Abstract

It is now evident that noncoding RNAs play key roles in regulatory networks determining cell fate and behavior, in a myriad of different conditions, and across all species. Among these noncoding RNAs are short RNAs, such as microRNAs, snoRNAs, and piRNAs, and the functions of those are relatively well understood. Other noncoding RNAs are longer and their modes of action and functions are also increasingly explored and deciphered. Short and long noncoding RNAs interact with each other with reciprocal consequences for their fates and functions. Here, I review the known types of such interactions, discuss their outcomes, and bring representative examples from studies in mammals.

Introduction

Studies profiling transcription on a genome-wide level over the past 15 years showed that regions between protein-coding genes are frequently transcribed into RNA molecules of various lengths. In addition, protein-coding genes are alternatively spliced and produce a variety of isoforms, some of which are unlikely to encode functional proteins. The majority of stable noncoding RNAs that are >200 nt are capped, spliced, and polyadenylated, and are collectively called long noncoding RNAs (lncRNAs). A minority of lncRNAs are processed into smaller RNAs that carry out relatively well-defined functions in cells, such as microRNAs (miRNAs), piRNAs, siRNAs, and snoRNAs. Other types of small noncoding RNAs, such as snRNAs and tRNAs are typically transcribed independently. Small RNAs are usually recognized on the basis of specific sequences and RNA structures by various proteins and form ribonucleoprotein (RNP) complexes. Many of these RNPs are then guided by the small RNA to other RNAs that carry short regions of sequence complementarity. For example, miRNAs are loaded into a RISC complex where they are bound by Argonaute (Ago) proteins and guide RISC to RNAs containing seed matches (defined by complementarity to positions 2–8 of the miRNA), mostly found in the 3' UTRs of protein-coding genes [1]. snoRNAs, on the other hand, guide complexes that deposit RNA modifications, such as 2'O-methylation and pseudouridylation, to specific RNA targets, usually in other noncoding RNAs, such as ribosomal RNAs [2].

The functions of the vast majority of lncRNAs remain unknown, but an increasing number is implicated in a myriad of biological processes [3–6]. Some lncRNAs are differentially expressed or genetically perturbed in a variety of human diseases [7,8], which further increases the interest in understanding lncRNA functions and mechanism of action. It is clear that the currently annotated lncRNAs are composed of a number of families that utilize drastically different
mechanisms, and which currently all bundled together under the “lncRNA” title due to our limited understanding which leads to very poor classification abilities. The common modes of action that were proposed have been reviewed extensively elsewhere [9,10], and include regulation of gene expression in cis and in trans, scaffolding of subcellular domains and complexes, and regulation of protein activity and abundance. Here I focus on the interface between lncRNAs and small RNAs, and the implications of the interactions between them on their functions. Most of the examples I will present come from mammalian cells, but the principles are likely applicable to other eukaryotic species as well, as while IncRNAs evolve fast, IncRNA features are largely similar in the species that have been profiled [11]. As known interactions between various ncRNAs have been recently quite exhaustively listed elsewhere [12], I will focus here on the general principles and possible outcomes of those interactions (Figure 1) and will not attempt to cover all reported examples.

Long noncoding RNAs as precursors for small RNAs

Short RNAs, including miRNAs and snoRNAs, are in many cases produced from introns or exons of longer “hosts”. Some of these hosts are protein-coding genes, but many are IncRNAs. If the small RNA is processed from exonic sequence of those hosts, the processing reaction typically exposes free RNA ends that lead to rapid exonucleolytic degradation of the host. When the small RNA is excised from an intron, the host RNA stability is typically not affected. Recent studies have assigned small RNA-independent functions for hosts of snoRNAs [13–16] and miRNAs [17–19]. In some cases, like H19, the IncRNA function was described before a miRNA was discovered to be encoded by the IncRNA locus [20–22]. H19 is also an intriguing example in which regulated processing of the host results in different relative abundances of the host RNA and the encoded small RNA in different cells [22]. In other cases, specific cellular decay pathways target the hosts and limit their accumulation. For example, nonsense mediated decay (NMD) [23] was shown to preferentially degrade snoRNA host genes in the cytoplasm [24].

