *in vitro* and *in vivo* using the chemical NMIA and its derivatives to detect flexible regions in RNA secondary structure [236,237]. This method has been proven valuable for Xist [238], RepA [239], and Rox1/2 [240]. PARIS (psoralen analysis of RNA interactions and structures) was recently developed to determine both RNA structure and interactions *in vivo* [241]. Using this approach, a model for the higher order structure of Xist was interrogated [241]. These approaches are critical for identifying structural motifs and enhance conservation studies, as well as these identified elements, can be used as novel targets for further exploration in precise intervention suitable for therapeutic applications.

5.3. Alternative splicing of lncRNAs

Alternative splicing (AS) significantly impacts the diversity of RNA isoforms produced, which in turn impacts the protein isoforms produced and can affect many aspects of the protein's biology including binding, intracellular localization, enzymatic activity, stability, post-translational modifications [242]. AS also impacts lncRNA genes which can have multiple isoforms depending on the cell/tissue, age, and disease state [243–245]. The UCSC genome browser [125], as well as NONCODEV5 [92], have all the annotated isoform transcripts for each gene. These tools will identify annotated transcript isoforms while RNA sequencing will provide information on which of these splicing events is utilized in a given cell or biological state. To date, there have only been a small number of papers focusing on the role that alternative splicing plays in controlling the immune system. One recent global study has shown that widespread shortening of 3′ untranslated regions and increased exon inclusion are evolutionarily conserved features of innate immune responses in primary human macrophages following *Listeria monocytogenes* and *Salmonella typhimurium* infection [78]. This is a transformative study for mRNAs but can be reanalyzed to examine AS and possible contributions from lncRNAs.

Publicly available tools for tissue isoform expression specificity is available on GTEx [97] and XENA [98] for human genes. Tabula Muris [100], a murine specific dataset, is now available as a UCSC genome browser track on mm10 and can be used to view cell-type specific splicing events and isoform expression. In order to identify isoforms in illumina RNA sequencing datasets, several tools can be utilized. MISO (Mixture of Isoforms) is software for a probabilistic model for RNA seq will identify specific the 5′ splice sites used for each isoform [246]. Two other tools highly used for splicing analysis are JuncBASE [247], MAJIQ-SPEL [248], and DRIMSeq [249]. These tools can be used to define if your candidate gene undergoes any alternative splicing (alternative start site, exon inclusion/exclusion or alternative last exon) during a specific biological process for example following inflammatory activation. These tools are limited because of their dependence on a fully annotated transcriptomes. Therefore if a lncRNA is unannotated or has unannotated transcriptional isoforms, these events will not be captured. To overcome the limitations of incomplete transcriptomes, researchers can perform *de novo* transcriptome assembly using short RNA-Seq reads [250–252].

The future of RNA sequencing is headed towards long read sequencing, which is being met by Pacific Biosciences Single Molecule Real Time sequencing technology (PacBio) and Oxford Nanopore Technology (ONT) [253,254]. While both powerful technologies perform long read sequencing, their platforms are very different. PacBio technology is dependent on sequencing-by-synthesis. A DNA polymerase incorporates nucleotides that each have a corresponding conjugated fluorescent dye. The DNA polymerase works at a rate of 1000 bp/s, which is beyond the capabilities of current technologies. However, by circularizing the DNA PacBio has overcome this limitation through continuous long read sequencing, resulting in ability to generate 500k-4million reads at an error rate of below 1% [255,256]. On the other hand, ONT's approach relies on a pore embedded in a membrane. As a long cDNA or RNA strand translocates through the nanopore at single nucleotide precision from enzymatic regulation, the ionic current across the membrane is recorded. This technology can sequence full-length transcripts and can yield up to 10 million reads on the MinION or up to 60 million reads on the PromethION for cDNA [257,258]. An initial limitation of ONT was the 5–10% per read error

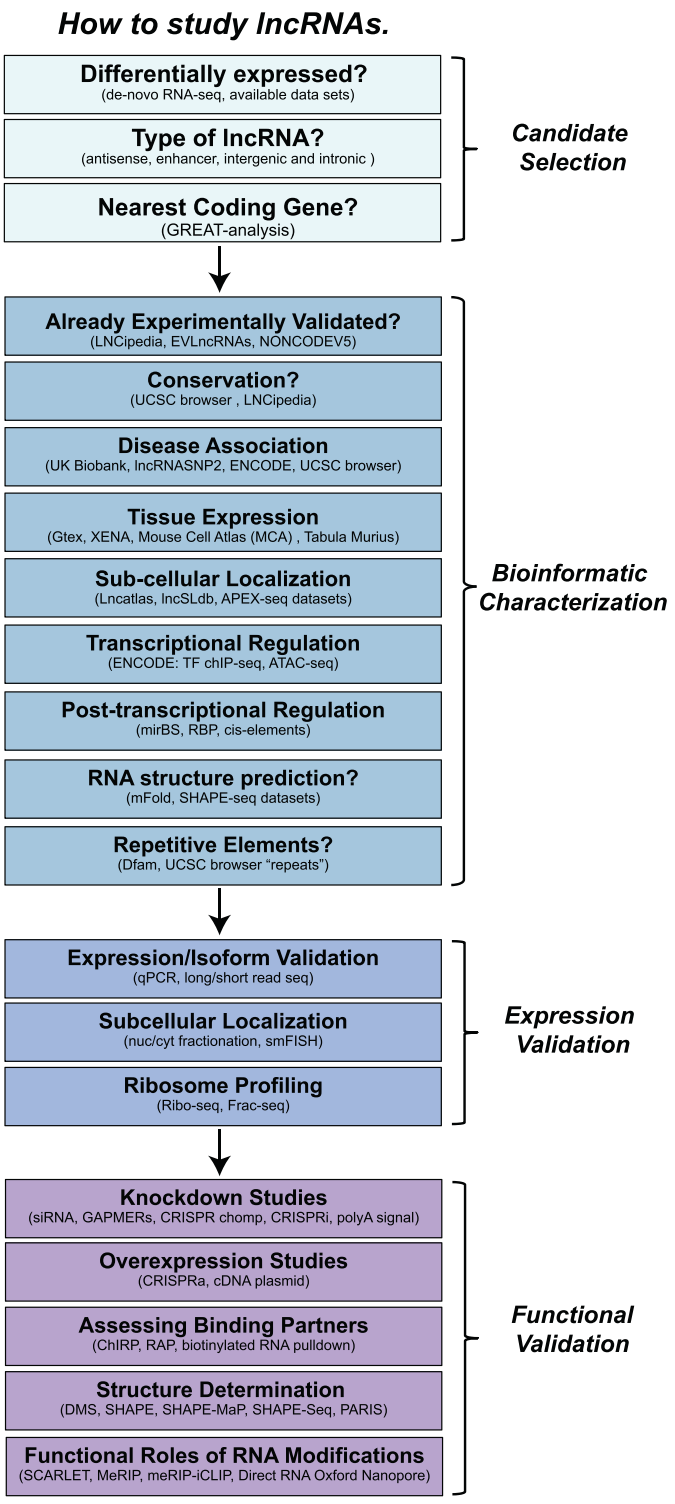


Fig. 4. How to study lncRNAs?.
The flow-chart provides a "beginning to end" guide to study lncRNAs. Not all suggested databases will be appropriate for all lncRNAs being studied. A. Selection of candidate lncRNAs should factor in changes in expression, type of lncRNA and nearby coding genes. B. Bioinformatic characterization of lncRNAs can be done using a variety of online databases including: LNCipedia to assess conservation or using the Mouse Cell Atlas (MCA) to assess tissue specific expression. C. Expression validation: This includes validation of expression and confirming that the lncRNA is in-fact non-protein coding. D. Functional validation dives into the final stage of mechanistic characterization of a lncRNAs, which involved manipulating its expression, as well as uncovering the specific cis-elements within the transcript important for its function.

rate, which has been overcome with a new technology called Rolling Circle to Concatemeric Consensus (R2C2) bringing the error rate down to 2.5% by increasing the read coverage. Overall, both of these technologies have overcome the transcriptome assembly and isoform identification limitations of short-read Illumina sequencing [259].

