Published in final edited form as:

Adv Drug Deliv Rev. 2015 June 29; 87: 15–24. doi:10.1016/j.addr.2015.05.012.

Basic Biology and Therapeutic Implications of lncRNA

O Khorkova^a, J Hsiao^a, and C Wahlestedt^{b,*}

^aOPKO Health Inc., 10320 USA Today Way, Miramar FL 33025, USA.

^bCenter for Therapeutic Innovation and the Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, 1501 NW 10th Avenue, Miami 33136, Florida, USA.

Abstract

Long non-coding RNAs (lncRNA), a class of non-coding RNA molecules recently identified largely due to the efforts of FANTOM, and later GENCODE and ENCODE consortia, have been a subject of intense investigation in the past decade. Extensive efforts to get deeper understanding of lncRNA biology have yielded evidence of their diverse structural and regulatory roles in protecting chromosome integrity, maintaining genomic architecture, X chromosome inactivation, imprinting, transcription, translation and epigenetic regulation. Here we will briefly review the recent studies in the field of lncRNA biology focusing mostly on mammalian species and discuss their therapeutic implications.

Graphical Abstract

A - Gene duplication and repurposing of pseudogenes is one of proposed routes of lncRNA evolution. **B** - Mobile genetic elements (MGE) frequently initiate formation and evolution of new transcriptional units (TU) some of which become lncRNA. **C** - lncRNA participate in long range DNA looping essential for transcription regulation in some loci. **D** - lncRNA are essential for telomere maintenance. **E** - lncRNA participate in X-chromosome inactivation and imprinting. **F** - lncRNA scaffold and regulate formation of nuclear paraspeckles, as a result also controlling the nuclear-cytoplasmic transport of mRNA. **G** - lncRNA are involved in formation of interchromatin granules enabling pre-mRNA splicing and maturation. **H** - Fine regulation of epigenetic modifications is assisted by tethering of epigenetic effectors and formation of polycomb bodies by lncRNA. **I** - Cytoplasmic roles of lncRNA include positive and negative regulation of mRNA stability. **J** - lncRNA are involved in regulation of translation and positive and negative regulation of nascent protein stability. **K** - lncRNA act as miRNA sponges blocking miRNA activity.

*To whom correspondence should be addressed: C.W. (CWahlestedt@med.miami.edu 305-243-1367).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Natural antisense transcripts; lincRNA; vlincRNA; telomeres; paraspeckles; transposable elements; X-chromosome inactivation; imprinting; epigenetic regulation; rare genetic diseases; AntagoNAT

1. Introduction

The scope of RNA's biological functions has been recently expanded from a single role of a messenger for protein synthesis to a wide spectrum of regulatory and structural roles in protecting chromosome integrity, maintaining genomic architecture, X chromosome inactivation, imprinting, transcription, translation and epigenetic regulation. This expansion occurred largely due to the pioneering efforts of the FANTOM, and later GENCODE and ENCODE consortia [1-5]. The newly discovered RNA functions are mostly carried out by non-coding RNA (ncRNA), including its most recently discovered class – long non-coding RNAs (lncRNAs).

lncRNAs are roughly defined as RNA molecules of more than 200 bases in length with no protein-coding capacity. Compared to protein-coding mRNAs, lncRNAs on average are present in lower abundance, frequently reside in the nucleus, are more tissue-specific and have poorer interspecies sequence conservation [6]. While individual species of lncRNA have low abundance, taken together lncRNA could comprise up to 64% of all non-ribosomal/non-mitochondrial/non-repeat transcripts in human cells [7].

Although classification of lncRNA is currently vastly incomplete, two large classes based on lncRNA position relative to protein-coding genes could be defined: 1) intergenic lncRNAs, including pseudogenes, long intergenic non-coding RNAs (lincRNAs), and very long intergenic non-coding RNAs (vlincRNAs/macroRNAs) among others; 2) coding gene-overlapping lncRNAs, represented by intronic RNAs, natural antisense transcripts (NATs),

promoter RNAs and a continuously expanding list of other subclasses. Efforts are currently underway to clarify and standardize the existing lncRNA nomenclature [8]. We will focus mostly on the relatively more studied lincRNAs and NATs. lincRNAs are localized in intergenic regions, usually span very large areas of the chromosome and regulate large sets of genes in *trans* [9,10]. NATs are localized within or immediately adjacent to protein-coding loci, usually on a strand opposite to the coding gene, and regulate a single gene or a small group of genes in *cis* [11]. NATs can be canonically processed (including a 5' 7-methylguanylate cap, intron removal, and 3' polyadenylation), RNA-edited and are usually transcribed by RNA-polymerase II [reviewed in 12]. However, some lncRNAs exhibit both *cis* and *trans* regulation and are difficult to assign to one class (e.g. ANRIL [13,14], Paupar [15], VAT [16]). In an even more complex combination of traits, lncRNA Nesp epigenetically regulates the GNAS locus as an RNA transcript but also encodes a protein [17]. Another example is steroid receptor RNA activator (SRA) which was initially thought to be non-coding, but was later found to produce a highly conserved small protein (SRAP; [18]).

Extensive efforts to get deeper understanding of lncRNA biology have yielded multiple publications and databases of functional lncRNA information including lncRNAdb, NONCODE, lncRNAtor, lncRNADisease, ncFANs, InCeDB, LNCipedia, psiDR [2,19-21]. Here we will briefly summarize the recent achievements in the studies of biology and therapeutic implications of lncRNA, focusing mostly on mammalian species (Fig.1).

2. Evolution of lncRNA

Multiple theories of lncRNA origins and evolution have been proposed. Two of them (gene duplication and generation by mobile genetic element insertion) are briefly described below. However, the full scale of achievements in this area is too vast and is comprehensively reviewed elsewhere [see for example 22].

2.1. Gene duplication and repurposing of pseudogenes

Gene, chromosome and genome duplication are thought to be the major moving forces of biological evolution. Gene duplication may be created by unequal crossing over, retroposition or chromosomal/whole genome duplication [23]. Resulting duplicate genes may become inactive (termed 'pseudogenes'), evolve a different protein function or give rise to a regulatory lncRNA transcript (Fig.1A)[2,21,24]. A well-studied example of this evolutionary route is lncRNA Xist derived from duplicated protein-coding gene [24].

2.2. New transcriptional unit formation by mobile genetic elements

Mobile genetic elements (MGE) discovered by Barbara McClintock more than 60 years ago, are believed to have contributed widely to the evolution of eukaryotic genomes. Of all currently recognized classes of MGE, transposable elements (also called TE or transposons) had the most immediate impact on mammalian genomes. As the name implies, originally TEs were able to transpose themselves into genomes and initiate transcription of genes necessary for their propagation, but have largely lost this ability during evolution. DNA remnants and some active copies of TEs constitute more than 50% of the mammalian

genome sequence [25]. TEs are divided into retrotransposons, DNA transposons and insertion elements. Retrotransposons are the most frequent type and include 1) long terminal repeat transposons (LTR), e.g. human endogenous retroviruses (HERVs), and 2) non-LTR transposons, with 2 subtypes, long and short interspersed nuclear elements (LINEs and SINEs). Some of the TEs which had a large impact on the human genome belong to LINE (L1, L2, L3) and SINE types (Alu).

While majority of the TE copies lost their capacity to propagate, some of the elements used by TEs to control their own transcription may create new transcriptional units in the host DNA, including lncRNAs (Fig.1B). The unmodified transcripts from active HERVs can regulate the coding genes in which they reside [92]. Additionally, TE-derived elements may act as exons, splice sites, enhancers, promoters and other RNA-, DNA-, and protein-binding domains of lncRNA, contributing up to 41% of lncRNA nucleotides [25,26].

2.3. Poor lncRNA sequence conservation among species

Overall, lncRNAs have higher conservation than the neutrally evolving regions of the genome, but lower conservation than the protein coding genes [26]. A relatively low level of conservation may be explained by the fact that structural and regulatory functions of the ncRNA molecules mostly depend on their 3D shape, as opposed to functions of protein-coding RNAs which mainly depend on their sequence. The likelihood of secondary-structure-driven evolution is supported by the overrepresentation of TE sequences (known to maintain complex and stable RNA secondary structures) in lncRNAs as compared to protein coding genes.

Another possible reason for low lncRNA sequence conservation among species could be the wide participation of TEs in their evolution. TE abundance and composition in mouse and human genomes are substantially different [26]. For example, primate-specific Alu elements are enriched in human lncRNA, while in mouse, where Alu are not present, the mouse-specific elements stand in [26], creating different conditions for lncRNA evolution in these species. Interestingly, cases of convergent evolution of lncRNA have been described. For example, formation of STAU1-binding sites by interaction of coding and non-coding RNAs through Alu stretches, found in humans, also evolved in mouse, where STAU1-binding is mediated by the mouse-specific B1, B2, and B4 elements [25].