Piwi-interacting RNAs (piRNAs), small RNAs expressed primarily in the germline, are also produced in many cases from IncRNA precursors [25,26]. In some cases these IncRNAs are expressed in tissues where the Piwi pathway is not active without being processed into piRNAs. It is thus likely that IncRNAs that are processed into small RNAs sometimes function independently as IncRNAs, with the processing event mediating their stability, and potentially offering an opportunity for post-transcriptional regulation of IncRNA accumulation.

Small RNAs related to termini of long noncoding RNAs

The vast majority of IncRNAs are capped at their 5' end and polyadenylated at their 3' end, by the same complexes and proteins that process mRNAs. However, there are notable exceptions where the termini of IncRNAs are specified and/or stabilized by pathways that typically produce small RNAs. For example, the 3' ends of the MALAT1 IncRNA and of the long isoform of the NEAT1 IncRNA are formed by cleavage by RNase P, that is typically processing the 5' ends of tRNAs [27]. This cleavage also specifies the 5' end of tRNA-like small RNAs (called mascRNA in the case of MALAT1), whose functions remain unknown. The 3' end of MALAT1 and NEAT1
is then stabilized by triple–helical structures that include a short genomically encoded poly(A) tail [28,29].

As mentioned above, snoRNAs are occasionally encoded in introns of lncRNAs. In most cases, the intron host of the snoRNA is rapidly degraded from both ends, and the snoRNA is stabilized by proteins that form the snoRNP complex. In some cases, however, a single intron can encode two snoRNAs, and following degradation, lncRNAs with snoRNAs in both ends are formed, denoted sno-IncRNAs [30]. These IncRNAs are stable, accumulate in the nucleus and can regulate alternative splicing globally by binding splicing regulators [30]. Interestingly, most of these cases occur in the region shown to be critical for the Prader-Willi Syndrome (PWS), and may be related to the pathogenesis of this disease. Another sno- IncRNA, called SLERT was recently shown to act in regulation of RNA Polymerase I activity through binding DDX21. In the PWS region, there are also transcripts that are 5′ snoRNA capped and 3′ polyadenylated (SPAs). These are formed when an snRNP that protects the 5′ end of the transcript allowing polymerase to continue until a polyadenylation site. These stable transcripts were also shown to bind several splicing regulators and regulate alternative splicing [31].

**Regulation of IncRNAs expression by small RNAs**

As IncRNAs are largely indistinguishable from mRNAs on the molecular level, including a cap, a polyA tail, and introns, it is expected that they would be also regulated by small RNAs in the same way as mRNAs. This indeed appears to be the case, and in some systems, it was shown that such regulation has interesting consequences for the IncRNA.

In *C. elegans*, ALG-1 argonature protein loaded with the let-7 miRNA binds the pri-let-7 precursor and promotes its processing, resulting in a positive feedback loop [32]. Conceptually similarly, in mammals, miR-709 localizes to the nucleus through an unknown mechanism and binds through an extensively complementary sequence to the polycistronic pri-miRNA of miR-15/16 miRNAs, inhibiting its processing [33].

Extensive complementarity between a miRNA and a IncRNA can also result in IncRNA cleavage, as first exemplified by the cleavage of the CDR1as circular RNA by miR-671 [34]. There are also numerous examples of IncRNAs that are targeted by miRNAs through conventional seed sites (recently listed in [12]), though the functional importance of these interactions remains mostly unclear. Mechanistically, regulation of IncRNA by miRNAs presumably occurs through the same pathway that acts on mRNA targets – recruitment of the cytoplasmic deadenylation complexes, followed by decapping and RNA degradation [35]. Two aspects of IncRNA biology may limit the relevance of regulation by miRNAs: IncRNAs are typically more nuclear than mRNAs [36,37], which makes them less accessible to cytoplasmic RISC complexes, and they are somewhat less stable (the observed difference in stability between IncRNAs and mRNAs is variable, largely due to differences in set of considered IncRNAs [38–40]). Less stable RNAs are less susceptible to regulation by miRNAs [41], and so miRNAs may have limited impact on expression levels of IncRNAs as a group.

Small RNAs other than miRNAs can also regulate IncRNAs accumulation. Many IncRNAs are expressed specifically in the testis, in particular in late-stage spermatocytes [26,42,43], where
the piRNA machinery is also active. The expression of hundreds of these lncRNAs is increased by more than twofold in the testis of Piwil1−/− mice which do not express piRNAs [26]. Only a minor fraction of the up-regulated lncRNAs are piRNA precursors, and their sequences match piRNAs antisense sequences, suggesting that piRNAs direct repress some lncRNAs. Similarly, piRNAs were shown to regulate lncRNA expression in flies [44].