As stated above, ONT and PacBio can perform long read cDNA sequencing [260] and even more exciting, both technologies can perform direct long read RNA sequencing [261,262]. The beauty and simplicity that ONT and PacBio offers is the ability to sequence the full expressed isoform (cDNA and direct RNA), without worries of misidentifying a complex splicing pattern, RNA cleavage events and also not relying on transcript annotation files will lead to the identification of novel isoforms which are problems faced using short-read Illumina sequencing. For lncRNAs there have been a couple of studies that focus on isoform specificity and function for a given lncRNA. Neat1 [263] and lncRNA-PXN-AS1 [264] are two studies that show how one gene can have different functions depending on the RNA isoform expressed. While these studies are not immunology specific, this field is still at the early stages and we anticipate it becoming more prevalent in future studies.

5.4. RNA modifications of lncRNAs

RNA modifications are widespread and diverse in chemical nature, as well as highly conserved in their occurrence and function throughout species. RNA modifications function to affect RNA stability, localization, alternative site of poly-adenylation, and more [265]. Since lncRNAs can function as decoys and scaffolds, which are highly dependent on RNA structure, a single modification can enhance or eradicate this RNA-protein interaction. As you study a lncRNA, the mechanism of this molecule could be dependent on a modified nucleotide.

There are many techniques used to determine a single RNA modification in a cell type and biology of choice. Site-specific Cleavage and Radioactive-labeling followed by ligation-assisted extraction and TLC (SCARLET) Technology give scientists the ability to probe for N6-methyladenosine RNA (m6A) modification status at single nucleotide resolution in mRNA and long noncoding RNA [266]. The significance of RNA modifications to the control of the immune response is beginning to be appreciated. A study by Winkler et al. showed that m6A modification controls the innate immune response to infection by targeting type I interferons [267]. A few recent studies have shown that lncRNAs do have RNA modifications such as MALAT1 containing m6A modifications [268,269], HOTAIR containing m5C and m6A [270,271] and XIST containing ψ, m6A and m5C modifications [269]. A study by Zhou et al. showed that the RNA modification, m6A, acts as a structural 'switch' in Malat1. When there is a modification at site 2515, it results in an increased ability to bind hnRNPG, while a modification at 2577 leads to an increase in binding to hnRNPC [269]. In clinical research, lncRNA RP11-139J23.1 is highly expressed in colorectal cancer cells (CRC), and this specific upregulation was controlled by m6A methylation [272]. The study showed that m6A could regulate the lncRNA, which in turn triggered the dissemination of CRC cells *via* post-translation upregulation of the protein Zeb1. This novel study, connecting the interplay of RNA modifications and lncRNAs, has paved the way for a novel predictive biomarker or therapeutic target in CRC [272].

There are over 160 identified RNA modifications, while only a few have been studied to any extent [273]. Of these RNA modifications, the way they are enriched for in analysis is through an assortment of techniques including methylated RNA immunoprecipitation (MeRIP), MeRIP-iCLIP (crosslinking and immunoprecipitation), Suicide enzyme trap and Clickable chemicals (Reviewed in [274]). These techniques have many limitations and biases, but hopefully, future studies using direct RNA nanopore sequencing will overcome all these pitfalls. In a recent study, direct RNA sequencing using nanopore technology showed detection of m6A modifications with a 97% accuracy with the design of synthetic sequences [275]. As the performance of the algorithm increases, use of this tool will be extremely insightful when

analyzing myeloid or lymphoid primary cells with and without a treatment to understand how RNA modifications are regulated in innate immunity and specifically as it relates to our long noncoding transcriptome.

6. Conclusions and future insights for the field

LncRNAs, including XIST and H19, have been studied intensely for decades [275,276]. At the time we had no idea that these genes would represent the largest family of RNA genes produced in the genome. As Louis Pasteur once said, "Chance favors the prepared mind," and this is especially true following the development of next-generation sequencing. RNA-sequencing provided an unprecedented insight into the human genome. We did not identify new proteins, instead we found a wealth of noncoding RNA transcripts. The lncRNA field is growing at a blistering pace with labs from all aspects of biology, and now immunology branching out to include questions about the regulatory impact of these pervasive long noncoding gene species. As detailed in this review, there are many publicly available datasets and web servers that will streamline how to begin a lncRNA project, from how to pick a lncRNA candidate by interrogating published RNA-sequencing data, to determine the best tools to use to study the function and mechanism of a candidate (Fig. 4). Since this field is still at an early stage in its development, there are some shortcomings, including poorly annotated lncRNA transcripts. However, this will be overcome with direct RNA sequencing using ONT and PacBio technology. These technologies will enable us to determine the exact isoforms of transcripts expressed in a particular cell and begin to catalog the different RNA modifications that exist basally and during a biological process such as activation of inflammation. Since lncRNAs are cell-type specific in their expression patterns continued development of single-cell sequencing technologies will provide a complete catalog of lncRNAs in the genome. As the list of annotated lncRNAs grows, characterizing the function of all these genes has become a definite bottle-neck in the field. However, high-throughput CRISPR screening provides an approach to quickly identify functional lncRNAs in a particular biological system. Utilizing all the tools outlined here should enable researchers to develop this field rapidly. For our research focus, gaining a better understanding of the role of lncRNAs in regulating immune responses will provide novel insights into the molecular mechanisms governing inflammation. This data will be critical for identifying new avenues for therapeutic intervention for infectious and inflammatory disease.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Acknowledgements

Support from : National Institute of Health AR070973 and the Tobacco Related Disease Research Program 27IP-0017H to Susan Carpenter.