The low lncRNA sequence conservation rule, however, is not all-inclusive. For example, lncRNA expressed from the 200–779 bp long transcribed-ultraconserved regions (T-UCRs) have 100% conservation in mouse, rat, and human genomes [reviewed in 28]. Additionally, lncRNA splice sites have much higher conservation than the rest of lncRNA sequence and can be used to trace the evolution of lncRNAs. Using this technique it has been shown that more than 85% of human GENCODE lncRNAs were present in early placental mammals [29].

2.4. High evolutionary activity of lncRNA

Low conservation of lncRNA among species also reflects their extremely fast evolution. It has been shown that HERVH elements associated with human lincRNAs are evolutionarily younger than HERVH elements genome-wide [26]. A good illustration of the lncRNA

evolutionary pace is ANRIL, which originated in the eutherian ancestor 100-150 million years ago and initially contained only a few exons. Consequently, multiple transposons were either inserted into existing ANRIL exons, or formed additional exons *de novo*. Interestingly, while ANRIL exon number progressively increased in simians, rodents gradually lost multiple exons during the evolution [30].

3. lncRNA biology

Studies of lncRNA biology have been gaining ground in the last 10 years since the pioneering work of the FANTOM, and later GENCODE and ENCODE consortia brought about the realization that a significantly larger part of the genome was being transcribed than previously believed. While arguably just scratching the surface, these studies have revealed lncRNA participation at all levels of nuclear architecture and gene expression, from telomere maintenance to long range DNA looping to X-chromosome inactivation and imprinting to fine-tuning single gene transcription, translation and mRNA stability.

3.1. lncRNA roles in replication: telomere stability

Telomeres are protective nucleoprotein structures at the end of chromosomes that include double-stranded 5'-TTAGGG-3' repeats terminating in 3' protruding single-stranded G-rich overhangs. Telomeres are capped with t loop structures, formed by telomeric repeat-binding factor 2 (TRF2), in which the telomeric 3' G overhang is tucked into the double-stranded stretch of telomeric DNA. This DNA loop is associated with shelterin proteins which in humans are represented by TRF1, TRF2, POT1, TIN2, TPP1, and Rap1 [31]. Due to peculiarities of the chromosome replication process and compounding environmental factors, telomeres shorten at each mitosis. Extreme telomere shortening or uncapping triggers a cell senescence response and apoptosis. Uncapping of telomeres can be caused by decreased TRF2 availability which allows two protein complexes, XPF/ERCC1 and MRE11/RAD50/NBS1 (the latter also involved in ATM-mediated checkpoint activation and cell cycle arrest) to digest the 3' G overhangs [32]. Damaged telomeres can be extended at 3' end by telomerase, a reverse transcriptase complex that carries its own RNA template for building telomere repeat sequences. Telomerase activity is high in stem cells and cancers. In adult somatic cells however telomerase expression is suppressed, which limits the number of divisions the cells can undergo [reviewed in 33]. As a consequence, regulation of telomere length has an important part in the treatment of cancers.

lncRNA TERRA has been implicated in chromosome replication and maintaining telomere homeostasis through at least 3 different mechanisms (Fig.1D). TERRA is an RNA polymerase II transcript containing telomeric UUAGGG repeats, conserved in eukaryotes, and transcribed from subtelomeres toward chromosome ends. TERRA expression is lowest in the late S phase of the cell cycle and peaks in early G1. Due to its homology to telomeric repeats, TERRA can sequester single strand DNA-binding protein hnRNPA1. Decrease in TERRA expression during S phase frees hnRNPA1 to bind single stranded telomeric DNA. Thus hnRNPA1 binding to telomeric ssDNA is timed to occur after completion of DNA replication. hnRNPA1 displaces RPA, a protein essential for DNA replication and activation of ATR checkpoint kinase. Release of RPA from telomeres also prevents erroneous activation of ATR by telomeric ssDNA during normal cell cycle progression. Because

hnRNPA1 only displaces RPA, but does not affect POT1, one of the proteins of the shelterin complex, POT1 occupancy at telomeric ssDNA eventually increases, enhancing telomere protection [34].

The second mechanism through which TERRA can participate in chromosome homeostasis is via direct regulation of telomerase activity, although this regulation is quite complex. In some cells TERRA has been shown to repress telomerase, thus slowing the telomere repair [35]. At the same time, telomere shortening has been shown to induce TERRA expression. Furthermore, TERRA may facilitate the nucleation of telomerase molecules into clusters prior to their recruitment at a short telomere [36]. TRF2 depletion, which leads to telomere uncapping, also induces TERRA expression which in turn may facilitate loading of the MRE11/RAD50/NBS1 complex onto DNA, digestion of the 3' G overhangs and DNA damage response [32].

Finally the possible third mechanism of TERRA-mediated chromosome homeostasis regulation is through TERRA's recently discovered role in the alternative telomere maintenance pathway, which employs chromosome recombination machinery [37].

3.2. lncRNA in control of nuclear architecture and transcription

3.2.1. Long range DNA looping and trans-chromosomal interactions—Recent discoveries in the lncRNA field were facilitated by the development of new methods including chromosome conformation capture (3C), live super-resolution (PALM) microscopy and fluorescence in situ hybridization (FISH). These methods have revealed that enhancer/promoter pairs, or several different promoters separated by large stretches on a chromosome, or even located on different chromosomes, may be brought together in the nucleus by chromosomal looping (Fig.1C)[reviewed in 38]. For example, in the NF- κ B-regulated multigene transcription complex that includes SAMD4A, TNFAIP2, and SLC6A5 genes, promoter-promoter interactions are shown to form prior to transcription start and to drive coordinated transcription of this gene cluster in HUVEC cells. It has been proposed that NF- κ B and RNAP II are passed on from the dominant to the subordinate promoters in this complex [39]. Such long range chromosomal interactions are believed to occur within discrete megabase-sized units, termed “topological domains”, delimited by high concentration of CTCF-binding sites [38]. CTCF is a protein which forms a ring-like complex with cohesin and other proteins to enable long-range chromosomal looping. The CTCF complex is loaded onto chromatin by NIPBL with possible participation of lncRNA [38,40,41].

Shorter range chromosomal interactions, for example between promoter and enhancer, can be mediated by lncRNAs expressed from the enhancer site (eRNA). Initially eRNAs bind Mediator protein complex to cause chromatin bending. Mediator then binds NIPBL, followed by cohesin at the chromatin contact site to stabilize the chromosomal loop. lncRNAs likely play a critical role in this mechanism since in SNAI1 and AURKA loci depletion of ncRNA-a7 reduces chromosomal looping to a similar extent as depletion of Mediator subunits MED1 or MED12 [41]. Additionally, lncRNA may provide positive feedback in enhancer-promoter interactions. For example, when estrogen receptor α (ESR1) binds to enhancers in ESR1-target coding genes, it causes increase in eRNA transcription.

The induced eRNAs in turn increase the strength of the enhancer/promoter looping [42]. eRNAs may also participate in the formation of multiple-promoter chromosomal loops. For example, eRNA from the NRIP1 enhancer was shown to facilitate interactions between the NRIP1 and TFF1 loci, located 27 Mb apart on human chromosome 21 [42].

lncRNA may also participate in trans-chromosomal interactions. For example lncRNA Firre, localized on X chromosome, interacts with one of the components of the nuclear matrix, hnRNPU, through a 156-bp repeating sequence and coordinates co-localization of at least three gene loci located on chromosomes 2 and 17 [9].

3.2.2. X-chromosome inactivation—lncRNAs play an important role in X-chromosome inactivation, which is necessary to reduce X-chromosome gene expression in female genomes (Fig.1E). lncRNA Xist allele located on the X chromosome targeted for inactivation (Xi), produces multiple copies of Xist transcripts that coat the entire length of Xi. Transcription repressor protein YY1, together with two lncRNAs, Jpx and Ftx, determines allele-specific binding of Xist to the Xi. Xist recruits the repressive chromatin modifier complex (polycomb repressive complex 2 or PRC2) that induces histone H3-K27 methylation (H3K27me3), a hallmark of gene inactivation, and silences gene expression from Xi. These processes are regulated by other lncRNAs, including Tsix, transcribed from Xist locus in antisense direction, and Xite [43].

