Decrypting noncoding RNA interactions, structures, and functional networks

Muller Fabbri,1,6 Leonard Girnita,2,6 Gabriele Varani,3 and George A. Calin4,5

The world of noncoding RNAs (ncRNAs) is composed of an enormous and growing number of transcripts, ranging in length from tens of bases to tens of kilobases, involved in all biological processes and altered in expression and/or function in many types of human disorders. The premise of this review is the concept that ncRNAs, like many large proteins, have a multidomain architecture that organizes them spatially and functionally. As ncRNAs are beginning to be imprecisely classified into functional families, we review here how their structural properties might inform their functions with focus on structural architecture–function relationships. We will describe the properties of “interactor elements” (IEs) involved in direct physical interaction with nucleic acids, proteins, or lipids and of “structural elements” (SEs) directing their wiring within the “ncRNA interactor networks” through the emergence of secondary and/or tertiary structures. We suggest that spectrums of “letters” (ncRNA elements) are assembled into “words” (ncRNA domains) that are further organized into “phrases” (complete ncRNA structures) with functional meaning (signaling output) through complex “sentences” (the ncRNA interactor networks). This semiotic analogy can guide the exploitation of ncRNAs as new therapeutic targets through the development of IE-blockers and/or SE-lockers that will change the interactor partners’ spectrum of proteins, RNAs, DNAs, or lipids and consequently influence disease phenotypes.

A quarter century after the cloning of the first human noncoding RNA (ncRNA), H19 (Zemel et al. 1992), the number of annotated ncRNAs is continuously increasing and greatly exceeds that of protein-coding genes (Iyer et al. 2015; Hon et al. 2017). An even larger set of noncoding transcripts, many of which are primate-specific, still awaits annotation (Necesula et al. 2014; Washietl et al. 2014; Rigoutsos et al. 2017). Over the last decade, advances in bioinformatics and deep sequencing technology have allowed the identification and annotation of tens of thousands of short and long ncRNAs (lncRNAs). These include endogenous microRNAs (miRNAs), small interfering RNAs (endo-siRNAs), PiWI-interacting RNAs (piRNAs), small nuclear RNAs (snRNAs), RNA-derived small RNAs (tsRNAs), natural antisense transcripts (NATs), circular RNAs (circRNAs), long intergenic noncoding RNAs (lincRNAs), enhancer noncoding RNAs (eRNAs), transcribed ultraconserved regions (T-UCRs), or primate-specific pyknon transcripts (Lee et al. 2009; Haussecker et al. 2010; Esteller 2011; Rigoutsos et al. 2017; Smith and Mattick 2017), and more. These discoveries have created a compelling need to understand the structure–function relationships that underlie the biological roles of ncRNAs.

A very well studied class of ncRNAs is the family of small (19- to 24-nucleotide [nt]) miRNAs (Ambros 2003). Mature miRNAs are generated by two sequential enzymatic cleavage reactions from pri-miRNAs, primary transcripts ranging from hundreds to thousands of nucleotides in length through precursor miRNAs (pre-miRNAs), stem-loop structures of 60–110 nt.

Functionally, a miRNA can regulate the expression of protein-coding or noncoding transcripts in a sequence-specific fashion mostly through the complementarity with the miRNA’s specific “seed” sequence (the first 2–8 nt at the 5’ end) (Bartel 2018). As a result of these interactions, mRNA stability and/or translation can be impaired, leading to a reduction in RNA or protein expression levels (Filipowicz et al. 2008). Yet, it is now apparent that the effects of miRNAs on gene expression are more varied than initially proposed (Dragomir et al. 2018). For instance, nuclear miRNAs can regulate transcription by acting at promoters (Hwang et al. 2007). Pri-miRNA processing to miRNA can be controlled by interactions with lncRNAs (Liz et al. 2014) that can also act as miRNA decoys, sequestering miRNAs or reducing their expression levels (Davis et al. 2017; Kleveland et al. 2018) and thus increasing the expression of genes that would otherwise be specifically repressed (Polissen et al. 2010).

LncRNAs (>200 nt in length) have cell-specific expression patterns and are mechanistically involved in many biological processes (Long et al. 2017). The length of lncRNAs, sometimes in the range of tens of kilobases, allows them to fold into potentially complex but poorly understood secondary and three-dimensional (3D) structures. It is generally believed that these structures affect the interaction of lncRNAs with regulatory DNA sequences; other lncRNAs, miRNAs, and messenger RNAs (mRNAs); various types of nuclear proteins, such as transcription factors, histones, or other chromatin-modifying enzymes; and perhaps even phospholipids (Wang and Chang 2011; Lin et al. 2017) and regulate complex regulatory networks composed of DNA, RNA, and proteins. The complexity of these networks allow alterations in lncRNA expression.
levels to affect a broad spectrum of genes via their multiple partners and orchestrate profound phenotypic changes (Wang and Chang 2013; Long et al. 2017). While the modular nature of IncRNAs is widely accepted, its regulatory principles remain largely unknown after >6 yr from the publication of an influential review (Guttman and Rinn 2012).