References

[1] M. J. Hangauer, I. W. Vaughn, and M. T. McManus, "Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs," PLoS Genet., vol. 9, no. 6, 2013.
[2] P. Carninci, T. Kasukawa, S. Katayama, J. Gough, M. Frith, N. Maede, The transcriptional landscape of the mammalian genome, Science (80-.). 309 (5740) (2005) 1559–1563.
[3] A.R.R. Forrest, et al., A promoter-level mammalian expression atlas, Nature 507 (7493) (2014) 462–470.
[4] T. Derrien, et al., The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression, Genome Res. 22 (9) (2012) 1775–1789.
[5] M.K. Atianand, K.A. Fitzgerald, Long non-coding rnas and control of gene expression in the immune system, Trends Mol. Med. 20 (11) (2014) 623–631.
[6] M. Wery, et al., Bases of antisense lncRNA-associated regulation of gene expression in fission yeast, PLoS Genet. 14 (7) (2018) 1–22.
[7] N. Hamazaki, K. Nakashima, K. Hayashi, T. Imamura, Detection of bidirectional promoter-derived lncRNAs from small-scale samples using pre-amplification-free directional RNA-seq method, Methods Mol. Biol. 1605 (2017) 83–103.
[8] M. Ding, Y. Liu, X. Liao, H. Zhan, Y. Liu, W. Huang, Enhancer RNAs (eRNAs): new insights into gene transcription and disease treatment, J. Cancer 9 (13) (2018) 2334–2340.
[9] I. Ulitsky and D. P. Bartel, "XLincRNAs: genomics, evolution, and mechanisms," Cell, vol. 154, no. 1, p. 26, 2013.
[10] Y. Tutar, Pseudogenes, BioEssays 2012 (2012) 253–258.
[11] J. Jarroux, A. Morillon, M. Pinskaya, History, discovery, and classification of lncRNAs, Adv. Exp. Med. Biol. 1008 (2017) 1–46.
[12] H. Rohrig, J. Schmidt, E. Miklashevichs, J. Schell, M. John, Soybean ENOD40 encodes two peptides that bind to sucrose synthase, Proc. Natl. Acad. Sci. 99 (4) (2002) 1915–1920.
[13] D.M. Anderson, et al., A micropeptide encoded by a putative long noncoding RNA regulates muscle performance, Cell 160 (4) (2015) 595–606.
[14] B.R. Nelson, et al., A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle, Sci. Reports 351 (6270) (2016) 271–275.
[15] A. Matsumoto, et al., MTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide, Nature 541 (7636) (2017) 228–232.
[16] I. Legnini et al., "Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis," Mol. Cell, vol. 66, no. 1, p. 22–37.e9, 2017.
[17] B. S. Razooky, B. Obermayer, J. B. O'May, and A. Tarakhovsky, "Viral infection identifies micropeptides differentially regulated in smORF-containing lncrnas," Genes (Basel)., vol. 8, no. 8, 2017.
[18] A. Iwasaki, R. Medzhitov, Control of adaptive immunity by the innate immune system, Nat. Immunol. 16 (4) (2015) 343–353.
[19] R. Medzhitov, C. Janeway, Innate immune recognition: mechanisms and pathways, Immunol. Rev. 173 (2000) 89–97.
[20] M. Dajon, K. Iribarren, I. Cremer, Toll-like receptor stimulation in cancer: a pro- and anti-tumor double-edged sword, Immunobiology 222 (1) (2017) 89–100.
[21] C. Susan, Long noncoding RNA: novel links between gene expression and innate immunity, Virus Res. 212 (2015) 137–145.
[22] R.C. Scanzello, Role of low-grade inflammation in osteoarthritis, Curr. Opin. Rheumatol. 29 (1) (2017) 79–85.
[23] S. Miranda-Hernandez, A.G. Baxter, Role of toll-like receptors in multiple sclerosis, Am. J. Clin. Exp. Immunol. 2 (1) (2013) 75–93.
[24] R. Wunderlich et al., "Interconnection between DNA damage, senescence, inflammation, and cancer.," Front. Biosci. (Landmark Ed.), vol. 22, pp. 348–369, 2017.
[25] A.T. Satpathy, H.Y. Chang, Long noncoding RNA in hematopoiesis and immunity, Immunity 42 (5) (May 2015) 792–804.
[26] P.J. Batista, H.Y. Chang, Long noncoding RNAs: cellular address codes in development and disease, Cell 152 (6) (2013) 1298–1307.
[27] A. Fatica, I. Bozzoni, Long non-coding RNAs: new players in cell differentiation and development, Nat. Rev. Genet. 15 (1) (2014) 7–21.
[28] J. Liu, H. Wang, N.H. Chua, Long noncoding RNA transcriptome of plants, Plant Biotechnol. J. 13 (3) (2015) 319–328.
[29] A. Rafiee, F. Riazi-Rad, M. Havaskary, F. Nuri, Long noncoding RNAs: regulation, function and cancer, Biotechnol. Genet. Eng. Rev. 34 (2) (2018) 153–180.
[30] N. Jariwala, D. Sarkar, Emerging role of lncRNA in cancer: a potential avenue in molecular medicine, Ann. Transl. Med. 4 (15) (2016) 286.
[31] J. J. Chan and Y. Tay, "Noncoding RNA: RNA regulatory networks in cancer," Int. J. Mol. Sci., vol. 19, no. 5, 2018.
[32] T. Tian et al., "The impact of lncRNA dysregulation on clinicopathology and survival of breast cancer: a systematic review and meta-analysis," Mol. Ther. - Nucleic Acids, vol. 12, no. September, pp. 359–369, 2018.
[33] Y. Mitobe, K. ichi Takayama, K. Horie-Inoue, and S. Inoue, "Prostate cancer-associated lncRNAs," Cancer Lett., vol. 418, pp. 159–166, 2018.
[34] M.J. Jiménez-Dalmaroni, M.E. Gerswhin, I.E. Adamopoulos, The critical role of toll-like receptors - from microbial recognition to autoimmunity: a comprehensive review, Autoimmun. Rev. 15 (1) (2016) 1–8.
[35] S. W. Brubaker, K. S. Bonham, I. Zanoni, and J. C. Kagan, Innate immune pattern recognition: a cell biological perspective, vol. 33, no. 1. 2015.
[36] R. Medzhitov, T. Horng, Transcriptional control of the inflammatory response, Nat. Rev. Immunol. 9 (10) (2009) 692–703.
[37] S.T. Smale, Transcriptional regulation in the innate immune system, Curr. Opin. Immunol. 24 (1) (2012) 51–57.
[38] K. Takeda, S. Akira, Toll-like receptors in innate immunity, Int. Immunol. 17 (1) (Jan. 2005) 1–14.
[39] M. Fukata, A.S. Vamadevan, M.T. Abreu, Toll-like receptors (TLRs) and Nod-like receptors (NLRs) in inflammatory disorders, Semin. Immunol. 21 (4) (Aug. 2009) 242–253.
[40] L. Nie, S.-Y. Cai, J.-Z. Shao, J. Chen, Toll-like receptors, associated biological roles, and signaling networks in non-mammals, Front. Immunol. 9 (2018) 1523.
[41] D.M. Underhill, A. Ozinsky, Toll-like receptors: key mediators of microbe detection, Curr. Opin. Immunol. 14 (1) (2002) 103–110.
[42] S. Carpenter, et al., A long noncoding RNA mediates both activation and repression of immune response genes, Science 341 (6147) (Aug. 2013) 789–792.
[43] G. Hu, et al., LincRNA-Cox2 promotes late inflammatory gene transcription in macrophages through modulating SWI/SNF-mediated chromatin remodeling, J. Immunol. 196 (6) (2016) 2799–2808.
[44] Q. Tong, et al., LincRNA-Cox2 modulates TNF-α-induced transcription of Il12b