3.2.3. lncRNA in imprinting—Some of the genes in the mammalian genome are expressed from only one allele, which could be maternal or paternal depending on the gene. The mechanism which determines which one of the alleles would be inactivated is called imprinting. In mammals allele inactivation is defined by epigenetic marks established in one of the parental germlines at imprinting control regions (ICRs). Deposition of these marks is believed to be guided by lncRNAs expressed from antisense strands of the imprinted coding loci. These lncRNAs tether chromatin-modifying complexes, such as PRC1 and PRC2, which create repressive chromatin marks at the imprinted loci [44, 45]. Participation in imprinting has been shown for multiple lncRNAs including Kcnq1ot1, Airn, Tsix, ANRIL, AIRN, and H19, which control the imprinted expression of Igf2, Igf2r, and Kcnq1 [43,46]

3.2.4. Formation of nuclear subcompartments—Some of the known nuclear subcompartments are assembled on lncRNAs. For example, nascent transcripts of lncRNA NEAT1 have been shown to scaffold nuclear paraspeckles, mammalian-specific ribonucleoprotein bodies (Fig.1F). Possible roles of paraspeckles include sequestering RNAs which contain inverted SINE repeats in their 3'UTRs, as well as hyper-A-to-I-edited RNAs, in the nucleus [47]. Additionally, NEAT1 sequesters SFPQ, a factor required for transcription of ADARB2, an enzyme involved in A-to-I editing. NEAT1 may also be further involved in transcriptional control through sequestration of other paraspeckle-associated transcriptional factors [47]. Biological importance of NEAT1 is reflected in its induction upon virus infection or by immune stimuli [48] and during early stages of ALS pathogenesis [49].

lncRNAs TUG1 and MALAT1, located in transcriptional repressive Polycomb bodies (PcGs, Fig.1H) and permissive interchromatin granules (ICGs, Fig.1G), respectively, control

relocation of growth-control genes between these 2 structures in response to growth stimuli. These ncRNAs also coordinate assembly of multiple corepressors/coactivators and can affect mark recognition by the readers of the histone code [50].

3.2.5. Fine regulation of epigenetic modification—The scaffolding potential of lncRNA is further employed in a more fine-grain regulation of transcription. lncRNAs can act in *cis* or *trans* to guide epigenetic modifier complexes to targeted genomic sites [51]. The effects of this tethering, frequently mediated by lncRNA from the NAT subclass, affect the expression of one gene or a small cluster of related genes and cause small but biologically highly significant changes in target mRNA levels (Fig.1H)[11]. The tethering can occur by the nascent lncRNA at the time of transcription [6] or after NAT transcription has been completed, by pairing with DNA or mRNA sequences. There are several possible configurations for the pairing, including base pairing between the lncRNA and ssDNA, formation of a RNA-DNA-DNA triplex or via RNA-RNA hybrids of lncRNAs with a nascent partner mRNA. It is likely that tethering requires participation of multiple protein factors [52]. PRC2 could be one of the main histone modification complexes tethered by lncRNAs since a substantial fraction of lncRNAs was found to associate with this complex [45,53,54]. PRC2 catalyzes the di- and tri-methylation of histone H3 at lysine 27 (H3K27me₂ or H3K27me₃), which can then be recognized by PRC1. PRC1 complex can further induce the monoubiquitylation of histone H2A, which in turn contributes to chromatin compaction and repression of the target locus [55]. This epigenetic regulatory model may explain functionality when the abundance of lncRNA molecules is extremely low, as has been shown for many lncRNA [54]. lncRNA-mediated modulation of epigenetic modifications is known to affect multiple biologically critical genes. In most cases lncRNAs act as transcriptional repressors (e.g. ANRIL, which recruits PRC2 to the CDKN2B locus causing H3K27-trimethylation and CDKN2B silencing, APOA1_NAT which represses several genes in the APOA cluster, BDNF-OS which suppresses expression of BDNF, SCAANT1 which inhibits ATXN7, and others [6,40,56-59]. There are, however, reports of lncRNAs acting as gene activators (e.g. Hottip, Mistral) [51].

A more complex case of epigenetic regulation in *trans*, employing direct interaction of a vlncRNA VAD with the histone variant H2A.Z (which regulates the chromatin recruitment of Suz12, a component of the PRC2 complex) and a p400 protein (which mediates H2A.Z incorporation into chromatin), has been reported recently [16]. VAD binding decreased H2A.Z occupancy at the INK4 promoters, thus leading to increased INK4 expression in senescent cells.

In some cases lncRNA may inhibit target coding gene expression by the act of transcription itself and not through their transcribed RNA product. For example, continuous expression of lncRNA Airn inhibits Igf2r mRNA in mice [reviewed in 6].

3.3 Post-transcriptional regulation of gene expression

3.3.1. Role in mRNA splicing and maturation—lncRNAs have been implicated in regulation of RNA splicing and maturation (Fig.1G)[60]. For example, lncRNA MALAT1 associates with serine/arginine splicing factors in the nuclear speckles, also known as

interchromatin granule clusters, and is involved in control of alternative splicing [50]. An interesting mechanism of alternative splicing regulation involves intron-derived long noncoding RNAs with snoRNA ends (sno-lncRNAs). sno-lncRNAs have snoRNA sequences at both ends to protect them from degradation. Expression of sno-lncRNAs is often associated with alternative splicing of exons within their parent genes. Expression patterns of sno-lncRNAs differ sharply between species. For example, sno-lncRNAs are well expressed in the RPL13A region in mouse, but are barely detectable in human. Additionally, snoRNAs from the Prader-Willi syndrome (PWS) region are highly expressed in human and monkey, but not mouse [61]. Association of PWS region sno-lncRNAs with Fox family splicing regulators and the resulting altered splicing patterns may play a role in the molecular pathogenesis of PWS [61].

lncRNA have also been shown to participate in RNA editing partially through their role in paraspeckles, regulation of ADAR activity (see above) and complementarity of their Alu repeats to Alu repeats in coding RNA which creates preferential editing sites [reviewed in 62]

3.3.2. mRNA transport/localization—lncRNA may control nuclear/cytoplasmic shuttling of mRNA, thus affecting their availability to translation machinery and the resulting protein levels, mainly by sequestration of mRNA in paraspeckles and other nuclear subcompartments (see above).

3.3.3 Control of mRNA and protein stability—mRNA degradation may occur for different reasons and through multiple mechanisms. A number of degradation pathways, for example nonsense-mediated decay (NMD) and STAU1-mediated decay (SMD), are initiated during translation and mainly have quality control purposes (Fig.1I). NMD is primarily responsible for degradation of mRNAs containing premature termination codons. Additionally, NMD degrades many non-mutated transcripts as part of normal cell metabolism. lncRNA GAS5 has been shown to participate in regulation of signaling downstream of NMD [63].

SMD affects mRNAs with a STAU1-recognition site in their 3' UTRs. STAU1 binds double-stranded RNA, formed in stem-loop structures or after duplexing of Alu elements in mRNAs and complementary lncRNAs. STAU1 binding leads to decreased mRNA half-life [64]. One lncRNA can bind a subset of SMD targets, and each target may be regulated by multiple lncRNAs [64].

Another mechanism of translational inhibition is illustrated by an lncRNA from the PU.1 locus. This lncRNA binds to the sense transcript and impedes its ribosomal entry or otherwise stops translation between initiation and elongation steps (Fig.1J)[65]. Furthermore, Huang et al. have shown that lncRNA UCA1 inhibits p27 (Kip1) mRNA translation by displacing it from the complex with hnRNPI, a protein which facilitates p27 (Kip1) translation [66].

Inversely, some of the lncRNAs can stabilize mRNAs (Fig 1I). For example, lncRNAs from the iNOS locus stabilized iNOS mRNA possibly by facilitating its binding to HuR protein,

which suppresses RNA degradation by inhibiting deadenylase or exonuclease enzymes [67]. Similarly, lncRNA BACE1-AS increased the stability of BACE1 mRNA through formation of an RNA duplex that may have altered the secondary or tertiary structure of BACE1 mRNA [68].

lncRNA-mediated regulation also affects protein stability (Fig.1J). For example, HOTAIR brings together E3 ubiquitin ligases Dzip3 and Mex3b and their respective substrates, Ataxin-1 and Snurportin-1, and thus accelerates their degradation [69]. Furthermore, proteasomal inhibition causes upregulation of paraspeckle associated lncRNA NEAT1 and NEAT1 expression in turn protects fibroblasts from cell death triggered by proteasome inhibition [47].

