Genomics Proteomics Bioinformatics 14 (2016) 42–54

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com

REVIEW

Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer

Yiwen Fang [1,a], **Melissa J. Fullwood** [1,2,3,4,*,b]

[1] *Cancer Science Institute of Singapore, National University of Singapore, Singapore 117599, Singapore*
[2] *School of Biological Sciences, Nanyang Technological University, Singapore 637551, Singapore*
[3] *Institute of Molecular and Cell Biology, Agency for Science, Technology and Research (A*STAR), Singapore 138673, Singapore*
[4] *Yale-NUS Liberal Arts College, Singapore 138527, Singapore*

Received 10 July 2015; revised 31 August 2015; accepted 17 September 2015
Available online 12 February 2016

Handled by Xiangdong Fang

KEYWORDS

lncRNAs;
Chromatin;
Transcription regulation;
Antisense oligonucleotides;
Cancer

Abstract Long non-coding RNAs (**lncRNAs**) play important roles in **cancer**. They are involved in **chromatin** remodeling, as well as transcriptional and post-transcriptional regulation, through a variety of chromatin-based mechanisms and via cross-talk with other RNA species. **lncRNAs** can function as decoys, scaffolds, and enhancer RNAs. This review summarizes the characteristics of **lncRNAs**, including their roles, functions, and working mechanisms, describes methods for identifying and annotating **lncRNAs**, and discusses future opportunities for lncRNA-based therapies using **antisense oligonucleotides**.

Introduction

Recent advances in sequencing technologies enabling more in-depth genomic and transcriptomic analyses have revealed that as much as 85% of the human genome is transcribed [1–3]. This was surprising, as studies of mammalian genomes have shown that a drastically low population of RNA transcripts code for protein products. Notably, from the recent Encyclopedia of DNA Elements (ENCODE) work, out of 41,204 called genes, only 56 genes (0.1%) showed mass spectrometric evidence consistent with protein expression, suggesting that the majority of RNA transcripts are non-coding [4]. The generation of such a large population of non-coding RNA (ncRNA) transcripts indicates that RNAs have a larger and more diverse role in biological processes than initially anticipated. ncRNAs can be roughly classified into two groups based on their size. One group includes short RNAs less than 200 nucleotides (nt) in length, such as microRNAs (miRNAs) that are small RNA (sRNA) molecules around 21–24 nt in length, as well as other classes such as piwi-interacting RNAs (piRNAs) [5]. The other group includes long ncNAs (lncRNAs) of around 200 nt or more [6]. While miRNAs have been heavily studied and are well understood for their function in gene regulation [7], lncRNAs in contrast are less understood.

Many lncRNAs have been functionally associated with human diseases, in particular, cancers [8] (**Table 1**). Dysregulation of lncRNAs has been implicated in glioblastoma [9–11],

* Corresponding author.
E-mail: melissa.fullwood@nus.edu.sg (Fullwood MJ).
[a] ORCID: 0000-0002-2611-8424.
[b] ORCID: 0000-0003-0321-7865.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

http://dx.doi.org/10.1016/j.gpb.2015.09.006

Table 1 Examples of lncRNAs in cancer

lncRNA	Description	Refs.
PTCSC3	Downregulated in thyroid cancers	[82]
HULC	Upregulated in hepatocellular carcinoma	[81,139–141]
XIST	Dysregulated in various cancers	[16,142]
GAPLINC	Associated with poor prognosis in gastric cancer	[143]
MALAT1	Associated with poor prognosis and metastasis in liver, lung and colorectal cancers	[144–147]
HOTAIR	Associated with metastasis in colorectal, liver, pancreatic, breast and gastric cancers	[12–14,148–150]
ANRIL	Upregulated in prostate cancer	[151]
KCNQ1OT1	Upregulated in colorectal cancer	[59]
PRNCR1	Upregulated in prostate cancer	[152]
H19	Highly expressed in hepatocellular carcinoma	[153]

breast cancer [12], colorectal cancer [9,13], liver cancer [14,15], and leukemia [9,16]. Commonly, dysregulation of lncRNAs exerts impacts on cellular functions such as cell proliferation, resistance to apoptosis, induction of angiogenesis, promotion of metastasis, and evasion of tumor suppressors [8,17].

