Biochemical and Biophysical Research Communications xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Review

Widespread genome transcription: New possibilities for RNA therapies

Hazuki Takahashi, Piero Carninci *

RIKEN Center for Life Science Technologies, Division of Genomic Technologies, RIKEN Yokohama Campus, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama City, Kanagawa 230-0045, Japan

ARTICLE INFO

Article history:
Received 21 August 2014
Available online xxxx

Keywords:
Antisense
Non-coding RNA
SINEUP
Gene therapy
Antiviral therapy

ABSTRACT

Comprehensive analysis of mammalian transcriptomes has surprisingly revealed that a major fraction of the RNAs produced by mammalian cells and tissues is comprised of long non-coding RNAs (lncRNAs). Such RNAs were previously disregarded as useless, but recent functional studies have revealed that they have multiple regulatory functions. A large subset of these lncRNAs are antisense to protein-coding genes; such RNAs are particularly attractive to researchers because their functions are better understood than other lncRNAs and their action can be easily modulated and engineered by modifying the antisense region. We discuss various aspects of regulation by antisense RNAs and other small nucleic acids and the challenges to bring these technologies to gene therapy. Despite several remaining issues related to delivery, RNA stability, side effects, and toxicity, the field is moving quickly towards future biotechnological and health applications. Therapies based on lncRNAs may be the key to increased cell-specificity of future gene therapies.

Contents

1. A multitude of regulatory RNAs . . . 00
2. Role of some lncRNA in diseases . . . 00
3. SINEUPs: surprising AS RNAs that enhance translation . . . 00
4. AS RNA in viral pathogenesis . . . 00
5. Use of exogenous RNAs to interfere with cell programs . . . 00
6. In search of stable expression . . . 00
7. Are RNA-based therapies feasible? . . . 00
8. We need to better understand RNA structures . . . 00
9. Perspectives . . . 00
Acknowledgments . . . 00
References . . . 00

1. A multitude of regulatory RNAs

One of the surprises of genomics has been the discovery that mammalian genomes produce a large amount of non-coding RNA. Apart from the production of small RNAs, such as microRNAs (miRNAs), which often function as negative regulators of RNA stability or translation, there are many long non-coding RNAs (lncRNAs) that have been well categorized both in mouse [1] and human [2–4]. Results of large-scale analyses suggest a striking but inconvenient truth: there are more non-coding RNAs than the ~20,000 protein-coding genes. The FANTOM3 project, by using physical cDNA clone analysis alone, identified more than 23,000 lncRNAs [1]. The GENCODE project (release 20), as part of the ENCODE project [5], identified almost 24,000 loci producing lncRNAs, but the total census of non-coding RNAs may be above 37,000 if miRNAs, RNA from processed pseudogene transcripts,

Abbreviations: Aβ, amyloid beta; AGO, argonaute; AON, antisense oligonucleotide; AS RNA, antisense RNA; ATL, adult T-cell leukemia; EBOV, Ebola virus; eRNA, enhancer RNA; ES, embryonic stem; HBZ, HTLV-1 basic leucine zipper factor; HTLV-1, human T-cell leukemia virus type 1; IL, interleukin; iPSC, induced pluripotent stem cell; lncRNA, long non-coding RNA; LTR, long terminal repeat; MARV, Marburg virus; miRNA, microRNA; NAT, natural antisense transcript; X-SCID, X-linked severe combined immunodeficiency; UTR, untranslated region.

* Corresponding author.
E-mail address: carninci@riken.jp (P. Carninci).

http://dx.doi.org/10.1016/j.bbrc.2014.08.139
0006-291X/

and other types of lncRNAs were to be included. The number of lncRNAs may still rise if next-generation sequencing studies focus on cell types that are not yet completely characterized. For example, in human embryonic stem (ES) cells and induced pluripotent stem cells (iPSCs), deep sequencing of cytoplasmic and nuclear transcripts has uncovered more than 3000 lncRNAs that were not previously identified in the GENCODE catalogue, including a large fraction of RNAs derived from long terminal repeats (LTRs) and other retrotransposon elements [6]. The function of the majority of lncRNAs is still unknown; however, whenever the function has been determined, lncRNAs show a remarkable diversity of function and mechanisms of action in various cell compartments.

After synthesis, a large number of lncRNAs never leave the nucleus, which makes this organelle a large source of uncharacterized lncRNAs [6,7]. A fraction of these lncRNAs closely interact with the chromatin and help to direct the epigenome. For instance, lncRNAs can mediate the interaction between polycomb repressive complex and target genes in ES cells [8], or mediate X-chromosome inactivation [9] by physically causing chromatin condensation around interacting regions in *cis* on X-chromosome [10,11]. Alternatively, lncRNAs can act in *trans*: e.g., *HOTAIR*, whose misexpression is associated with cancer [12]. To further complicate the interpretation and analysis of lncRNAs, active enhancers also produce lncRNAs, often called eRNAs [6,13], whose function is mostly unknown. At least in some cases, eRNAs could be essential structural molecules that promote nuclear interactions through the multi-protein complex called Mediator, thus enhancing the set of potential regulatory lncRNAs [14]. LncRNAs can also have other structural functions in the nucleus: for instance, *NEAT1* is an essential RNA component of structures known as paraspeckles [15]. There are numerous cases where lncRNAs are processed to produce smaller RNAs, which have various regulatory functions, including not only miRNAs [16] and small nucleolar RNAs (commonly known as snoRNAs) [17] but also other classes of small RNAs [18,19]. Due to space limitations we cannot extensively review all aspects of RNA processing and we apologize for not citing some important works of colleagues: please see other reviews [20–23] for extensive discussions of lncRNAs and their processing.