3.3.4. Regulation of miRNA activity—miRNA, a class of short non-coding RNAs, have been shown to regulate translation of mRNAs by capturing them for degradation by RISC complex or by sequestering them into GW/P-bodies, stress granules or other structures for later translation. In turn, lncRNAs have been shown to bind both miRNA themselves and miRNA binding sites on their target mRNAs, thus modulating miRNA activity (Fig.1K). Competing endogenous lncRNAs (termed ceRNAs [20]), which act as miRNA sponges, include, for example lncRNA CHRF, which has been shown to bind miR-489, reducing its availability for downregulating Myd88 mRNA [70]. In other examples, linc ROR blocks miR-145 [71]; lncRNA PTENP1 modulates endogenous PTEN transcript levels by acting as a molecular sponge for PTEN-targeting miRNAs [72]; linc-MD1 is a muscle-specific lncRNA which activates the expression of MAML1 and MEF2C by sponging up miR-133 [73]. Alternatively, miR-485-5p and lncRNA BACE1-AS compete for the same binding site in BACE1 mRNA, a crucial enzyme in Alzheimer's disease pathophysiology [74].

In addition to blocking miRNA binding sites, lncRNAs can also regulate pri-miRNA processing. For example, a long ncRNA Uc.283+A binds the lower stem region of the pri-miR-195 transcript and prevents its cleavage by Drosha [75].

3.3.5. siRNA formation—One of the first proposed mechanisms of lncRNA-based regulation of gene expression postulated that RNA duplex formation between the complementary regions of coding and lncRNA may lead to dicer-dependent endogenous siRNA production. Known examples of this mechanism include CEND1, FERM and FRMPD4 loci that generate endogenous siRNA by their NAT/sense transcript pairing [76,77].

3.4. Circulating lncRNA

In an important development for lncRNA-based diagnostics, multiple lncRNAs have been observed in exosomes, recently discovered cell-derived vesicles that are present in such easily accessible biological fluids as blood, urine, and cultured medium of cell cultures [78]. For example, linc-ROR was highly represented in extracellular RNA released by hepatocellular cancer cells. TGF β treatment selectively increased linc-RoR concentration within extracellular vesicles and enhanced cellular resistance to chemotherapeutic agents. Incubation of naïve cells with extracellular vesicle preparations reduced chemotherapy-induced cell death in these cells [79].

3.5. Regulation of lncRNA expression

lncRNA production is believed to be regulated by mechanisms similar to these of protein-coding genes, including transcription factor binding, DNA and histone modifications and RNA splicing, but this area of lncRNA biology remains relatively unexplored.

Jiang et al. [80] have compiled a database of transcription factors found to be associated with lncRNA promoter regions based on ChIP-Seq data. For example, expression of lncRNA TERRA has been shown to be repressed by DNMT1 and DNMT3b-mediated methylation of CpG dinucleotides present in the human subtelomeric region. H3K9 histone methyltransferase SUV39H1 and HP1 α protein, which binds H3K9me3, also inhibit TERRA transcription [32].

lncRNA stability can be affected by RNA-binding proteins. For example, HuD associated with and stabilized lncRNA BACE1-AS, which in turn enhanced BACE1 expression [81]. Zhang et al. [82] demonstrated a reciprocally negative regulation between GAS5 and miR-21. While miR-21 represses GAS5 by targeting a sequence encoded by its exon 4, GAS5 in turn inhibits miR-21 expression.

3.6. Interaction of lncRNA-mediated mechanisms

Understandably, in real life multiple aspects of RNA biology described above are involved in regulation of gene expression at the same time. An example of such complexity may be seen in regulation of PTEN expression by its pseudogene PTENP1. PTENP1 locus transcribed in forward direction produces lncRNA PTENP1. When transcribed in reverse direction, PTENP1 locus creates two isoforms of another lncRNA, α and β . The α isoform binds PTEN promoter and recruits DNMT3A and histone methyltransferases EZH2 and G9A, which mediate methylation of H3K27 and H3K9, promoting the epigenetic silencing of PTEN expression. The β isoform forms an RNA/RNA duplex with the forward lncRNA PTENP1, which increases stability of PTENP1 and facilitates its role as a miRNA sponge. As a result, β isoform activates PTEN expression [72].

4. Therapeutic aspects of lncRNAs

Several aspects of lncRNA biology make them highly attractive as therapeutic targets, including the access to multiple previously 'undruggable' proteins and ease with which their abundance can be modulated using synthetic antisense oligonucleotides. The two main areas for which lncRNA-based therapies are now being developed are cancers and rare genetic disorders. Mainly two approaches are being used – knockdown of the lncRNA from the NAT subclass and interfering with lncRNA/PRC2 interaction.

4.1. Disease association of lncRNA SNPs

Historically, studies of rare genetic disorders and of the genetic roots of common diseases have implicated SNPs not associated with protein coding or known regulatory regions. This phenomenon has been mostly explained by the discovery of lncRNAs and elucidation of their involvement in pathogenesis of many diseases [10,11,83]. Additionally, some of the disease-causing mutations in protein-coding regions are detrimental because they affect

protein regulation by lncRNA. For example, two mutations in MED12 protein have been linked to Opitz-Kaveggia syndrome, which causes physical anomalies and developmental delays. It has been found that although these mutations do not affect MED12 binding to other Mediator subunits, they impair its association with 2 lncRNAs (ncRNA-a1 and ncRNA) [41]. Another example, lncRNA SNRPN, is involved in Prader Willi syndrome due to its unique feature, namely two large clusters of snoRNAs (SNORD115 and SNORD116). These clusters are predominantly expressed in the brain and, in the case of SNORD116, contribute to many of the PWS symptoms [17]. Further examples of ncRNAs likely associated with rare genetic disorders have been recently reviewed [59,84].

4.2. Use of lncRNA for gene upregulation

One of the main advantages of lncRNAs as therapeutic targets is that they allow manipulation of proteins previously considered undruggable. The largest portion of undruggable targets are proteins upregulation/activation of which would be beneficial for disease treatment. Historically, efforts to create small molecule drugs to induce protein upregulation/activation were not very successful. In case of lncRNA targets, though, their inhibition, easily achieved through oligonucleotide-based drugs, can lead to the desired gene upregulation. Significant advances in oligonucleotide technologies in recent years have largely solved the delivery problems which plagued early efforts in oligonucleotide drug development. As a reflection of these achievements, the number of clinical trials involving oligonucleotides has considerably increased in recent years [85].

Another problem which may be solved by lncRNA-targeted drugs is poor specificity of small molecules, especially in cases of proteins with multiple closely related family members, for example, ion channels or protein kinases. The high target specificity of NATs, that normally regulate one gene or a small group of related genes, and of the oligonucleotide-based drugs, which can be used for NAT modulation, makes them highly suitable for therapeutic intervention in these cases. A growing number of NATs in disease-relevant loci adds further value to the approach [59]. Because most NATs act as gene repressors, their oligonucleotide-mediated knockdown usually increases the expression of specific genes in a highly targeted manner [11]. Another approach to lncRNA-mediated gene upregulation is blocking lncRNA sites of interaction with PRC2 using synthetic oligonucleotides [45]. These two strategies are being put into action by two recently formed biotechnology companies, OPKO CURNA (Miramar, FL) and RaNA (Cambridge, MA), respectively. Significant interest also exists in modulating lncRNA function using small molecules [86].

4.3. Disease-specific applications of lncRNAs

4.3.1. Cancer and lncRNA—Participation of lncRNA in chromosome maintenance, transcription regulation and mRNA and protein quality control mechanisms described above, as well as their newly defined roles in apoptosis and senescence, ensure their prominent effect on cancer development and treatment [reviewed in 72,87]. A large proportion of vlnRNAs serve as hallmarks of pluripotency and cancer [88]. lncRNAs transcribed from non-coding ultraconserved regions are involved in many biological functions including hypoxia response [89] and regulation of pri-miRNA processing [75] and

are also being studied as therapeutic targets and diagnostic tools in cancer. The current cancer applications are mostly based on manipulating lincRNAs [reviewed in 90].

4.3.2. Rare diseases and lncRNA—Rare monogenic diseases probably represent the optimum site of lncRNA-mediated therapeutic intervention. Many of the known disease-causing mutations result in decreased protein expression and/or activity and can be treated through upregulation of the affected protein. Additionally, small numbers of patients affected by rare disorders make drug development for them even less attractive for the pharmaceutical companies using conventional random small molecule screening approaches. At the same time, NAT-targeted oligonucleotide-based drugs can be identified after less extensive screening, and induce highly specific protein upregulation. Nonsurprisingly, the current gene upregulation targets of RaNA include spinal muscular atrophy and Freidrich's ataxia, while OPKO works on Dravet syndrome, a childhood epilepsy caused by SCN1A haploinsufficiency.