gene in intestinal epithelial cells through regulation of Mi-2/NuRD-mediated epigenetic histone modifications, FASEB J. 30 (3) (2016) 1187–1197.
[45] S. Covarrubias, et al., CRISPR/Cas-based screening of long non-coding RNAs (lncRNAs) in macrophages with an NF-κB reporter, J. Biol. Chem. 292 (51) (2017) 20911–20920.
[46] R. Elling et al., "Genetic models reveal cis and trans immune-regulatory activities for lincRNA-Cox2," Cell Rep., vol. 25, no. 6, p. 1511–1524.e6, 2018.
[47] Hadjicharalambous and Lindsay, "Long non-coding RNAs and the innate immune response," Non-Coding RNA, vol. 5, no. 2, p. 34, 2019.
[48] S. Carpenter and K. A. Fitzgerald, "Cytokines and long noncoding RNAs," Cold Spring Harb. Perspect. Biol., vol. 10, no. 6, 2018.
[49] L. Ma, V.B. Bajic, Z. Zhang, On the classification of long non-coding RNAs, RNA Biol. 10 (6) (2013) 924–933.
[50] G.St. Laurent, C. Wahlestedt, P. Kapranov, The landscape of long noncoding RNA classification, Trends Genet. 31 (5) (2015) 239–251.
[51] V.E. Villegas, P.G. Zaphiropoulos, Neighboring gene regulation by antisense long non-coding RNAs, Int. J. Mol. Sci. 16 (2) (2015) 3251–3266.
[52] J.M. Engreitz, et al., Local regulation of gene expression by lncRNA promoters, transcription and splicing, Nature 539 (7629) (2016) 452–455.
[53] Q. Zhang et al., "The long noncoding RNA ROCKI regulates inflammatory gene expression.," EMBO J., vol. 38, no. 8, Apr. 2019.
[54] J. Chan, et al., Cutting edge: a natural antisense transcript, AS-IL1α, controls inducible transcription of the proinflammatory cytokine IL-1α, J. Immunol. 195 (4) (2015) 1359–1363.
[55] H. Cui, et al., The human long noncoding RNA lnc-IL7R regulates the inflammatory response, Eur. J. Immunol. 44 (7) (2014) 2085–2095.
[56] N.E. Ilott, et al., Corrigendum: long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human monocytes, Nat. Commun. 6 (Apr. 2015) 6814.
[57] M. Guttman, et al., Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals, Nature 458 (7235) (2009) 223–227.
[58] J.L. Rinn, H.Y. Chang, Genome regulation by long noncoding RNAs, Annu. Rev. Biochem. 81 (2012) 145–166.
[59] P. Yan, S. Luo, J.Y. Lu, X. Shen, Cis- and trans-acting lncRNAs in pluripotency and reprogramming, Curr. Opin. Genet. Dev. 46 (2017) 170–178.
[60] M.K. Atianand, et al., A long noncoding RNA lincRNA-EPS acts as a transcriptional brake to restrain inflammation, Cell 165 (7) (2016) 1672–1685.
[61] A. Castellanos-rubio, et al., A long noncoding RNA associated with susceptibility to celiac disease, Science (80-.). 352 (6281) (2016) 91–96.
[62] M.K. Atianand, D.R. Caffrey, K.A. Fitzgerald, Immunobiology of long noncoding RNAs, Annu. Rev. Immunol. 35 (1) (2017) 177–198.
[63] A. Frishberg, et al., Cell composition analysis of bulk genomics using single-cell data, Nat. Methods 16 (4) (2019) 327–332.
[64] Y. Lavin et al., "Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses.," Cell, vol. 169, no. 4, p. 750–765.e17, 2017.
[65] S. A. MacParland et al., "Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations.," Nat. Commun., vol. 9, no. 1, p. 4383, 2018.
[66] F. Wimmers et al., "Single-cell analysis reveals that stochasticity and paracrine signaling control interferon-alpha production by plasmacytoid dendritic cells.," Nat. Commun., vol. 9, no. 1, p. 3317, 2018.
[67] T.M. Gierahn, et al., Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput, Nat. Methods 14 (4) (Apr. 2017) 395–398.
[68] X. Peng et al., "Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling.," MBio, vol. 1, no. 5, Oct. 2010.
[69] Z. Altboum, et al., Digital cell quantification identifies global immune cell dynamics during influenza infection, Mol. Syst. Biol. 10 (2014) 720.
[70] D. Lara-Astiaso, et al., Chromatin state dynamics during blood formation HHS Public Access, Science (80-.). 345 (6199) (2014) 943–949.
[71] Y. Lavin, et al., Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment, Cell 159 (6) (Dec. 2014) 1312–1326.
[72] A.-J. Tong, et al., A Stringent Systems Approach Uncovers Gene-Specific Mechanisms Regulating Inflammation, Cell 165 (1) (Mar. 2016) 165–179.
[73] V. M. Link et al., "Analysis of genetically diverse macrophages reveals local and domain-wide mechanisms that control transcription factor binding and function.," Cell, vol. 173, no. 7, p. 1796–1809.e17, 2018.
[74] D.M. Bhatt, et al., Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions, Cell 150 (2) (Jul. 2012) 279–290.
[75] S. Roy et al., "Transcriptional landscape of *Mycobacterium tuberculosis* infection in macrophages.," Sci. Rep., vol. 8, no. 1, p. 6758, Apr. 2018.
[76] A.-J. Tong, et al., A stringent systems approach uncovers gene-specific mechanisms regulating inflammation, Cell 165 (1) (Mar. 2016) 165–179.
[77] D. Z. Eichenfield et al., "Tissue damage drives co-localization of NF-κB, Smad3, and Nrf2 to direct Rev-erb sensitive wound repair in mouse macrophages.," Elife, vol. 5, 2016.
[78] A.A. Pai, et al., Widespread shortening of 3′ untranslated regions and increased exon inclusion are evolutionarily conserved features of innate immune responses to infection, PLoS Genet. 12 (9) (2016) 1–24.
[79] P. Wang, et al., The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation, Science 344 (6181) (Apr. 2014) 310–313.
[80] S. Saeed et al., "Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity.," Science, vol. 345, no. 6204, p. 1251086, Sep. 2014.
[81] B. Novakovic et al., "β-Glucan reverses the epigenetic state of LPS-induced immunological tolerance.," Cell, vol. 167, no. 5, p. 1354–1368.e14, 2016.
[82] A. Dix, et al., Specific and novel microRNAs are regulated as response to fungal infection in human dendritic cells, Front. Microbiol. 8 (2017) 270.
[83] K. Wu et al., "An interferon-related signature in the transcriptional core response of human macrophages to *Mycobacterium tuberculosis* infection.," PLoS One, vol. 7, no. 6, p. e38367, 2012.
[84] Z. Li, et al., The long noncoding RNA THRIL regulates TNFα expression through its interaction with hnRNPL, Proc. Natl. Acad. Sci. U. S. A. 111 (3) (Jan. 2014) 1002–1007.
[85] B.T. Roux, J.A. Heward, L.E. Donnelly, S.W. Jones, M.A. Lindsay, Catalog of differentially expressed long non-coding RNA following activation of human and mouse innate immune response, Front. Immunol. 8 (2017) 1038.
[86] N. Zhang et al., "Whole transcriptome analysis reveals differential gene expression profile reflecting macrophage polarization in response to influenza A H5N1 virus infection.," BMC Med. Genomics, vol. 11, no. 1, p. 20, 2018.
[87] M.D. Schultz, et al., Human body epigenome maps reveal noncanonical DNA methylation variation, Nature 523 (7559) (Jul. 2015) 212–216.
[88] M.J. Ciancanelli, et al., Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency, Science 348 (6233) (Apr. 2015) 448–453.
[89] S. Hong, et al., Longitudinal profiling of human blood transcriptome in healthy and lupus pregnancy, J. Exp. Med. 216 (5) (May 2019) 1154–1169.
[90] R.E. Thurman, et al., The accessible chromatin landscape of the human genome, Nature 489 (7414) (Sep. 2012) 75–82.
[91] B. Zhou, et al., EVLncRNAs: a manually curated database for long non-coding RNAs validated by low-throughput experiments, Nucleic Acids Res. 46 (D1) (2018) D100–D105.
[92] S. Fang, et al., NONCODEV5: a comprehensive annotation database for long non-coding RNAs, Nucleic Acids Res. 46 (D1) (2018) D308–D314.
[93] P.-J. Volders, et al., LNCipedia 5: towards a reference set of human long non-coding RNAs, Nucleic Acids Res. 47 (D1) (Jan. 2019) D135–D139.
[94] S. Djebali, et al., Landscape of transcription in human cells, Nature 489 (7414) (2012) 101–108.
[95] M.N. Cabili, et al., Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses, Genes Dev. 25 (18) (Sep. 2011) 1915–1927.
[96] S. J. Liu et al., "CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells.," Science, vol. 355, no. 6320, 2017.
[97] F. Aguet, et al., Genetic effects on gene expression across human tissues, Nature 550 (7675) (2017) 204–213.
[98] M. Goldman, B. Craft, A. Kamath, A. Brooks, J. Zhu, and D. Haussler, "The UCSC Xena Platform for cancer genomics data visualization and interpretation," bioRxiv, no. Schroeder 2015, p. 326470, 2018.
[99] X. Han et al., "Mapping the Mouse Cell Atlas by Microwell-Seq.," Cell, vol. 172, no. 5, p. 1091–1107.e17, 2018.
[100] N. Schaum, et al., Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris, Nature 562 (7727) (2018) 367–372.
[101] C.A. Davis, et al., The Encyclopedia of DNA elements (ENCODE): data portal update, Nucleic Acids Res. 46 (D1) (2018) D794–D801.
[102] M. de Hoon, J.W. Shin, P. Carninci, Paradigm shifts in genomics through the FANTOM projects, Mamm. Genome 26 (9–10) (2015) 391–402.
[103] W. S. Bush and J. H. Moore, "Chapter 11: genome-wide association studies.," PLoS Comput. Biol., vol. 8, no. 12, p. e1002822, 2012.
[104] P.M. Visscher, et al., 10 years of GWAS discovery: biology, function, and translation, Am. J. Hum. Genet. 101 (1) (2017) 5–22.
[105] J.S. Witte, P.M. Visscher, N.R. Wray, The contribution of genetic variants to disease depends on the ruler, Nat. Rev. Genet. 15 (11) (2014) 765–776.
[106] V. Kumar et al., "Human disease-associated genetic variation impacts large intergenic non-coding RNA expression," PLoS Genet., vol. 9, no. 1, 2013.
[107] C. Bycroft, et al., The UK Biobank resource with deep phenotyping and genomic data, Nature 562 (7726) (2018) 203–209.
[108] J.S. Witte, Genome-wide association studies and beyond, Annu. Rev. Public Health 31 (2010) 9–20.
[109] Y.R. Miao, W. Liu, Q. Zhang, A.Y. Guo, LncRNASNP2: an updated database of functional SNPs and mutations in human and mouse lncRNAs, Nucleic Acids Res. 46 (D1) (2018) D276–D280.
[110] C. Ren, G. An, C. Zhao, Z. Ouyang, X. Bo, W. Shu, Lnc2Catlas: an atlas of long noncoding RNAs associated with risk of cancers, Sci. Rep. 8 (1) (2018) 1–8.
[111] D. D. Zhang et al., "Long noncoding RNA LINC00305 promotes inflammation by activating the AHRR-NF-k B pathway in human monocytes," Sci. Rep., vol. 7, no. April, pp. 1–12, 2017.
[112] A. Castellanos-Rubio, S. Ghosh, Disease-associated SNPs in inflammation-related lncRNAs, Front. Immunol. 10 (2019) 420.
[113] M.F. Lin, I. Jungreis, M. Kellis, PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions, Bioinformatics 27 (13) (2011) 275–282.
[114] P. Johnsson, L. Lipovich, D. Grandér, K.V. Morris, Evolutionary conservation of long non-coding RNAs; sequence, structure, function, Biochim. Biophys. Acta - Gen. Subj. 1840 (3) (2014) 1063–1071.
[115] J.J. Quinn, H.Y. Chang, Unique features of long non-coding RNA biogenesis and function, Nat. Rev. Genet. 17 (1) (2016) 47–62.
[116] T.C. Roberts, K.V. Morris, M.S. Weinberg, Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs, Epigenetics 9 (1) (2014) 13–20.
[117] A. Nitsche, P.F. Stadler, Evolutionary clues in lncRNAs, Wiley Interdiscip. Rev. RNA 8 (1) (2017) 14–17.
[118] M.T.Y. Lam, W. Li, M.G. Rosenfeld, C.K. Glass, Enhancer RNAs and regulated