The full repertoire of ncRNAs and a mechanistic understanding of their functional involvement in the regulation of cellular processes, and by extension in the onset and progression of human disease, remain largely unknown (Kapranov et al. 2007; Cech and Steitz 2014; Ling et al. 2015), as is the molecular and structural basis for their function. We analyze together the short miRNAs and the long IncRNAs, as similar structural principles can be applied to both categories. We propose that two classes of functional elements can be identified in ncRNAs: first, the “interactor elements” (IEs), necessary for direct physical interaction with various partners through base complementarity (with other nucleic acids) and sequence-specific recognition by RNA-binding proteins (RBPs) (Table 1); and, second, the “structural elements” (SEs), governing the emergence of secondary and/or tertiary 3D ncRNA structures, that direct their functional interactions with other cellular partners (Table 1). A noncoding transcript can have multiple IEs and multiple SEs, located either separately or overlapping within the RNA sequence. IEs might be structured as well. For instance, many RBPs bind to double-strand RNA or to secondary structures such as a hairpin, a bulged nucleotide, or even more complex 3D architecture, and these structures generally influence binding affinity and specificity (Helder et al. 2016). IEs and SEs are components of more complex and structured segments of RNA named “domains,” containing only IEs or only SEs or combinations of both (Fig. 1). We anticipate that most of ncRNAs have both IE and SE, one exception being the mature miRNAs that, due to their short size, lack SEs. By analogy with the “universal grammar” (Hauser et al. 2002; Chomsky 2017; Yang et al. 2017; Box 1), we envision that the patterns of nucleotides that we name “elements” are important for the structure (SEs) and function (IEs) of ncRNA, like patterns of letters and words are important for the structure of a language. Such patterns are structured in a genomic syntax composed of ncRNA interactor networks (NINs), which can be targeted in new therapeutic applications (Fig. 2 and Box 1).

### The IEs in ncRNAs

IEs are generally short and include stretches of functional nucleotides

Many well-characterized IEs are located in miRNAs, which follow the paradigm of target recognition through sequence complementarity (base-pairing) to other RNAs (e.g., miRNAs or IncRNAs) or DNAs (e.g., promoters). The archetypal IE is the “seed” sequence, a conserved heptamer including the nucleotides at positions 2–8 at the 5′-end of miRNAs, which provides most of the binding energy between the miRNA and its target (Fig. 1A). Only the “seed” region is perfectly complementary to the target miRNA’s response element on messenger RNAs or IncRNAs, while the remaining base pairs do not necessarily match perfectly. The principles underlying the interaction between the seed regions of miRNAs and their target miRNAs are well covered in other reviews (Bartel 2004, 2009, 2018). Other types of miRNA interactions through

---

**Table 1. Examples of experimentally supported interactor elements and structural elements in noncoding RNAs**

| ncRNA (size) | Function (example) | Interactor elements | Structural elements | References |
|-------------|-------------------|---------------------|---------------------|-----------|
| BCAR4 (multiple transcripts from 610–1314 nt) | Breast cancer metastasis | Interaction with SNIP1 through IE at position 235–288 and with PPP1R10 through second IE at position 991–1044 | Not known | Xing et al. (2014) |
| Bvht (606 nt) | Cardiomyocytes differentiation | Interaction with CNBP/ZNF9 through a G-rich motif (AGIL) | Not known | Xue et al. (2016) |
| miR-16a precursor (89 nt) | Promotes apoptosis; tumor suppressor; the IE with SRSF3 is mutated in CLL | Interaction through the “mature” miR sequence, including the “seed” region, and with the splicing factor SRSF3 through the “CNNC” motif in precursor | Hairpin loop nucleotides | Calin and Croce (2006); Auyeung et al. (2013) |
| miR-21 mature (22 nt) | Block of apoptosis and induction of invasion and metastasis | Interaction with coding and ncRNA targets through the “mature” miR sequence, including the “seed” region, and GU-rich motifs at the position 18–21 interaction with TLR8 (homo) or TLR7 (mouse) | None | Fabbri et al. (2012) |
| miR-328 precursor (75 nt) | Myeloid maturation | An C-rich IE in the precursor interacting with PCBP2, and a distinct IE in “mature” miRNA interaction with the 3′ UTR of PIM1 kinase mRNA | Hairpin loop nucleotides | Eiring et al. (2010) |
| uc.339 (849 nt) | Cell cycle G1/S transition and promotes lung carcinogenesis | Three IEs for miR-339-3p, miR-663b-3p, and miR-95-5p | Two 4-nt-long SEs that impair the binding of miR-339, miR-663b, and miR-95 to uc.339 | Vannini et al. (2017) |
| XIST (19,296 nt) | Inactivation of X Chromosome | Multiple IE sponging various miRNAs, such as an GTACTGT site for miR-101 or a GCACCTT site for miR-106-5p | Conserved A-repeat element | Pintacuda et al. (2017); Cheng et al. (2019); Sun et al. (2019) |