4.3.3. Infectious diseases and lncRNA—Endogenous and exogenous viral-encoded lncRNAs have been shown to regulate both host and pathogen gene expression which makes them attractive therapeutic targets in the treatment of infectious diseases [91,92]. In a recent example, variations in lncRNA LOC284889 and MIF-794CATT repeats were associated with malaria susceptibility in Indian populations [93]. Furthermore, lncRNA NeST was shown to bind WDR5, a component of the histone H3 lysine 4 methyltransferase complex, and regulate IFN- γ expression. NeST alone was sufficient to determine mouse strain susceptibility to viral and bacterial pathogens [94].

4.3.3. Common diseases and lncRNA—Application of lncRNA-based treatments in common diseases is currently thwarted by the poor knowledge of their genetic determinants. However, with advances in genetics this issue becomes less and less prominent. GWAS studies, in particular, highlight the roles of lncRNAs in development of such common diseases as diabetes, atherosclerosis, Alzheimer's, schizophrenia, cardiovascular diseases [83,95,96]. For example, 17 of the 23 SNPs already implicated in type 1 diabetes were observed within a long non-coding RNA region in mouse [97]. Furthermore, one of the SNPs strongly associated with Parkinson's, rs356219, is located in the transcription factor YY1 binding site, which regulates expression of RP11-115D19.1, an antisense ncRNA in SNCA 3'-flanking region, which in turn negatively regulates SNCA expression [98]. ANRIL may be involved in development of atherosclerosis because the Chr9p21 atherosclerosis risk allele is associated with increased ANRIL expression [14]. However extensive functional studies and genetic stratification of patients are needed before successful lncRNA-based treatments are possible in common diseases.

4.4. lncRNA for diagnostics

Presence of lncRNA in biological fluids and their deep involvement in disease pathogenesis described above makes them ideal candidates for the development of efficient diagnostic assays, especially in neurological and neurodegenerative diseases where disease site is largely inaccessible. For example, Alzheimer's disease-specific neural transcripts were present at increased levels in the blood of affected individuals [99]. Furthermore, Liu et al.

[78] have identified several lncRNA in the blood which were differentially expressed in subjects with major depressive disorder compared to healthy controls and could be used for diagnostics. In other examples, lncRNA HOTAIR was shown to play a role in cancer metastasis and may be a marker of poor prognosis in patients with primary breast cancer. Levels of lncRNA MALAT1 were significantly higher in metastasizing adenocarcinomas compared to non-metastasizing ones [reviewed in 100]. These findings have inspired the development of a diagnostic assay for prostate cancer by Gen-Probe/Hologic (Bedford, MA).

5. Conclusions

Although recent years brought significant progress in the understanding of the biology of lncRNA, and even some initial developments in their therapeutic application, the lncRNA field is still in its infancy. Partially the progress is impeded by lack of efficient methodologies for studying RNA secondary and tertiary structure and nuclear ultrastructure. More emphasis is needed on studies of regulation of lncRNA expression, their isoform composition and developmental- and tissue-specificity. However, even with the current limited knowledge, the importance of lncRNA studies for the understanding of cell biology and mechanisms of disease is obvious and should fuel the active development of the field.

References

1. Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J, Suzuki H, Carninci P, Hayashizaki Y, Wells C, Frith M, Ravasi T, Pang KC, Hallinan J, Mattick J, Hume DA, Lipovich L, Batalov S, Engström PG, Mizuno Y, Faghihi MA, Sandelin A, Chalk AM, Mottagui-Tabar S, Liang Z, Lenhard B, Wahlestedt C. RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium. Antisense transcription in the mammalian transcriptome. *Science*. Sep 2; 2005 309(5740):1564–6. PMID: 16141073. [PubMed: 16141073]
2. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigó R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res*. 2012; 22(9):1775–89. PMID: 22955988. [PubMed: 22955988]
3. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest AR, Zavolan M, Davis MJ, Wilming LG, Aidinis V, Allen JE, Ambesi-Impiombato A, Apweiler R, Aturaliya RN, Bailey TL, Bansal M, Baxter L, Beisel KW, Bersano T, Bono H, Chalk AM, Chiu KP, Choudhary V, Christoffels A, Clutterbuck DR, Crowe ML, Dalla E, Dalrymple BP, de Bono B, Della Gatta G, di Bernardo D, Down T, Engstrom P, Fagiolini M, Faulkner G, Fletcher CF, Fukushima T, Furuno M, Futaki S, Gariboldi M, Georgii-Hemming P, Gingeras TR, Gojobori T, Green RE, Gustincich S, Harbers M, Hayashi Y, Hensch TK, Hirokawa N, Hill D, Huminiecki L, Iacono M, Ikeo K, Iwama A, Ishikawa T, Jakt M, Kanapin A, Katoh M, Kawasawa Y, Kelso J, Kitamura H, Kitano H, Kollias G, Krishnan SP, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau LF, Lazarevic D, Lipovich L, Liu J, Liuni S, McWilliam S, Madan Babu M, Madera M, Marchionni L, Matsuda H, Matsuzawa S, Miki H, Mignone F, Miyake S, Morris K, Mottagui-Tabar S, Mulder N, Nakano N, Nakauchi H, Ng P, Nilsson R, Nishiguchi S, Nishikawa S, Nori F, Ohara O, Okazaki Y, Orlando V, Pang KC, Pavan WJ, Pavesi G, Pesole G, Petrovsky N, Piazza S, Reed J, Reid JF, Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C, Schönbach C, Sekiguchi K, Semple CA, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H, Shimada K, Silva D, Sinclair B, Sperling S, Stupka E, Sugiura K, Sultana R, Takenaka Y, Taki K, Tammoja K, Tan SL, Tang S, Taylor MS, Tegner J,

Teichmann SA, Ueda HR, van Nimwegen E, Verardo R, Wei CL, Yagi K, Yamanishi H, Zabarovsky E, Zhu S, Zimmer A, Hide W, Bult C, Grimmond SM, Teasdale RD, Liu ET, Brusic V, Quackenbush J, Wahlestedt C, Mattick JS, Hume DA, Kai C, Sasaki D, Tomaru Y, Fukuda S, Kanamori-Katayama M, Suzuki M, Aoki J, Arakawa T, Iida J, Imamura K, Itoh M, Kato T, Kawaji H, Kawagashira N, Kawashima T, Kojima M, Kondo S, Konno H, Nakano K, Ninomiya N, Nishio T, Okada M, Plessy C, Shibata K, Shiraki T, Suzuki S, Tagami M, Waki K, Watahiki A, Okamura-Oho Y, Suzuki H, Kawai J, Hayashizaki Y, FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). The transcriptional landscape of the mammalian genome. *Science*. Sep 2; 2005 309(5740):1559–63. PMID: 16141072. [PubMed: 16141072]

4. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakraborty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Luo OJ, Park E, Persaud K, Preall JB, Ribeca P, Risk B, Robyr D, Sammeth M, Schaffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Ruan X, Hayashizaki Y, Harrow J, Gerstein M, Hubbard T, Reymond A, Antonarakis SE, Hannon G, Giddings MC, Ruan Y, Wold B, Carninci P, Guigó R, Gingeras TR. Landscape of transcription in human cells. *Nature*. Sep 6; 2012 489(7414):101–8. PMID: 22955620. [PubMed: 22955620]
5. ENCODE Project Consortium. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489(7414):57–74. PMID: 22955616. [PubMed: 22955616]
6. Nakagawa S, Kageyama Y. Nuclear lncRNAs as epigenetic regulators-beyond skepticism. *Biochim Biophys Acta*. Mar; 2014 1839(3):215–22. PMID: 24200874. [PubMed: 24200874]
7. Kapranov P, St Laurent G, Raz T, Ozsolak F, Reynolds CP, Sorensen PH, Reaman G, Milos P, Arcenci RJ, Thompson JF, Triche TJ. The majority of total nuclear-encoded non-ribosomal RNA in a human cell is 'dark matter' unannotated RNA. *BMC Biol*. Dec 21.2010 8:149. [PubMed: 21176148]
8. St Laurent G, Wahlestedt C, Kapranov P. The Landscape of long noncoding RNA classification. *Trends Genet*. May; 2015 31(5):239–251. PMID: 25869999. [PubMed: 25869999]
9. Hacisuleyman E, Goff LA, Trapnell C, Williams A, Henao-Mejia J, Sun L, McClanahan P, Hendrickson DG, Sauvageau M, Kelley DR, Morse M, Engreitz J, Lander ES, Guttman M, Lodish HF, Flavell R, Raj A, Rinn JL. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat Struct Mol Biol*. Feb; 2014 21(2):198–206. PMID: 24463464. [PubMed: 24463464]
10. Hangauer MJ, Vaughn IW, McManus MT. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet*. Jun.2013 9(6):e1003569. PMID: 23818866. [PubMed: 23818866]
11. Wahlestedt C. Targeting long non-coding RNA to therapeutically upregulate gene expression. *Nat Rev Drug Discov*. Jun; 2013 12(6):433–46. PMID: 23722346. [PubMed: 23722346]
12. Wood EJ, Chin-Inmanu K, Jia H, Lipovich L. Sense-antisense gene pairs: sequence, transcription, and structure are not conserved between human and mouse. *Front Genet*. Sep 26.2013 4:183. PMID: 24133500. [PubMed: 24133500]
13. Kotake Y, Nakagawa T, Kitagawa K, Suzuki S, Liu N, Kitagawa M, Xiong Y. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene*. Apr 21; 2011 30(16):1956–62. PMID: 21151178. [PubMed: 21151178]
14. Holdt LM, Hoffmann S, Sass K, Langenberger D, Scholz M, Krohn K, Finstermeier K, Stahringer A, Wilfert W, Beutner F, Gielen S, Schuler G, Gäbel G, Bergert H, Bechmann I, Stadler PF, Thierry J, Teupser D. Alu elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. *PLoS Genet*. 2013; 9(7):e1003588. PMID: 23861667. [PubMed: 23861667]