**Regulation of small RNA activity by long noncoding RNAs**

While regulation of long RNAs by small RNAs has so far received relatively limited attention in the scientific literature, the reverse activity – regulation of small RNA activity by lncRNAs, has been subject of extensive study, and non-negligible controversy, in the last few years. The main reason for this extensive interest is the relative ease with which one can predict possible interactions between lncRNAs and miRNAs, and the considerable understanding of the functions and targets of individual miRNAs. Therefore, when one is faced with the formidable problem of hypothesizing a mode of action or the regulatory targets of a lncRNA, it is often appealing to propose that the lncRNA regulates a particular pathway through binding and affecting the activity of a miRNA. Regulation of miRNA stability by lncRNAs is also an appealing mode of action, since turnover of miRNAs remains quite poorly understood. On the one hand, miRNAs are typically very stable [45], presumably protected from general RNA decay pathways by the Argonaute proteins. On the other hand, developmental transitions and response to stimuli sometimes result in abrupt down-regulation of some miRNAs [46–48], suggesting active and specific turnover, and making target-dependant decay an attractive possibility.

The interest in lncRNAs as potential “competing endogenous RNAs” (ceRNAs) increased in 2010 following a report from the Pandolfi lab that PTENP1, a transcribed pseudogene of the PTEN tumor suppressor, can compete with PTEN mRNA for binding of miRNAs [49]. This report also began the skepticism about this phenomenon, as PTENP1 is expressed at much lower levels than PTEN [50], and has to compete for miRNA binding not just with PTEN, but also with tens of thousands of other binding sites each miRNA has throughout the transcriptome. Since individual miRNA binding sites confer limited repression, it has been proposed that multiple shared sites result in more efficient crosstalk [51,52], but this does not resolve the stoichiometric concerns about the “ceRNA hypothesis”. Several recent studies used theoretical and experimental tools to try and address the question of what magnitude of changes in abundance of a single RNA species are required for affecting expression of other genes through competition for binding of short RNAs. Jens and Rajewsky [53] estimated ~22,700 binding sites compete for miR-20a binding in unperturbed monocytes. Under these conditions, thousands of new binding sites need to be introduced for meaningfully altering the occupancy of miR-20a on any of its targets. Indeed, artificial “miRNA sponges” introducing such numbers of sites were shown to lead to increases in levels of individual targets without markedly affecting miRNA expression levels [54]. In stark contrast, changes in expression of endogenous genes, in particular the typically lowly expressed lncRNAs, almost never reach levels that are predicted to have regulatory impact via simple competition for binding.

Consistent with these predictions, an experimental study in mouse liver and hepatocytes [55] found that target overexpression that effectively doubles by the number of available binding
sites in the transcriptome is needed for detectable changes in gene expression by competition. Specifically, for miR-122, which is expressed at $1.2 \times 10^5$ copies per cell in the liver, addition of at least $\sim 200,000$ of copies of the AldoA target, which contains three potent binding sites for miR-122, was required for detectable up-regulation of miR-122 targets without affecting miR-122 levels. Similar results were obtained when miR-122 was reduced by $\sim 3$-fold using antagomiRs, suggesting that miRNA levels are less important for the threshold of expression above which competition becomes observable [55]. In an in vivo setting, 20-fold increase in AldoA levels, adding thousands of new binding sites did not have any detectable effect on miR-122 target expression levels [55]. These findings were recently corroborated in a follow-up study using mouse embryonic stem cells (mESCs) as an additional system and testing other microRNAs with different abundance ranges [56], supporting the concept that the threshold above which a ceRNA can start influence abundance of other miRNA targets is determined not by miRNA abundance, but rather by the total number of miRNA binding sites, including low-affinity ones, throughout the transcriptome.