Antisense transcripts [24] are a broad and very important component of the non-coding transcriptome. The mechanisms of action and regulation of antisense lncRNAs are somehow better understood than those of other lncRNAs. Because antisense sequences can easily be designed, lncRNA action based on sense-antisense transcript pairing is highly flexible in principle. Accordingly, we will focus on the biology and current applications of antisense lncRNAs, and how these transcripts could be used for future therapies.

2. Role of some lncRNA in diseases

Some antisense lncRNAs can positively regulate transcription [25,26]. A prime example is the regulation of the gene encoding β-secretase 1 (BASE1) (also known as β-site amyloid precursor protein-cleaving enzyme), which is implicated in the pathogenesis of Alzheimer's disease. Transcription of *BASE1* is positively regulated by its noncoding antisense transcript (*BASE1*-AS). Knockdown experiments with small interfering RNA (siRNA) against *BASE1*-AS RNA caused reduction of amyloid β (Aβ) 1–40 and Aβ 1–42 protein production in human SH-SY5Y cells. These results suggest that *BASE1*-AS transcript is positive regulator of transcription of the coding gene. *In vivo* experiments showed that 14 days continuous treatment with siRNA targeting *BASE1*-AS transcript reduces the *BACE1* mRNA expression level in mouse brain regions (cortex, striatum, dorsal hippocampus, and ventral hippocampus). Cell stress with high temperature, serum starvation, Aβ 1–42 accumulation, exposure to H_2O_2, or treatment with a high glucose concentration caused a 30–130% increase in *BACE1*-AS levels associated with a 20–60% increase in *BACE1* mRNA levels. When *BACE1*-AS transcript levels in the cerebellum, hippocampus, entorhinal cortex, and superior frontal gyrus were compared between Alzheimer's disease patients and control subjects, the *BACE1*-AS levels were elevated in Alzheimer's disease subjects by up to six fold, with an average of about two fold across all brain regions [25]. This study is particularly remarkable because the lncRNA expression level can positively control the mRNA expression level.

Some of the mechanisms of gene regulation based on sense-antisense pairing in the nucleus appear to involve transcriptional interference [27] or the siRNA pathway. For instance, in the nucleus, argonaute (AGO) proteins, a part of the RNA-induced silencing complex (RISC), are involved both in specific transcriptional activation and repression [28], suggesting broad involvement of the siRNA machinery on the chromatin. Importantly, although AS RNAs are commonly believed to be mostly negative regulators of sense mRNA counterparts [25], there are many more mechanisms of action, many of which are not yet fully understood.

Apart from regulating transcript stability, antisense lncRNAs can regulate transcription initiation [24,27]. This phenomenon could be used in future AS RNA therapies: in fact, inhibition and degradation of natural antisense transcripts (NATs) could upregulate some specific mRNAs for therapeutic purposes. The group led by Claes Wahlestedt found that inhibition of the NAT of the gene for brain-derived neurotrophic factor (BDNF) increases the level of the sense mRNA transcript and BDNF protein expression, leading to neuronal outgrowth and differentiation *in vitro* and *in vivo*. In a follow-up study, the same group investigated two additional RNAs, which are antisense to the mRNAs encoding glial-derived neurotrophic factor (GDNF) and ephrin receptor B2 (EPHB2). Inhibition of these NATs resulted in increased sense mRNA transcript levels [29], showing that this is a general phenomenon. The application of these approaches for increasing protein production is becoming recognized as important for human therapy. Accordingly, the company OPKO-CURNA in USA is exploiting this approach, which was named "Inhibition of Antisense Transcripts to Upregulate its Sense mRNA". These new RNAs, named AntagoNATs, are designed to target the NAT region overlapping the mRNA and thus inhibit the sense-antisense RNA pairing (Fig. 1A).

Since many antisense and other lncRNAs are regulated by cellular stress, it is very important to understand their dynamic expression and regulation when cells or tissues are under stress, including during heat shock [30] or immune challenge [31].

Genetic mutations can affect antisense regulation. For example, α-thalassemia is a disorder caused by reduced production of functional globins due to gene mutation. In particular cases, the mutation increases expression of an antisense transcript, called *LUC7L*, which induces transcriptional silencing of globin genes by methylation of the CpG island [32].

3. SINEUPs: surprising AS RNAs that enhance translation

Reduction in the expression of the *UCHL1* gene (also called *PARK5*), is positively correlated with familial Parkinson's disease [33]. A surprising novel mechanism of sense-antisense action came from the study of AS *Uchl1* RNA (AS-*Uchl1*) in mouse [34]. Large-scale analysis of full-length cDNA sequences by the FANTOM3 consortium reported widespread sense-antisense transcription (>72% of identified genes in mouse) [24]. Using the FANTOM3 cDNA clones, Carrieri et al. investigated the functional role of AS RNAs, including AS-*Uchl1*, in the mouse dopaminergic neuronal cell-line, MN9D. AS-*Uchl1* overlaps the 5′ untranslated region (UTR) and translational start site of the *Uchl1* mRNA [34]. Surprisingly, while *Uchl1*-AS does not have any effect on the stability of the mRNA, it

Please cite this article in press as: H. Takahashi, P. Carninci, Widespread genome transcription: New possibilities for RNA therapies, Biochem. Biophys. Res. Commun. (2014), http://dx.doi.org/10.1016/j.bbrc.2014.08.139