15. Vance KW, Sansom SN, Lee S, Chalei V, Kong L, Cooper SE, Oliver PL, Ponting CP. The long non-coding RNA Paupar regulates the expression of both local and distal genes. *EMBO J.* Feb 18; 2014 33(4):296–311. PMID: 24488179. [PubMed: 24488179]
16. Lazorthes S, Vallot C, Briois S, Aguirrebengoa M, Thuret JY, St Laurent G, Rougeulle C, Kapranov P, Mann C, Trouche D, Nicolas E. A vlincRNA participates in senescence maintenance by relieving H2AZ-mediated repression at the INK4 locus. *Nat Commun.* Jan 20.2015 6:5971. doi: 10.1038/ncomms6971 PMID: 25601475. [PubMed: 25601475]
17. Plagge A. Non-Coding RNAs at the Gnas and Snrpn-Ube3a Imprinted Gene Loci and Their Involvement in Hereditary Disorders. *Front Genet.* Nov 26.2012 3:264. PMID: 23226156. [PubMed: 23226156]
18. Leygue E. Steroid receptor RNA activator (SRA1): unusual bifaceted gene products with suspected relevance to breast cancer. *Nucl Recept Signal.* Aug 3.2007 5:e006. PMID: 17710122. [PubMed: 17710122]
19. Park C, Yu N, Choi I, Kim W, Lee S. IncRNAator: a comprehensive resource for functional investigation of long non-coding RNAs. *Bioinformatics.* May 9.2014 pii: btu325. PMID: 24813212.
20. Das S, Ghosal S, Sen R, Chakrabarti J. InCeDB: Database of Human Long Noncoding RNA Acting as Competing Endogenous RNA. *PLoS One.* Jun 13.2014 9(6):e98965. PMID: 24926662. [PubMed: 24926662]
21. Frankish A, Harrow J. GENCODE Pseudogenes. *Methods Mol Biol.* 2014; 1167:129–55. PMID: 24823776. [PubMed: 24823776]
22. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell.* Jul 3; 2013 154(1): 26–46. doi: 10.1016/j.cell.2013.06.020. Review. PMID: 23827673. [PubMed: 23827673]
23. Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. Gene duplication as a major force in evolution. *J Genet.* Apr; 2013 92(1):155–61. PMID: 23640422. [PubMed: 23640422]
24. Duret L, Chureau C, Samain S, Weissenbach J, Avner P. The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science.* Jun 16; 2006 312(5780):1653–5. PMID: 16778056. [PubMed: 16778056]
25. Johnson R, Guigó R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA.* Jul; 2014 20(7):959–76. PMID: 24850885. [PubMed: 24850885]
26. Kelley D, Rinn J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* Nov 2.2012 1(11):R10. PMID: 23181609.
27. Park E, Maquat LE. Staufen-mediated mRNA decay. *Wiley Interdiscip Rev RNA.* Jul-Aug;2013 4(4):423–35. PMID: 23681777. [PubMed: 23681777]
28. Scaruffi P. The transcribed-ultraconserved regions: a novel class of long noncoding RNAs involved in cancer susceptibility. *ScientificWorldJournal.* Feb 3.2011 11:340–52. PMID: 21298224. [PubMed: 21298224]
29. Nitsche A, Rose D, Fasold M, Reiche K, Stadler PF. Comparison of splice sites reveals that long noncoding RNAs are evolutionarily well conserved. *RNA.* May; 2015 21(5):801–12. PMID: 25802408. [PubMed: 25802408]
30. He S, Gu W, Li Y, Zhu H. ANRIL/CDKN2B-AS shows two-stage clade-specific evolution and becomes conserved after transposon insertions in simians. *BMC Evol Biol.* Nov 13.2013 13:247. PMID: 24225082. [PubMed: 24225082]
31. de Lange T. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* Sep 15; 2005 19(18):2100–10. PMID: 16166375. [PubMed: 16166375]
32. Porrot A, Feuerhahn S, Lingner J. TERRA-reinforced association of LSD1 with MRE11 promotes processing of uncapped telomeres. *Cell Rep.* Feb 27; 2014 6(4):765–76. PMID: 24529708. [PubMed: 24529708]
33. Webb CJ, Wu Y, Zakian VA. DNA repair at telomeres: keeping the ends intact. *Cold Spring Harb Perspect Biol.* Jun 1.2013 5(6) pii: a012666. PMID: 23732473.
34. Flynn RL, Centore RC, O'Sullivan RJ, Rai R, Tse A, Songyang Z, Chang S, Karlseder J, Zou L. TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature.* Mar 24; 2011 471(7339):532–6. PMID: 21399625. [PubMed: 21399625]

35. Redon S, Reichenbach P, Lingner J. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res. Sep*; 2010 38(17):5797–806. PMID: 20460456. [PubMed: 20460456]
36. Cusanelli E, Romero CA, Chartrand P. Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. *Mol Cell. Sep 26*; 2013 51(6):780–91. PMID: 24074956. [PubMed: 24074956]
37. Yu TY, Kao YW, Lin JJ. Telomeric transcripts stimulate telomere recombination to suppress senescence in cells lacking telomerase. *Proc Natl Acad Sci U S A. Mar 4*; 2014 111(9):3377–82. PMID: 24550456. [PubMed: 24550456]
38. Shibayama Y, Fanucchi S, Magagula L, Mhlanga MM. lncRNA and gene looping: what's the connection? *Transcription. Apr 7*.2014 5 PMID: 24802898.
39. Fanucchi S, Shibayama Y, Burd S, Weinberg MS, Mhlanga MM. Chromosomal contact permits transcription between coregulated genes. *Cell. 2013*; 155:606–20. PMID: 24243018. [PubMed: 24243018]
40. Halley P, Kadakkuzha BM, Faghihi MA, Magistri M, Zeier Z, Khorkova O, Coito C, Hsiao J, Lawrence M, Wahlestedt C. Regulation of the apolipoprotein gene cluster by a long noncoding RNA. *Cell Rep. Jan 16*; 2014 6(1):222–30. PMID: 24388749. [PubMed: 24388749]
41. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature. 2013*; 494:497–501. PMID: 23417068. [PubMed: 23417068]
42. Li AW, Seyoum G, Shiu RP, Murphy PR. Expression of the rat BFGF antisense RNA transcript is tissue-specific and developmentally regulated. *Mol Cell Endocrinol. Apr 19*; 1996 118(1-2):113–23. PMID: 8735597. [PubMed: 8735597]
43. Froberg JE, Yang L, Lee JT. Guided by RNAs: X-inactivation as a model for lncRNA function. *J Mol Biol. Oct 9*; 2013 425(19):3698–706. PMID: 23816838. [PubMed: 23816838]
44. Khalil AM, Wahlestedt C. Epigenetic mechanisms of gene regulation during mammalian spermatogenesis. *Epigenetics. Jan-Feb*;2008 3(1):21–8. PMID: 18416029. [PubMed: 18416029]
45. Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, Song JJ, Kingston RE, Borowsky M, Lee JT. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol Cell. Dec 22*; 2010 40(6):939–53. PMID: 21172659. [PubMed: 21172659]
46. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature. Jan 20*; 2011 469(7330):343–9. PMID: 21248841. [PubMed: 21248841]
47. Hirose T, Virnicchi G, Tanigawa A, Naganuma T, Li R, Kimura H, Yokoi T, Nakagawa S, Bénard M, Fox AH, Pierron G. NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol Biol Cell. Jan*; 2014 25(1):169–83. PMID: 24173718. [PubMed: 24173718]
48. Imamura K, Imamachi N, Akizuki G, Kumakura M, Kawaguchi A, Nagata K, Kato A, Kawaguchi Y, Sato H, Yoneda M, Kai C, Yada T, Suzuki Y, Yamada T, Ozawa T, Kaneki K, Inoue T, Kobayashi M, Kodama T, Wada Y0, Sekimizu K, Akimitsu N. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL expression upon immune stimuli. *Mol Cell. Feb 6*; 2014 53(3):393–406. PMID: 24507715. [PubMed: 24507715]
49. Nishimoto Y, Nakagawa S, Hirose T, Okano HJ, Takao M, Shibata S, Suyama S, Kuwako K, Imai T, Murayama S, Suzuki N, Okano H. The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. *Mol Brain. Jul 8*.2013 6:31. PMID: 23835137. [PubMed: 23835137]
50. Yang L, Lin C, Liu W, Zhang J, Ohgi KA, Grinstein JD, Dorrestein PC, Rosenfeld MG. ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell. Nov 11*; 2011 147(4):773–88. PMID: 22078878. [PubMed: 22078878]
51. Roberts TC, Morris KV, Weinberg MS. Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics. Jan 1*; 2014 9(1):13–20. PMID: 24149621. [PubMed: 24149621]
52. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat Struct Mol Biol. Nov*; 2013 20(11):1250–7. PMID: 24077223. [PubMed: 24077223]

53. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL. Many human large intergenic noncoding RNAs associate with chromatinmodifying complexes and affect gene expression. *Proc Natl Acad Sci U S A*. Jul 14; 2009 106(28):11667–72. PMID: 19571010. [PubMed: 19571010]
54. Faghihi MA, Wahlestedt C. Regulatory roles of natural antisense transcripts. *Nat Rev Mol Cell Biol*. Sep; 2009 10(9):637–43. PMID: 19638999. [PubMed: 19638999]
55. Magistri M, Faghihi MA, St Laurent G 3rd, Wahlestedt C. Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. *Trends Genet*. Aug; 2012 28(8):389–96. PMID: 22541732. [PubMed: 22541732]
56. Peschansky VJ, Wahlestedt C. Non-coding RNAs as direct and indirect modulators of epigenetic regulation. *Epigenetics*. Jan 1; 2014 9(1):3–12. PMID: 24739571. [PubMed: 24739571]
57. Modarresi F, Faghihi MA, Lopez-Toledano MA, Fatemi RP, Magistri M, Brothers SP, van der Brug MP, Wahlestedt C. Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. *Nat Biotechnol*. Mar 25; 2012 30(5):453–9. PMID: 22446693. [PubMed: 22446693]
58. Sopher BL, Ladd PD, Pineda VV, Libby RT, Sunkin SM, Hurley JB, Thienes CP, Gaasterland T, Filippova GN, La Spada AR. CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncoding RNA. *Neuron*. Jun 23; 2011 70(6):1071–84. PMID: 2168959. [PubMed: 21689595]
59. Khorkova O, Myers AJ, Hsiao J, Wahlestedt C. Natural antisense transcripts. *Hum Mol Genet*. May 16.2014 pii: ddu207. PMID: 24838284.
60. Yoon JH, Abdelmohsen K, Gorospe M. Posttranscriptional gene regulation by long noncoding RNA. *J Mol Biol*. Oct 9; 2013 425(19):3723–30. PMID: 23178169. [PubMed: 23178169]
61. Zhang XO, Yin QF, Wang HB, Zhang Y, Chen T, Zheng P, Lu X, Chen LL1, Yang L. Species-specific alternative splicing leads to unique expression of snolncRNAs. *BMC Genomics*. Apr 16.2014 15:287. PMID: 24734784. [PubMed: 24734784]
62. Singh M. Dysregulated A to I RNA editing and non-coding RNAs in neurodegeneration. *Front Genet*. Jan 22.2013 3:326. PMID: 23346095. [PubMed: 23346095]
63. Mourtada-Maarabouni M, Williams GT. Growth arrest on inhibition of nonsense-mediated decay is mediated by noncoding RNA GAS. *Biomed Res Int*. 2013; 2013:358015. PMID: 2431968294. [PubMed: 24319682]
64. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 5'UTRs via Alu elements. *Nature*. Feb 10; 2011 470(7333):284–8. PMID: 21307942. [PubMed: 21307942]
65. Ebralidze AK, Guibal FC, Steidl U, Zhang P, Lee S, Bartholdy B, Jorda MA, Petkova V, Rosenbauer F, Huang G, Dayaram T, Klupp J, O'Brien KB, Will B, Hoogenkamp M, Borden KL, Bonifer C, Tenen DG. PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. *Genes Dev*. Aug 1; 2008 22(15):2085–92. PMID: 18676813. [PubMed: 18676813]
66. Huang J, Zhou N, Watabe K, Lu Z, Wu F, Xu M, Mo YY. Long non-coding RNA UCA1 promotes breast tumor growth by suppression of p27 (Kip1). *Cell Death Dis*. Jan 23.2014 5:e1008. PMID: 24457952. [PubMed: 24457952]
67. Matsui K, Nishizawa M, Ozaki T, Kimura T, Hashimoto I, Yamada M, Kaibori M, Kamiyama Y, Ito S, Okumura T. Natural antisense transcript stabilizes inducible nitric oxide synthase messenger RNA in rat hepatocytes. *Hepatology*. Feb; 2008 47(2):686–97. PMID: 18161049. [PubMed: 18161049]
68. Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, Finch CE, St Laurent G 3rd, Kenny PJ, Wahlestedt C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med*. Jul; 2008 14(7):723–30. PMID: 18587408. [PubMed: 18587408]
69. Yoon JH, Abdelmohsen K, Kim J, Yang X, Martindale JL, Tominaga-Yamanaka K, White EJ, Orjalo AV, Rinn JL, Kreft SG, Wilson GM, Gorospe M. Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat Commun*. 2013; 4:2939. PMID: 24326307. [PubMed: 24326307]