transcriptional programs, Trends Biochem. Sci. 39 (4) (Apr. 2014) 170–182.
[119] B. Font-Cunill, L. Arnes, J. Ferrer, L. Sussel, and A. Beucher, “Long non-coding RNAs as local regulators of pancreatic islet transcription factor genes,” Front. Genet., vol. 9, no. November, pp. 1–9, 2018.
[120] J.H. Cheng, D.Z.C. Pan, Z.T.Y. Tsai, H.K. Tsai, Genome-wide analysis of enhancer RNA in gene regulation across 12 mouse tissues, Sci. Rep. 5 (2015) 1–9.
[121] G. K. Rai, N. P. Rai, S. Rathaur, S. Kumar, and M. Singh, “Expression of rd29A:: AtDREB1A/CBF3 in tomato alleviates drought-induced oxidative stress by regulating key enzymatic and non-enzymatic antioxidants,” Plant Physiol. Biochem., vol. 69, pp. 90–100, 2013.
[122] C.A. Melo, et al., ERNAs are required for p53-dependent enhancer activity and gene transcription, Mol. Cell 49 (3) (2013) 524–535.
[123] R.C.A. Tavares, A.M. Pyle, S. Somarowthu, Phylogenetic analysis with improved parameters reveals conservation in lncRNA structures, J. Mol. Biol. 431 (8) (2019) 1592–1603.
[124] X.C. Quek, et al., lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs, Nucleic Acids Res. 43 (D1) (2015) D168–D173.
[125] D. Karolchik, S.A. Hinrichs, W.J. Kent, The UCSC genome browser, Curr. Protoc. Bioinforma. (2009) 1–34.
[126] S. Ma, et al., A long noncoding RNA, lincRNA-Tnfaip3, acts as a coregulator of NF-κB to modulate inflammatory gene transcription in mouse macrophages, FASEB J. 31 (3) (2017) 1215–1225.
[127] M. Krawczyk, B.M. Emerson, p50-associated COX-2 extragenic RNA (PACER) activates COX-2 gene expression by occluding repressive NF-κB complexes, Elife 3 (Apr. 2014) e01776.
[128] G.E. Crawford, et al., Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS), Genome Res. 16 (1) (2006) 123–131.
[129] J. Buenrostro, B. Wu, H. Chang, W. Greenleaf, ATAC-seq: a method for assaying chromatin accessibility genome-wide, Curr. Protoc. Mol. Biol. 2015 (2016) 1–10.
[130] H. He, Z. Hu, H. Xiao, F. Zhou, B. Yang, The tale of histone modifications and its role in multiple sclerosis, Hum. Genomics 12 (1) (2018) 1–12.
[131] Y. Araki, T. Mimura, The histone modification code in the pathogenesis of autoimmune diseases, Mediat. Inflamm. 2017 (2017) 1–12.
[132] B. Alaskhar Alhamwe et al., “Histone modifications and their role in epigenetics of atopy and allergic diseases,” Allergy, Asthma Clin. Immunol., vol. 14, no. 1, 2018.
[133] A.J. Bannister, T. Kouzarides, Regulation of chromatin by histone modifications, Cell Res. 21 (3) (2011) 381–395.
[134] V.G. Allfrey, R. Faulkner, A.E. Mirsky, Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis, Proc. Natl. Acad. Sci. U. S. A. 51 (May 1964) 786–794.
[135] E. Platanitis and T. Decker, “Regulatory networks involving STATs, IRFs, and NFκB in inflammation,” Front. Immunol., vol. 9, no. NOV, pp. 1–16, 2018.
[136] C. G. Palii et al., “Single-cell proteomics reveal that quantitative changes in co-expressed lineage-specific transcription factors determine cell fate,” Cell Stem Cell, vol. 24, no. 5, p. 812–820.e5, 2019.
[137] S. Carpenter, E.P. Ricci, B.C. Mercier, M.J. Moore, K.A. Fitzgerald, Post-transcriptional regulation of gene expression in innate immunity, Nat. Rev. Immunol. 14 (6) (2014) 361–376.
[138] T. H. Chang, H. Y. Huang, J. B. K. Hsu, S. L. Weng, J. T. Horng, and H. Da Huang, “An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs,” BMC Bioinformatics, vol. 14, no. Suppl 2, 2013.
[139] V. Boeva, “Analysis of genomic sequence motifs for deciphering transcription factor binding and transcriptional regulation in eukaryotic cells,” Front. Genet., vol. 7, no. FEB, 2016.
[140] S. Heinz, et al., Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities, Mol. Cell 38 (4) (2010) 576–589.
[141] M.C. Frith, et al., Discrimination of non-protein-coding transcripts from protein-coding mRNA, RNA Biol. 3 (1) (2006) 40–48.
[142] M.I. Galindo, J.I. Pueyo, S. Fouix, S.A. Bishop, J.P. Couso, Peptides encoded by short ORFs control development and define a new eukaryotic gene family, PLoS Biol. 5 (5) (2007) 1052–1062.
[143] H. Cheng, W.S. Chan, Z. Li, D. Wang, S. Liu, Y. Zhou, Small open reading frames: current prediction techniques and future prospect, Curr. Protein Perspecitive Sci. 12 (6) (2011) 503–507.
[144] S.A. Slavoff, et al., Peptidomic discovery of short open reading frame-encoded peptides in human cells, Nat. Chem. Biol. 9 (1) (2013) 59–64.
[145] N.T. Ingolia, et al., Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes, Cell Rep. 8 (5) (2014) 1365–1379.
[146] N. T. Ingolia, S. Ghaemmaghami, J. R. S. Newman, and J. Weissman, “Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling,” Science (80-.)., vol. 1168978, no. April, pp. 218–324, 2009.
[147] M. Guttman, P. Russell, N.T. Ingolia, J.S. Weissman, E.S. Lander, Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins, Cell 154 (1) (2013) 240–251.
[148] J. Carlevaro-Fita, A. Rahim, R. Guigó, L.A. Vardy, R. Johnson, Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells, RNA 22 (6) (2016) 867–882.
[149] G.-L. Chew, A. Pauli, J.L. Rinn, A. Regev, A.F. Schier, E. Valen, Ribosome profiling reveals resemblance between long non-coding RNAs and 5′ leaders of coding RNAs, Development 140 (13) (2013) 2828–2834.
[150] R. Jackson, et al., The translation of non-canonical open reading frames controls mucosal immunity, Nature 564 (7736) (2018) 434–438.
[151] L. Ma, et al., Lncbook: a curated knowledgebase of human long non-coding rnas, Nucleic Acids Res. 47 (D1) (2019) D128–D134.
[152] M. Boettcher, M.T. McManus, Choosing the right tool for the job: RNAi, TALEN, or CRISPR, Mol. Cell 58 (4) (May 2015) 575–585.
[153] J. Kurreck, RNA interference: from basic research to therapeutic applications, Angew. Chemie - Int. Ed. 48 (8) (2009) 1378–1398.
[154] L.S. Lambeth, C.A. Smith, Short hairpin RNA-mediated gene silencing, Methods Mol. Biol. 942 (2013) 205–232.
[155] K.A. Lennox, M.A. Behlke, Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides, Nucleic Acids Res. 44 (2) (2016) 863–877.
[156] X. Zong, et al., Knockdown of nuclear-retained long noncoding RNAs using modified DNA antisense oligonucleotides, Methods Mol. Biol. 1262 (2015) 321–331.
[157] G.F. Deleavey, M.J. Damha, Designing chemically modified oligonucleotides for targeted gene silencing, Chem. Biol. 19 (8) (2012) 937–954.
[158] S.M. Cerritelli, R.J. Crouch, Ribonuclease H: the enzymes in eukaryotes, FEBS J. 276 (6) (2009) 1494–1505.
[159] M.L. Hvam, et al., Fatty acid-modified gapmer antisense oligonucleotide and serum albumin constructs for pharmacokinetic modulation, Mol. Ther. 25 (7) (2017) 1710–1717.
[160] D.E. Greydanus, H.D. Pratt, M. Pryson, Angelman syndrome, Heal. Care People with Intellect. Dev. Disabil. Across Lifesp. (2016) 815–819.
[161] L. Meng, A.J. Ward, S. Chun, C.F. Bennett, A.L. Beaudet, F. Rigo, Towards a therapy for Angelman syndrome by targeting a long non-coding RNA, Nature 518 (7539) (2015) 409–412.
[162] J. Stanisławska, W.L. Olszewski, RNA interference—significance and applications, Arch. Immunol. Ther. Exp. 53 (1) (2005) 39–46.
[163] S.T. Crooke, Molecular mechanisms of antisense oligonucleotides, Nucleic Acid Ther. 27 (2) (2017) 70–77.
[164] S.E. Castel, R.A. Martienssen, RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond, Nat. Rev. Genet. 14 (2) (2013) 100–112.
[165] N.J. Proudfoot, Transcriptional termination in mammals: stopping the RNA polymerase juggernaut, Science (80-.). 352 (6291) (2016) 1291–1300.
[166] S.E. Castel, et al., Dicer promotes transcription termination at sites of replication stress to maintain genome stability, Cell 159 (3) (2014) 572–583.
[167] M.L. Hochstrasser, J.A. Doudna, Cutting it close: CRISPR-associated endoribonuclease structure and function, Trends Biochem. Sci. 40 (1) (2015) 58–66.
[168] S.H. Sternberg, J.A. Doudna, Expanding the Biologist's toolkit with CRISPR-Cas9, Mol. Cell 58 (4) (2015) 568–574.
[169] T. T. Ho et al., “Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines,” Nucleic Acids Res., vol. 43, no. 3, p. e17, 2015.
[170] S. Zhu, et al., Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library, Nat. Biotechnol. 34 (12) (2016) 1279–1286.
[171] A. R. Bassett et al., “Considerations when investigating lncRNA function in vivo,” Elife, vol. 3, no. August2014, pp. 1–14, 2014.
[172] Y. Liu, et al., Genome-wide screening for functional long noncoding RNAs in human cells by Cas9 targeting of splice sites, Nat. Biotechnol. 36 (12) (2018) 1203–1210.
[173] S. Lalonde, et al., Frameshift indels introduced by genome editing can lead to in-frame exon skipping, PLoS One 12 (6) (2017) 1–13.
[174] L.A. Gilbert, et al., Genome-scale CRISPR-mediated control of gene repression and activation, Cell 159 (3) (2014) 647–661.
[175] M.H. Larson, L.A. Gilbert, X. Wang, W.A. Lim, J.S. Weissman, L.S. Qi, CRISPR interference (CRISPRi) for sequence-specific control of gene expression, Nat. Protoc. 8 (11) (2013) 2180–2196.
[176] Y. Ying, et al., The Krüppel-associated box repressor domain induces reversible and irreversible regulation of endogenous mouse genes by mediating different chromatin states, Nucleic Acids Res. 43 (3) (2015) 1549–1561.
[177] G. Prelich, Gene overexpression: uses, mechanisms, and interpretation, Genetics 190 (3) (2012) 841–854.
[178] A.W. Cheng, et al., Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system, Cell Res. 23 (10) (2013) 1163–1171.
[179] L. A. Gilbert et al., “CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes,” Cell, vol. 154, no. 2, p. 442, 2013.
[180] J. Joung, et al., Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood, Nature 548 (7667) (2017) 343–346.
[181] L.A. Goff, J.L. Rinn, Linking RNA biology to lncRNAs, Genome Res. 25 (10) (2015) 1456–1465.
[182] S. W. Cho et al., “Promoter of lncRNA gene PVT1 is a tumor-suppressor DNA boundary element,” Cell, vol. 173, no. 6, p. 1398–1412.e22, 2018.
[183] V.R. Paralkar, et al., Unlinking an lncRNA from its associated cis element, Mol. Cell 62 (1) (2016) 104–110.
[184] M. Sauvageau, et al., Multiple knockout mouse models reveal lincRNAs are required for life and brain development, Elife 2013 (2) (2013) 1–24.
[185] Y. Marahrens, B. Panning, J. Dausman, W. Strauss, R. Jaenisch, Xist-deficient mice are defective in dosage compensation but not spermatogenesis, Genes Dev. 11 (2) (1997) 156–166.
[186] M.A. Ripoche, C. Kress, F. Poirier, L. Dandolo, Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element, Genes Dev. 11 (12) (1997) 1596–1604.
[187] V.H. Meller, B.P. Rattner, The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex, EMBO J. 21 (5) (2002) 1084–1091.
[188] M. Eißmann, et al., Loss of the abundant nuclear non-coding RNA MALAT1 is