---

Additional information and other examples can be found in recent references (Zampetaki et al. 2018; Qian et al. 2019). (CLL) Chronic lymphocytic leukemia; (CNBP/ZNF9) CCHC-type zinc finger nucleic acid binding protein; (PCBP2) also known as hnRNP-E2) poly(rC) binding protein 2; (PIM1) Pim-1 proto-oncogene, serine/threonine kinase; (PPP1R10) also known as PnU53) protein phosphatase 1 regulatory subunit 10; (SNIP1) Smad nuclear interacting protein 1; (TLR7 or -8) toll-like receptor 7 or 8.

---
specific elements outside the “seed region” have been described, and their functional role investigated as well (Tay et al. 2008).

Single-nucleotide mismatches due to single-nucleotide polymorphisms (SNP) located on either the target miRNA response element (Nicoloso et al. 2010; Saridaki et al. 2014) or on miRNA IEs (Króliczewski et al. 2018) suffice to affect the function of the IE and can have clear biological consequences, such as increased risk for cancer. This is the case for the insulin-like growth factor-1 receptor (IGF1R) that promotes cancer cell growth and survival and is essential for malignant transformation by many classical cancer-related genes, such as $\text{TP53}$ and $\text{BRCA1}$ (Suleymanova et al. 2017; Worrall et al. 2017). The SNP rs28674628, located in the $\text{IGF1R}$ $3'$ UTR, associates with increased cancer risk and earlier age at diagnosis of breast cancer in Jewish Ashkenazi carriers of the 185delAG $\text{BRCA1}$ mutation (Gilam et al. 2013). Functional studies demonstrated that $\text{IGF1R}$ expression is directly down-regulated by the precursor miR-328 harbors three C-rich elements that resemble the inhibitory PCBP2 binding site in the intercistronic mRNA region of the transcription factor $\text{CEBPA}$. Mechanistically, pre-miR-328 directly