Perhaps the most striking candidate for an endogenous “miRNA sponge” is the CDR1as circular IncRNA that contains >60 binding sites for the miR-7 miRNA, and is expected to be resilient to repression by miR-7 due to its circular structure [57,58]. CDR1as indeed acts as a miR-7 sponge in artificial settings [57,58], but loss of CDR1as in mice is surprisingly associated with decrease, rather than increase, in miR-7 levels in the brain, and with increased levels of miR-7 targets [59], suggesting that even the abundant CDR1as with its dozens of high affinity sites for a single miRNA, likely does not act as a miRNA sponge in the endogenous setting.

Despite the doubt cast on the prevalence of ceRNA activity, there is a rapidly growing number of studies reporting ceRNA effect of individual IncRNAs. Most of these studies are performed in cancer cell lines. For example, as of March 2018, there are at least 17 studies reporting ceRNA activity of PVT1 [60–76], a moderately abundant IncRNA, that is almost exclusively nuclear [77,78], and therefore not expected to effectively bind miRNAs. Strikingly, these studies collectively implicate 10 different miRNAs as being “sponged” by PVT1. The experimental evidence in such studies is typically limited to over-expression of the competitor (which typically pushes its levels way above the physiologically relevant levels), or knockdown followed by qRT-PCR of selected targets, which is typically difficult to interpret, as changes in expression can result from other, miRNA-unrelated effects [53]. The suggested “gold-standard” for proving ceRNA activity has been editing of endogenous miRNA target sites (e.g., using CRISPR/Cas9) [53,56] and comprehensive evaluation of the effect on other targets (e.g., by RNA-seq followed by Sylamer analysis [79] showing specific de-repression of the miRNA targets), but to the best of my knowledge, such experiments have not yet been performed for any ceRNA candidate.

**LncRNAs that degrade miRNAs through extensively complementary binding sites**

One way through which a relatively non-abundant IncRNA can nevertheless regulate the activity of typically more abundant miRNAs is through “special” binding sites, that would trigger miRNA degradation upon binding [80]. Indeed, the first example of IncRNAs acting on a microRNA was the IPS1 IncRNA in plants that binds the phosphate starvation-induced miR-399 through an extensively complementary, yet uncleavable binding site [81]. This activity leads to up-regulation
of PHO2, which is an endogenous target of miR-399. The same mechanism could be used to
design inhibitors for other plant miRNAs [81].

Although animal miRNAs typically do not act through target cleavage and rarely have extensive
complementarity with their targets, there is accumulating evidence that such target sites can
efficiently affect miRNA accumulation in animal cells. In 2010, Phil Zamore and colleagues have
shown through experiments in flies that binding of miRNAs loaded in Ago1 to targets with
extensive sequence complementarity triggers tailing of the miRNA with non-templated
nucleotides (mostly adenines and uridines), miRNA trimming, and eventual miRNA degradation
[82], a phenomenon referred to as target RNA-directed miRNA degradation, or TDMD. Similar
results were shown in HeLa cells in vitro [82]. Artificial constructs containing highly
complementary microRNA binding sites were shown to direct efficient microRNA destruction in
liver cells and mouse neurons, which was also correlated with tailing and trimming of the
microRNA [83,84].

Recent studies have described endogenous targets that cause strong TDMD through
extensively complementary sites. The lab of Alena Shkumatava found that a conserved RNA
region, part of what the *libra* IncRNA in fish and of the 3' UTR of *Nrep* protein-coding gene in
mammals [85], binds and degrades the miR-29b microRNA. This has functional consequences
in vivo, as animal behavior is altered in zebrafish and mouse mutants where this binding site is
lost [86]. A preprint from Bartel lab [87] describes similar activity by the highly conserved *Cyrano
(5IPS-AS1)* IncRNA, which harbors an extensively complementary binding site for the miR-7
microRNA [85] (an additional recent study suggested that Cyrano inhibits miR-7 also in mESCs
[88], but changes in miR-7 abundance were not demonstrated in those cells).