An emerging view of lncRNAs is that they are fundamental regulators of transcription. This view has led to an intense focus on elucidating the molecular mechanisms that underlie their function [18]. Many lncRNAs have been characterized and several models of action have been proposed, such as functioning as signal, decoy, scaffold, guide, enhancer RNAs, and short peptides [19,20], as mentioned below. The main function of a signal lncRNA is to serve as a molecular signal to regulate transcription in response to various stimuli. Thus its production and presence can serve as an indicator of transcriptional activity [21]. Decoy lncRNAs limit the availability of regulatory factors by presenting "decoy" binding sites. These lncRNAs modulate transcription by sequestering regulatory factors including transcription factors, catalytic proteins, subunits of larger chromatin modifying complexes, as well as miRNAs, thereby reducing their availability [22]. Transcripts from the scaffold class of lncRNAs play a structural role by providing platforms for assembly of multiple-component complexes, such as ribonucleoprotein (RNP) complexes [23]. In the case of RNP complexes, once the complexes have been fully assembled, either transcriptional activation or repression could be conferred depending on the nature of proteins and RNAs present [24,25]. Guide lncRNAs interact with RNPs and direct them to specific target genes. These guide lncRNAs are essential for the proper localization of RNPs [26]. Next, enhancer RNAs (eRNAs) are produced from enhancer regions and may work by influencing the 3-dimensional (3D) organization of DNA, also known as "chromatin interactions". One hypothesized model of action is that these lncRNAs may possibly work as "tethers" as they may not be released from the enhancer regions when functioning, thus tethering the interacting proteins to enhancer regions [19]. In addition, lncRNAs can encode short peptides, which may also have functions [27]. It is likely that additional mechanisms will be discovered in the future.

This review introduces lncRNAs, focusing on lncRNAs with cancer-associated roles, and discusses proposed mechanisms by which these lncRNAs function in chromatin remodeling, chromatin interactions, and as competing endogenous RNAs (ceRNAs). In addition, we highlight approaches for the identification and annotation of lncRNAs, cross-talk between these mechanisms, and methods for perturbing specific lncRNAs, which could eventually provide lncRNA-based therapies for diseases. We then conclude by highlighting challenges and future research topics.

Characteristics of lncRNAs

lncRNAs are defined as RNA molecules with more than 200 nucleotides. This distinction, while somewhat arbitrary and based on technical aspects of RNA isolation methods, serves to distinguish lncRNAs from miRNAs and other sRNAs. lncRNAs are present in large numbers in genome [28,29]. They typically do not possess functional open reading frames (ORFs). However, this distinction is blurred by the discovery of bifunctional RNAs that can have both protein-coding and coding-independent functions [30,31], raising the possibility that many protein-coding genes may also have non-coding functions. Many lncRNAs are lowly expressed [32], posing a challenge in terms of exploration of lncRNAs and explaining why lncRNAs had been thought to be only "transcriptional noise" until recently. RNA-seq studies in different tetrapods show that most (81%) lncRNAs are poorly conserved in DNA sequence and are primate-specific. However, it should be noted that several lncRNAs are ultra-conserved in DNA sequence—3% of lncRNAs appear to have originated more than 300 million years ago and can be found from organisms ranging from *Xenopus* and chicken to man [33]. It is possible that lncRNAs might be fast-evolving RNA species that can play key roles in specifying lineages. In support of this idea, a comparison of matched tissues in *Mus musculus domesticus, Mus musculus castaneus*, and *Rattus norvegicus* shows that the emergence or extinction of intergenic lncRNAs is associated with changes in transcription levels of proximal protein-coding genes [34]. In addition, there are examples of lncRNAs exhibiting conserved biological function but low sequence conservation, such as megamind/TUNA, which is associated with brain development in zebrafish, mouse, and human [35,36], as well as X-inactive specific transcript (*Xist*), which is involved in X-inactivation [37]. It is possible that RNA molecules need less sequence conservation to retain their function compared to proteins. Conversely, there is high sequence conservation of lncRNA promoters, which is even higher than that of protein-coding gene promoters [29], suggesting that regulation of lncRNA expression is important.

Many lncRNAs possess features reminiscent of protein-coding genes, such as 5′ cap and alternative splicing [32].