compatible with life and development, RNA Biol. 9 (8) (2012) 1076–1087.
[189] L. Li, et al., Targeted disruption of Hotair leads to homeotic transformation and gene derepression, Cell Rep. 5 (1) (2013) 3–12.
[190] P. Grote, et al., The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse, Dev. Cell 24 (2) (2013) 206–214.
[191] A.V. Dharmadhikari, P. Szafranski, V.V. Kalinichenko, P. Stankiewicz, Genomic and epigenetic complexity of the FOXF1 locus in 16q24.1: implications for development and disease, Curr. Genomics 16 (2) (2015) 107–116.
[192] S. H. et al., "Transcription factor IRF8 plays a critical role in the development of murine basophils and mast cells," Blood, vol. 125, no. 2, pp. 358–369, 2015.
[193] A.M. Bond, et al., Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry, Nat. Neurosci. 12 (8) (2009) 1020–1027.
[194] E.G. Berghoff, M.F. Clark, S. Chen, I. Cajigas, D.E. Leib, J.D. Kohtz, Evf2 (Dlx6as) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes, Development 140 (21) (2013) 4407–4416.
[195] S. Carmona, B. Lin, T. Chou, K. Arroyo, S. Sun, LncRNA Jpx induces Xist expression in mice using both trans and cis mechanisms, PLoS Genet. 14 (5) (2018) 1–21.
[196] R. E. Andersen et al., "The long noncoding RNA Pnky is a trans-acting regulator of cortical development in vivo.," Dev. Cell, vol. 49, no. 4, p. 632–642.e7, May 2019.
[197] C. Chu, H.Y. Chang, ChIRP-MS: RNA-directed proteomic discovery, Methods Mol. Biol. 1861 (2018) 37–45.
[198] C. Chu, J. Quinn, and H. Y. Chang, "Chromatin isolation by RNA purification (ChIRP)," J. Vis. Exp., no. 61, pp. 1–6, 2012.
[199] C.A. McHugh, M. Guttman, RAP-MS: a method to identify proteins that interact directly with a specific RNA molecule in cells, Methods Mol. Biol. 1649 (2018) 473–488.
[200] J.M. Engreitz, et al., RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites, Cell 159 (1) (2014) 188–199.
[201] M. Torres et al., "RNA pull-down procedure to identify RNA targets of a long noncoding RNA.," J. Vis. Exp., no. 134, 2018.
[202] S. Li, et al., Zbtb7b engages the long noncoding RNA Blnc1 to drive brown and beige fat development and thermogenesis, Proc. Natl. Acad. Sci. 114 (34) (2017) E7111–E7120.
[203] F. Agostini, A. Zanzoni, P. Klus, D. Marchese, D. Cirillo, G.G. Tartaglia, CatRAPID omics: a web server for large-scale prediction of protein-RNA interactions, Bioinformatics 29 (22) (2013) 2928–2930.
[204] V.J. Henry, A.E. Bandrowski, A.S. Pepin, B.J. Gonzalez, A. Desfeux, OMICtools: an informative directory for multi-omic data analysis, Database (Oxford). 2014 (13) (2014) 1–5.
[205] Q. Fan, et al., The emerging role of exosome-derived non-coding RNAs in cancer biology, Cancer Lett. 414 (2018) 107–115.
[206] J. Carlevaro-Fita, R. Johnson, Global positioning system: understanding long noncoding RNAs through subcellular localization, Mol. Cell 73 (5) (2019) 869–883.
[207] J. Heon Noh, K. Mi Kim, W. McClucky, K. Abdelmohsen, M. Gorospe, Cytoplasmic functions of lncRNAs, Wiley Interdiscip Rev RNA 118 (24) (2018) 6072–6078.
[208] Q. Sun, Q. Hao, K.V. Prasanth, Nuclear long noncoding RNAs: key regulators of gene expression, Trends Genet. 34 (2) (2018) 142–157.
[209] B. Wan, et al., Plasma long noncoding RNA IL-7R as a prognostic biomarker for clinical outcomes in patients with acute respiratory distress syndrome, Clin. Respir. J. 12 (4) (2018) 1607–1614.
[210] I. Legnini, M. Morlando, A. Mangiavacchi, A. Fatica, I. Bozzoni, A feedforward regulatory loop between HuR and the long noncoding RNA linc-MD1 controls early phases of myogenesis, Mol. Cell 53 (3) (2014) 506–514.
[211] M. Cesana, et al., A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA, Cell 147 (2) (2011) 358–369.
[212] Z.D. Su, et al., ILoc-lncRNA: predict the subcellular location of lncRNAs by incorporating octamer composition into general PseKNC, Bioinformatics 34 (24) (2018) 4196–4204.
[213] F. M. Fazal et al., "Atlas of subcellular RNA localization revealed by APEX-seq," bioRxiv, p. 454470, 2018.
[214] P.-J. A. and B. D.L., Co-transcriptional splicing of constitutive and alternative exons, Rna 15 (10) (2009) 1896–1908.
[215] P.G. Sadowski, A.J. Groen, P. Dupree, K.S. Lilley, Sub-cellular localization of membrane proteins, Proteomics 8 (19) (2008) 3991–4011.
[216] C. Lesnik, A. Golani-Armon, Y. Arava, Localized translation near the mitochondrial outer membrane: an update, RNA Biol. 12 (8) (2015) 801–809.
[217] A. M. Femino, F. S. Fay, K. Fogarty, and R. H. Singer, "Visualization of single RNA transcripts in situ stable," Am. Assoc. Adv. Sci., vol. 280, no. 5363, pp. 585–590, 2018.
[218] S. Shah, E. Lubeck, W. Zhou, L. Cai, In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus, Neuron 92 (2) (2016) 342–357.
[219] U. Schnell, F. Dijk, K.A. Sjollema, B.N.G. Giepmans, Immunolabeling artifacts and the need for live-cell imaging, Nat. Methods 9 (2) (2012) 152–158.
[220] I. Ulitsky, Evolution to the rescue: using comparative genomics to understand long non-coding RNAs, Nat. Rev. Genet. 17 (10) (2016) 601–614.
[221] X. Zhang, et al., Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions, Endocrinology 151 (3) (2010) 939–947.
[222] C. Sherpa, J.W. Rausch, S.F.J. Le Grice, Structural characterization of maternally expressed gene 3 RNA reveals conserved motifs and potential sites of interaction with polycomb repressive complex 2, Nucleic Acids Res. 46 (19) (2018) 10432–10447.