There appear to be numerous parallels between the TDMD caused by *Nrep* and *Cyrano*. Both
RNAs are quite abundant and predominantly cytoplasmic [86–88] and both contain unusually
complementary sites - the highly conserved region in *Cyrano* contains an 8mer pairing to the 5'
of miR-7 and another 13 bases pairing to its 3', thus pairing with all bases of miR-7 except 9 and
10 (Figure 2) [85,87]. *Nrep*/*libra* conserved region pairs with 11 bases at the 5' end of miR-29
and nine bases at the 3' end, thus binding all bases of miR-29b except 12–14 (Figure 2). These
binding sites lead to very efficient microRNA degradation. Scrambling of the miR-29b binding
site in *Nrep* leads to a sharp increase in miR-29b levels in the cerebellar granule cell layer in
mice, and ~5-fold increase of miR-29b in *in vitro* differentiated neuronal progenitors [86]. Loss of
*Cyrano* or small changes in the seed of the miR-7 binding site leads to a >40-fold increase in
mature miR-7 levels in the mouse cerebellum and appreciable increases in other tissues where
miR-7 is expressed, as well as in cultured neurons from *Cyrano*-deficient animals [87]. In both
cases, pri- or pre-miRNA levels are not affected. The activity of *Cyrano* appears much more
efficient than other described examples of TDMD when copy numbers are considered, with a
single molecule of *Cyrano* accounts for loss of ~17 molecules of miR-7, presumably because of
other elements in this IncRNA, or because of the specific neuronal context in which it is active
[87]. In any case, consistently with previous results [83], a target with an extensively
complementary binding site can cause degradation of multiple miRNA molecules. *Cyrano*
activity is associated with tailing and trimming of miR-7, though tailing does not appear to
contribute to trimming or miR-7 degradation [87]. *Nrep* is required for miR-29b trimming (no
substantial tailing was observed), and it is not clear if this trimming is needed for miR-29b degradation [86].

Interestingly, the main consequence of Cyrano loss is reduction in the levels of Cdr1as, a circular RNA, which as mentioned above harbors a large number of miR-7 sites. We observed a similar reduction with transient knockdown of Cyrano using siRNAs in SH-Y5Y cells (Hezroni and Ulitsky, unpublished results). The mechanism through which this reduction in Cdr1as occurs is still largely unclear, but it appears to involve miR-671 that cleaves Cdr1as [87].

TDMD is also used by some viral ncRNAs, including the *Herpesvirus saimiri* HSUR1 ncRNA and the murine cytomegalovirus (MCMV) m169 mRNA both containing binding sites for miR-27 that trigger miRNA degradation [89–91]. These degradation processes also trigger trimming and tailing of miR-27. Similarly, the human CMV UL144-145 transcript causes degradation of miR-17 and miR-20a though an extensively complementary binding site ([Figure]) [92].

The prevalence of TDMD by endogenous mammalian IncRNAs remains unclear, but its likely rare, as efficient TDMD requires both seed and extensive 3’ complementarity [83,87] which is exceedingly rare. The endogenous transcripts shown to cause TDMD indeed both have highly highly conserved binding sites, with perfect conservation of at least 8 bases complementary to the 3’ end of the microRNA and extensive sequence conservation outside of the miRNA binding site. Further, TDMD appears to be much more efficient in primary neurons than in other cell types [83,87]. Therefore, the vast majority of typical miRNA binding sites in IncRNAs are not expected to trigger TDMD.

**Conclusions and Future prospects**

As miRNAs and related small RNAs are already known to act in virtually every biological process in mammalian cells, and the spread of IncRNA influence of is also increasing, it is likely that we will also see a dramatic increase in the known interactions between members of these two RNA classes. As IncRNAs are in general very similar in their structure and modifications to mRNAs, the modes and outcomes of their interactions with small RNAs also resemble those already seen with mRNAs, and indeed, none of the examples presented here, be it TDMD or cleavage by piRNAs appear to be unique to IncRNAs. As mentioned above, IncRNAs and mRNAs differ in their average abundance, stability, and localization, and these properties may affect the prevalence of their interactions with small RNAs, but it is important to keep in mind that there are thousands of IncRNAs that closely resemble mRNAs in each of those properties. Thus, the small-long RNA network, that is just now beginning to be uncovered, is expected to remain vibrant and fertile ground for future discoveries, and potentially even therapeutic interventions in a wide array of contexts.
Figures

IncRNA as small RNA precursor

Proteins recognizing the small RNA stabilize the IncRNA

small RNA regulates IncRNA expression

IncRNA regulates small RNA stability

Figure 1. Modes of possible interactions between small and long RNAs.
Figure 2. Base pairing patterns between targets (top) and miRNA (bottom) that result in TDMD. The seed pairing is highlighted in bold.
References


splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res.* **22**, 1616–1625.