Box 1. The ‘universal grammar’ of the ncRNA structure

In analogy with linguistics, we consider the two elements of ncRNAs, IEs and SEs, as analogous to two types of letters, such as consonants and vowels. Just like words are composed of both, structural domains contain both IEs and SEs in various combinations. By not considering SEs in short RNAs such as precursor miRNAs and the interactor miRNAs or long ncRNAs could explain why target prediction programs, developed by many scientists to be based mostly on sequence complementarity, never performed well enough. Diverse combinations of multiple structural domains compose full-length ncRNAs in a similar way that phrases are composed of different words (see the definition of a phrase and the difference with sentence at https://www.eurocentres.com/blog/clause-phrase-sentence-learn-the-difference/). The domains do not have equal influence on the final structure, as those containing more SEs than IEs are expected to participate mostly in the secondary and 3D structure of a ncRNA, like different words have different grammar roles: Without a verb and a subject, no sentence can be assembled, while adjectives, for example, are not mandatory for the sentence structure. The IEs are important for the context of ncRNA interactor networks (NINs). Just like the universal grammar (UG) concept (Chomsky 2005, 2017; Dabrowska 2015) that considers the initial state of language development to be determined by “the genetic endowment,” the ncRNA language assembles according to rules transmitted hereditarily. For example, the initial mutations identified in the essential tumor-suppressor mir-15a and mir-16-1 for the development of CLL were germline, distributed in a family with multiple cases of cancer (Calin et al. 2005), and located within an IE for the splicing factor SRSF3 (Auyeung et al. 2013). As the number of confirmed ncRNAs of any types (including short or long) already is much larger than protein-coding genes (Ezkurdia et al. 2014), we anticipate that the “sentences” (i.e., NINs) will contain more noncoding than coding words. Protein-coding genes are usually more conserved phylogenetically and, therefore, can be considered “archaic” words (archaism); many ncRNAs are recently evolved (as pyknon transcripts are primate- and/or human-specific), and therefore, these could be considered the “neologisms” of the language. These different parts compose a ncRNA UC essential for the construction of the ncRNA functional language. The composition of structural domains from already constructed IEs and SEs can be called a “merge” function, a fundamental operation of structure building in human language (Fig. 2B; Yang et al. 2017). A “merge” (combine) function can be extraordinary conserved through hundreds of millions of years of evolution: The oncogenic THORLNC uc.283 contains an ultraconserved region responsible for the interaction with the IGF2BP1, and this interaction is conserved from zebrafish to humans (Hosono et al. 2017), meaning that the IE and the necessary SE are merged (combined) in both homologous transcripts, although zebrafish is thought to have arisen ~340 Myr ago from a common ancestor with human (Howe et al. 2013). By “merge” operations, the various elements can be combined during evolution of species in homologous transcripts that are analogous to redundant language phrases and sentences that have similar meanings (contribution to the same NINs). An example is provided by uc.339. The miRNA binding elements for mir-339-3p, mir-663b-3p, and mir-95-5p represent IEs, whose disruption prevents the interaction between uc.339 and the three miRNAs, thereby affecting the expression levels of CCNE2. Even if these IEs are intact, a mutation in the TREs modifies the secondary structure of the uc.339 in such a way that the miRNAs cannot interact with their target, even if the IEs are left intact. Therefore, TREs act as bona fide SEs. Let us assume that a hypothetical RBP (RB1P1) recognizes the complex of uc.339 bound to mir-663b-3p. In this case, the combination of IE, SE, and miR-663b-3p bound to the IX would constitute a structural domain (SD1), whose integrity is necessary for RB1P1’s function. Now, let us assume that another hypothetical RBP (RB2P2) recognizes another structural domain (SD2) composed of the IE necessary to interact with mir-95-5p, the TRE (SE) that affects this interaction and the miR-95-5p itself. In this case, a protein working only if RB1P1 and RB2P2 are properly bound to their structural domains could accurately function only by recognizing NIN (composed of SD1 + SD2). Similarly, in linguistics a sentence “5” (NIN5) is composed of phrases (SD), written with words (IE and SE) (Fig. 2B).

The interaction of IEs with proteins can regulate the production of small RNAs

Long ncRNAs such as pri-miRNAs rely on IEs for processing. The murine knockout model of Dleu2, the primary noncoding transcript in which the sequence of mir-15/16 is embedded (Bullrich et al. 2001), showed a more aggressive chronic lymphocytic leukemia (CLL) phenotype compared with the miR-15/16 knockout model, suggesting that Dleu2 participates in the development of CLL, the most common form of leukemia (Klein et al. 2010). Reduced miR-16 expression associated with CLL is typically due to deletions involving both mir-15a and mir-16-1 (Calin et al. 2002; Calin and Croce 2006). However, there are cases of CLL that retain high levels of pri-miRNAs in malignant leukemic cells while carrying a C-to-T germline mutation near the 3’ region of the mir-16-1 hairpin (Calin et al. 2005). This mutation corresponds to the first C in the “CNNC” motif, an IE for the splicing factor SRSF3 (named also Srp20) binding to pre-mir-16-1 (Auyeung et al. 2013). When the mutation occurs, the interaction between the IE and SRSF3 protein is disrupted, pri-miRNA processing is impaired, and the expression of mature miR-16 is reduced, leading to CLL. SRSF3 is also involved in mRNA export from the nucleus and translation (Corbo et al. 2013). Further supporting the importance of this interaction, the SRSF3-binding motif (CNNC) was found downstream from most pre-miRNA hairpins in bilateralian animals. In fact, a miR-16 mutation in the same IE was also found in the mouse synthetic region of New Zealand black mice, a strain that naturally develops CLL at older ages but not in other strains, including its nearest neighbor strain, New...
IEs might even interact with lipids

Recent findings expanded the interactor capacities of IncRNAs through the potential identification of lipid-binding IncRNAs in cancer cells (Lin et al. 2017). Phosphatidylinositol (3,4,5)-trisphosphate (PIP3), abbreviated to PIP3, functions to activate the AKT pleckstrin homology domain and PIP3 at the single-nucleotide level, facilitating AKT–PIP3 interaction and subsequent enzymatic activation. Genetic deletions of the LINC01139 PIP3-binding motif sensitized breast cancer cells to AKT inhibitors, while LINC01139-dependent AKT hyperactivation led to resistance to AKT inhibitors and increased tumorigenesis (Fig. 1C; Lin et al. 2017). It is possible that this IE is in fact a combined IE-SE as this element might have also a defined structure that facilitates the interaction.