[223] D. Sprague, et al., Non-linear sequence similarity between the Xist and Rsx long noncoding RNAs suggests shared functions of tandem repeat domains, RNA, 2019.
[224] L. Shang, W. Xu, S. Ozer, and R. R. Gutell, "Structural constraints identified with covariation analysis in ribosomal RNA," PLoS One, vol. 7, no. 6, 2012.
[225] R. C. A. Tavares, A. M. Pyle, and S. Somarowthu, "Covariation analysis with improved parameters reveals conservation in lncRNA structures The existence of phylogenetic covariation in base-pairing is strong evidence for functional," bioRxiv, pp. 1–27, 2018.
[226] P. Tijerina, S. Mohr, R. Russell, Dms footprinting of structured rnas and rna-protein complexes, Nat. Protoc. 2 (10) (2007) 2608–2623.
[227] P. Tijerina, S. Mohr, R. Russell, Dms footprinting of structured rnas and rna-protein complexes, Nat. Protoc. 2 (10) (2007) 2608–2623.
[228] S. Rouskin, M. Zubradt, S. Washietl, M. Kellis, J.S. Weissman, Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo, Nature 505 (7485) (2014) 701–705.
[229] F. Liu, S. Somarowthu, A.M. Pyle, Visualizing the secondary and tertiary architectural domains of lncRNA RepA, Nat. Chem. Biol. 13 (3) (2017) 282–289.
[230] Z. Xue, et al., A G-rich motif in the lncRNA Braveheart interacts with a zinc-finger transcription factor to specify the cardiovascular lineage, Mol. Cell 64 (1) (2016) 37–50.
[231] R. Fang, W.N. Moss, M. Rutenberg-Schoenberg, M.D. Simon, Probing Xist RNA structure in cells using targeted structure-Seq, PLoS Genet. 11 (12) (2015) 1–29.
[232] N.A. Siegfried, S. Busan, G.M. Rice, J.A.E. Nelson, K.M. Weeks, RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP), Nat. Methods 11 (9) (2014) 959–965.
[233] K. E. Watters, T. R. Abbott, and J. B. Lucks, "Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq," Nucleic Acids Res., vol. 44, no. 2, p. e12, 2016.
[234] R.A. Flynn, Q.C. Zhang, R.C. Spitale, B. Lee, M.R. Mumbach, H.Y. Chang, Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE, Nat. Protoc. 11 (2) (2016) 273–290.
[235] K.M. Weeks, D.M. Mauger, Exploring RNA structural codes with SHAPE chemistry, Acc. Chem. Res. 44 (12) (2011) 1280–1291.
[236] K.A. Wilkinson, E.J. Merino, K.M. Weeks, Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution, Nat. Protoc. 1 (3) (2006) 1610–1616.
[237] M.J. Smola, et al., SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells, Proc. Natl. Acad. Sci. 113 (37) (2016) 10322–10327.
[238] B. Lee, R.A. Flynn, A. Kadina, J.K. Guo, E.T. Kool, H.Y. Chang, Comparison of SHAPE reagents for mapping RNA structures inside living cells, RNA 23 (2) (2017) 169–174.
[239] I.A. Ilik, et al., Tandem stem-loops in roX RNAs act together to mediate X chromosome dosage compensation in Drosophila, Mol. Cell 51 (2) (2013) 156–173.
[240] Z. Lu, et al., RNA duplex map in living cells reveals higher-order transcriptome structure, Cell 165 (5) (2016) 1267–1279.
[241] S. Stamm, et al., Function of alternative splicing, Gene 344 (2005) 1–20.
[242] J. Wang, Y. Chen, K. Xu, Y. Mo, and Y. Zhou, "Comprehensive network analysis reveals alternative splicing-related lncRNAs in hepatocellular carcinoma," bioRxiv, pp. 1–28, 2019.
[243] E. A. Kiegle, A. Garden, E. Lacchini, and M. M. Kater, "A genomic view of alternative splicing of long non-coding RNAs during rice seed development reveals extensive splicing and lncRNA gene families," Front. Plant Sci., vol. 9, no. February, pp. 1–12, 2018.
[244] C. Ziegler and M. Kretz, "The more the merrier—complexity in long non-coding RNA loci," Front. Endocrinol. (Lausanne)., vol. 8, no. April, pp. 1–6, 2017.
[245] Y. Katz, E.T. Wang, E.M. Airoldi, C.B. Burge, Analysis and design of RNA sequencing experiments for identifying isoform regulation, Nat. Methods 7 (12) (2010) 1009–1015.
[246] A.N. Brooks, et al., Conservation of an RNA regulatory map between Drosophila and mammals, Genome Res. 21 (2) (Feb. 2011) 193–202.
[247] C.J. Green, M.R. Gazzara, Y. Barash, MAJIQ-SPEL: web-tool to interrogate classical and complex splicing variations from RNA-Seq data, Bioinformatics 34 (2) (2018) 300–302.
[248] M. Nowicka and M. D. Robinson, "DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics," F1000Research, vol. 5, no. May, p. 1356, 2016.
[249] J.M. Hibberd, R. Patro, S. Kelly, C. Boursnell, R. Smith-Unna, TransRate: reference-free quality assessment of de novo transcriptome assemblies, Genome Res. 26 (8) (2016) 1134–1144.
[250] D. Grün, et al., De novo prediction of stem cell identity using single-cell transcriptome data, Cell Stem Cell 19 (2) (2016) 266–277.
[251] Z. Chang, et al., Bridger: a new framework for de novo transcriptome assembly using RNA-seq data, Genome Biol. 16 (1) (2015) 1–10.
[252] J. L. Weirather et al., "Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis," F1000Research, vol. 6, no. May, p. 100, 2017.
[253] K. Loit et al., "Relative performances of Oxford Nanopore MinION vs. Pacific Biosciences Sequel third-generation sequencing platforms in identification of agricultural and forest pathogens.," bioRxiv, 2019.
[254] A. Rhoads, K.F. Au, PacBio sequencing and its applications, Genomics, Proteomics Bioinforma. 13 (5) (2015) 278–289.
[255] M. Ferrarini et al., "An evaluation of the PacBio RS platform for sequencing and de novo assembly of a chloroplast genome," BMC Genomics, vol. 14, no. 1, 2013.
[256] W. De Coster, S. D'Hert, D.T. Schultz, M. Cruts, C. Van Broeckhoven, NanoPack: visualizing and processing long-read sequencing data, Bioinformatics 34 (15) (2018) 2666–2669.