70. Wang K, Liu F, Zhou LY, Long B, Yuan SM, Wang Y, Liu CY, Sun T, Zhang XJ, Li PF. A Long Noncoding RNA, CHRF Regulates Cardiac Hypertrophy by Targeting miR-489. *Circ Res.* Apr 25; 2014 114(9):1377–88. PMID: 24557880. [PubMed: 24557880]
71. Takahashi K, Yan IK, Haga H, Patel T. Modulation of hypoxia-signaling pathways by extracellular linc-RoR. *J Cell Sci.* Apr 1; 2014 127(Pt 7):1585–94. PMID: 24463816. [PubMed: 24463816]
72. Deng G, Sui G. Noncoding RNA in oncogenesis: a new era of identifying key players. *Int J Mol Sci.* Sep 5; 2013 14(9):18319–49. PMID: 24013378. [PubMed: 24013378]
73. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, Tramontano A, Bozzoni I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell.* Oct 14; 2011 147(2):358–69. PMID: 22000014. [PubMed: 22000014]
74. Faghihi MA, Zhang M, Huang J, Modarresi F, Van der Brug MP, Nalls MA, Cookson MR, St-Laurent G 3rd, Wahlestedt C. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol.* 2010; 11(5):R56. PMID: 20507594. [PubMed: 20507594]
75. Liz J, Portela A, Soler M, Gómez A, Ling H, Michlewski G, Calin GA, Guil S, Esteller M. Regulation of pri-miRNA Processing by a Long Noncoding RNA Transcribed from an Ultraconserved Region. *Mol Cell.* Jul 3; 2014 55(1):138–47. PMID: 24910097. [PubMed: 24910097]
76. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, Sasaki H. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature.* May 22; 2008 453(7194):539–43. PMID: 18404146. [PubMed: 18404146]
77. Kawaji H, Nakamura M, Takahashi Y, Sandelin A, Katayama S, Fukuda S, Daub CO, Kai C, Kawai J, Yasuda J, Carninci P, Hayashizaki Y. Hidden layers of human small RNAs. *BMC Genomics.* Apr 10; 2008 9:157. PMID: 18402656. [PubMed: 18402656]
78. Liu Z, Li X, Sun N, Xu Y, Meng Y, Yang C, Wang Y, Zhang K. Microarray profiling and co-expression network analysis of circulating lncRNAs and mRNAs associated with major depressive disorder. *PLoS One.* Mar 27; 2014 9(3):e93388. PMID: 24676134. [PubMed: 24676134]
79. Takahashi K, Yan IK, Kogure T, Haga H, Patel T. Extracellular vesicle-mediated transfer of long non-coding RNA ROR modulates chemosensitivity in human hepatocellular cancer. *EBS Open Bio.* May 9; 2014 4:458–67. PMID: 24918061.
80. Jiang Q, Wang J, Wang Y, Ma R, Wu X, Li Y. TF2LncRNA: Identifying Common Transcription Factors for a List of lncRNA Genes from ChIP-Seq Data. *Biomed Res Int.* 2014; 2014:317642. PMID: 24729968. [PubMed: 24729968]
81. Kang MJ, Abdelmohsen K, Hutchison ER, Mitchell SJ, Grammatikakis I, Guo R, Noh JH, Martindale JL, Yang X, Lee EK, Faghihi MA, Wahlestedt C, Troncoso JC, Pletnikova O, Perrone-Bizzozero N, Resnick SM, de Cabo R, Mattson MP, Gorospe M. HuD Regulates Coding and Noncoding RNA to Induce APP→A β Processing. *Cell Rep.* Jun 12; 2014 7(5):1401–9. PMID: 24857657. [PubMed: 24857657]
82. Zhang Z, Zhu Z, Watabe K, Zhang X, Bai C, Xu M, Wu F, Mo YY. Negative regulation of lncRNA GAS5 by miR-21. *Cell Death Differ.* Nov; 2013 20(11):1558–68. PMID: 23933812. [PubMed: 23933812]
83. St Laurent G, Vyatkin Y, Kapranov P. Dark matter RNA illuminates the puzzle of genome-wide association studies. *BMC Med.* Jun 12; 2014 12:97. PMID: 24924000. [PubMed: 24924000]
84. van Devondervoort II, Gordebeke PM, Khoshab N, Tiesinga PH, Buitelaar JK, Kozicz T, Aschrafi A, Glennon JC. Long non-coding RNAs in neurodevelopmental disorders. *Front Mol Neurosci.* Dec 30; 2013 6:53. PMID: 24415997. [PubMed: 24415997]
85. Khorkova O, Hsiao J, Wahlestedt C. Oligonucleotides for upregulating gene expression. *Pharm Pat Anal.* Mar; 2013 2(2):215–29. PMID: 24237027. [PubMed: 24237027]
86. Fatemi RP, Velmshch D, Faghihi MA. De-repressing lncRNA-Targeted Genes to Upregulate Gene Expression: Focus on Small Molecule Therapeutics. *Mol Ther Nucleic Acids.* Nov 18; 2014 3:e196. PMID: 25405465. [PubMed: 25405465]
87. Rossi MN, Antonangeli F. lncRNAs: New Players in Apoptosis Control. *Int J Cell Biol.* 2014; 2014:473857. PMID: 24627686. [PubMed: 24627686]

88. St Laurent G, Shtokalo D, Dong B, Tackett MR, Fan X, Lazorthes S, Nicolas E, Sang N, Triche TJ, McCaffrey TA, Xiao W, Kapranov P. VlinRNAs controlled by retroviral elements are a hallmark of pluripotency and cancer. *Genome Biol.* Jul 22.2013 14(7):R73. PMID: 23876380. [PubMed: 23876380]
89. Ferdin J, Nishida N, Wu X, Nicoloso MS, Shah MY, Devlin C, Ling H, Shimizu M, Kumar K, Cortez MA, Ferracin M, Bi Y, Yang D, Czerniak B, Zhang W, Schmittgen TD, Voorhoeve MP, Reginato MJ, Negrini M, Davuluri RV, Kunej T, Ivan M, Calin GA. HINCUTs in cancer: hypoxia-induced noncoding ultraconserved transcripts. *Cell Death Differ.* Dec; 2013 20(12):1675–87. PMID: 24037088. [PubMed: 24037088]
90. Ling H, Fabbri M, Calin GA. MicroRNAs and other non-coding RNAs as targets for anticancer drug development. *Nat Rev Drug Discov.* Nov; 2013 12(11):847–65. PMID: 24172333. [PubMed: 24172333]
91. Saayman S, Ackley A, Turner AM, Famiglietti M, Bosque A, Clemson M, Planelles V, Morris KV. An HIV-Encoded Antisense Long Noncoding RNA Epigenetically Regulates Viral Transcription. *Mol Ther.* Jun; 2014 22(6):1164–75. PMID: 24576854. [PubMed: 24576854]
92. Gosenca D, Gabriel U, Steidler A, Mayer J, Diem O, Erben P, Fabarius A, Leib- Mösch C, Hofmann WK, Seifarth W. HERV-E-mediated modulation of PLA2G4A transcription in urothelial carcinoma. *PLoS One.* 2012; 7(11):e49341. PMID: 23145155. [PubMed: 23145155]
93. Jha AN, Sundaravadivel P, Pati SS, Patra PK, Thangaraj K. Variations in ncRNA gene LOC284889 and MIF-794CATT repeats are associated with malaria susceptibility in Indian populations. *Malar J.* Sep 25.2013 12:345. PMID: 24066864. [PubMed: 24066864]
94. Gomez JA, Wapinski OL, Yang YW, Bureau JF, Gopinath S, Monack DM, Chang HY, Brahic M, Kirkegaard K. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon- γ locus. *Cell.* Feb 14; 2013 152(4):743–54. PMID: 23415224. [PubMed: 23415224]
95. Szafranski K, Abraham KJ, Mekhail K. Non-coding RNA in neural function, disease, and aging. *Front Genet.* Mar 9.2015 6:87. PMID: 25806046. [PubMed: 25806046]
96. Uchida S, Dimmeler S. Long noncoding RNAs in cardiovascular diseases. *Circ Res.* Feb 13; 2015 116(4):737–50. PMID: 25677520. [PubMed: 25677520]
97. Pundhir S, Hannibal TD, Bang-Berthelsen CH, Wegener AM, Pociot F, Holmberg D, Gorodkin J. Spatially conserved regulatory elements identified within human and mouse Cd247 gene using high-throughput sequencing data from the ENCODE project. *Gene.* Jul 15; 2014 545(1):80–7. PMID: 24797614. [PubMed: 24797614]
98. Mizuta I, Takafuji K, Ando Y, Satake W, Kanagawa M, Kobayashi K, Nagamori S, Shinohara T, Ito C, Yamamoto M, Hattori N, Murata M, Kanai Y, Murayama S, Nakagawa M, Toda T. YY1 binds to α -synuclein 3'-flanking region SNP and stimulates antisense noncoding RNA expression. *J Hum Genet.* Nov; 2013 58(11):711–9. PMID: 24026176. [PubMed: 24026176]
99. Koh W, Pan W, Gawad C, Fan HC, Kerchner GA, Wyss-Coray T, Blumenfeld YJ, El-Sayed YY, Quake SR. Noninvasive in vivo monitoring of tissue-specific global gene expression in humans. *Proc Natl Acad Sci U S A.* May 20; 2014 111(20):7361–6. PMID: 24799715. [PubMed: 24799715]
100. Kim T, Reitmair A. Non-Coding RNAs: Functional Aspects and Diagnostic Utility in Oncology. *Int J Mol Sci.* Mar 1; 2013 14(3):4934–68. PMID: 23455466. [PubMed: 23455466]

Figure 1. lncRNA biology and mechanisms of action

A - Gene duplication and repurposing of pseudogenes is one of proposed routes of lncRNA evolution. **B** - Mobile genetic elements (MGE) frequently initiate formation and evolution of new transcriptional units (TU) some of which become lncRNA. **C** - lncRNA participate in long range DNA looping essential for transcription regulation in some loci. **D** - lncRNA are essential for telomere maintenance. **E** - lncRNA participate in X-chromosome inactivation and imprinting. **F** - lncRNA scaffold and regulate formation of nuclear paraspeckles, as a result also controlling the nuclear-cytoplasmic transport of mRNA. **G** - lncRNA are involved in formation of interchromatin granules enabling pre-mRNA splicing and maturation. **H** - Fine regulation of epigenetic modifications is assisted by tethering of epigenetic effectors and formation of polycomb bodies by lncRNA. **I** - Cytoplasmic roles of

lncRNA include positive and negative regulation of mRNA stability. **J** – lncRNA are involved in regulation of translation and positive and negative regulation of nascent protein stability. **K** - lncRNA act as miRNA sponges blocking miRNA activity.

Figure 2. Therapeutic applications of lncRNA

Dark circles – disease groups; gray boxes – lncRNA functions; white boxes – biotech companies employing the strategy.