[257] N. O. Technology, "RNA and gene expression analysis using direct RNA and cDNA sequencing," 2019. [Online]. Available: https://nanoporetech.com/applications/rna-sequencing. [Accessed: 26-May-2019].
[258] R. Volden, et al., Improving nanopore read accuracy with the R2C2 method enables the sequencing of highly multiplexed full-length single-cell cDNA, Proc. Natl. Acad. Sci. U. S. A. 115 (39) (2018) 9726–9731.
[259] Z. Boldogkői, N. Moldován, A. Szűcs, D. Tombácz, Transcriptome-wide analysis of a baculovirus using nanopore sequencing, Sci. data 5 (2018) 180276.
[260] K. C. Maier, S. Gressel, P. Cramer, and B. Schwalb, "Native molecule sequencing by nano-ID reveals synthesis and stability of RNA isoforms," bioRxiv, 2019.
[261] Q. Xu et al., "Transcriptome profiling using single-molecule direct RNA sequencing approach for in-depth understanding of genes in secondary metabolism pathways of *Camellia sinensis*," Front. Plant Sci., vol. 8, no. July, pp. 1–11, 2017.
[262] Y. Lin, B.F. Schmidt, M.P. Bruchez, C.J. McManus, Structural analyses of NEAT1 lncRNAs suggest long-range RNA interactions that may contribute to paraspeckle architecture, Nucleic Acids Res. 46 (7) (2018) 3742–3752.
[263] J.H. Yuan, et al., The MBNL3 splicing factor promotes hepatocellular carcinoma by increasing PXN expression through the alternative splicing of lncRNA-PXN-AS1, Nat. Cell Biol. 19 (7) (2017) 820–832.
[264] S. Ke, et al., A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation, Genes Dev. 29 (19) (Oct. 2015) 2037–2053.
[265] Q. Dai, C. He, T. Pan, G. Zheng, N. Liu, M. Parisien, Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA, RNA 19 (12) (2013) 1848–1856.
[266] R. Winkler, et al., m6A modification controls the innate immune response to infection by targeting type I interferons, Nat. Immunol. 20 (2) (2019) 173–182.
[267] K. I. Zhou et al., "N(6)-methyladenosine modification in a long noncoding RNA hairpin predisposes its conformation to protein binding.," J. Mol. Biol., vol. 428, no. 5 Pt A, pp. 822–833, Feb. 2016.
[268] H. Coker, G. Wei, N. Brockdorff, m6A modification of non-coding RNA and the control of mammalian gene expression, Biochim. Biophys. Acta - Gene Regul. Mech. 1862 (3) (2019) 310–318.
[269] D. Dominissini, et al., Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq, Nature 485 (7397) (2012) 201–206.
[270] K.D. Meyer, Y. Saletore, P. Zumbo, O. Elemento, C.E. Mason, S.R. Jaffrey, Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons, Cell 149 (7) (2012) 1635–1646.
[271] Y. Wu, et al., M 6 A-induced lncRNA RP11 triggers the dissemination of colorectal cancer cells via upregulation of Zeb1, Mol. Cancer 18 (1) (2019) 1–16.
[272] P. Boccaletto, et al., MODOMICS: a database of RNA modification pathways. 2017 update, Nucleic Acids Res. 46 (D1) (2018) D303–D307.
[273] M. Helm, Y. Motorin, Detecting RNA modifications in the epitranscriptome: predict and validate, Nat. Rev. Genet. 18 (5) (2017) 275–291.
[274] H. Liu et al., "Accurate detection of m6A RNA modifications in native RNA sequences," bioRxiv, p. 525741, 2019.
[275] C.J. Brown, et al., A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome, Nature 349 (6304) (Jan. 1991) 38–44.
[276] C.I. Brannan, E.C. Dees, R.S. Ingram, S.M. Tilghman, The product of the H19 gene may function as an RNA, Mol. Cell. Biol. 10 (1) (Jan. 1990) 28–36.