IEs signal the interaction of miRNAs with receptors

It was reported that extracellular secreted miRNAs can act as paracrine agonists of Toll-like receptors (TLRs) in immune cells, triggering a TLR-mediated pro-metastatic inflammatory response that leads to tumor growth and metastasis (Fabbi et al. 2012) and increased resistance to cisplatin (Challagundla et al. 2015). Specific sequences were found to be relevant to this process: GU-rich motifs at positions 18–21 (GUUG for miR-21 and GGUU for miR-29a) seem to be important for the activation of the downstream signaling of human TLR8 (or its murine ortholog TL117). The induced NF-kB signaling leads to increased secretion of IL6 and TNF by the TAMS, promoting cancer growth and metastasis (Fabbi et al. 2012). By performing a systematic point mutagenesis of the miRNA GU region, the nucleotide in position 20 was identified as critical for the TLR8 binding activity of both miR-21 and miR-29a, whereas substituting G with U in position 18 of miR-21 increased the activation of TLR8.


cellular vesicles) are able to bind to TLR8 and activate it in a MYD88-dependent manner. The induced NF-kB signaling leads to increased secretion of IL6 and TNF by the TAMS, promoting cancer growth and metastasis (Fabbi et al. 2012). By performing a systematic point mutagenesis of the miRNA GU region, the nucleotide in position 20 was identified as critical for the TLR8 binding activity of both miR-21 and miR-29a, whereas substituting G with U in position 18 of miR-21 increased the activation of TLR8.

Genomic deletions of the LINC01139 PIP3-binding motif sensitized breast cancer cells to AKT inhibitors, while LINC01139-dependent AKT hyperactivation led to resistance to AKT inhibitors and increased tumorigenesis (Fig. 1C; Lin et al. 2017). It is possible that this IE is in fact a combined IE-SE as this element might have also a defined structure that facilitates the interaction.

IEs signal the interaction of miRNAs with receptors

It was reported that extracellular secreted miRNAs can act as paracrine agonists of Toll-like receptors (TLRs) in immune cells, triggering a TLR-mediated pro-metastatic inflammatory response that leads to tumor growth and metastasis (Fabbi et al. 2012) and increased resistance to cisplatin (Challagundla et al. 2015). Specific sequences were found to be relevant to this process: GU-rich motifs at positions 18–21 (GUUG for miR-21 and GGUU for miR-29a) seem to be important for the activation of the downstream signaling of human TLR8 (or its murine ortholog TL117). The induced NF-kB signaling leads to increased secretion of IL6 and TNF by the TAMS, promoting cancer growth and metastasis (Fabbi et al. 2012). By performing a systematic point mutagenesis of the miRNA GU region, the nucleotide in position 20 was identified as critical for the TLR8 binding activity of both miR-21 and miR-29a, whereas substituting G with U in position 18 of miR-21 increased the activation of TLR8.

Genomic deletions of the LINC01139 PIP3-binding motif sensitized breast cancer cells to AKT inhibitors, while LINC01139-dependent AKT hyperactivation led to resistance to AKT inhibitors and increased tumorigenesis (Fig. 1C; Lin et al. 2017). It is possible that this IE is in fact a combined IE-SE as this element might have also a defined structure that facilitates the interaction.

IEs signal the interaction of miRNAs with receptors

It was reported that extracellular secreted miRNAs can act as paracrine agonists of Toll-like receptors (TLRs) in immune cells, triggering a TLR-mediated pro-metastatic inflammatory response that leads to tumor growth and metastasis (Fabbi et al. 2012) and increased resistance to cisplatin (Challagundla et al. 2015). Specific sequences were found to be relevant to this process: GU-rich motifs at positions 18–21 (GUUG for miR-21 and GGUU for miR-29a) seem to be important for the activation of the downstream signaling of human TLR8 (or its murine ortholog TL117). The induced NF-kB signaling leads to increased secretion of IL6 and TNF by the TAMS, promoting cancer growth and metastasis (Fabbi et al. 2012). By performing a systematic point mutagenesis of the miRNA GU region, the nucleotide in position 20 was identified as critical for the TLR8 binding activity of both miR-21 and miR-29a, whereas substituting G with U in position 18 of miR-21 increased the activation of TLR8.
Some IEs serve to ‘signal’ the intracellular location of ncRNAs

Some miRNAs are confined to specific subcellular compartments; for example, miR-29b (but not its close homolog miR-29a) is enriched in the nucleus and contains a particular hexanucleotide terminal motif (5′-AGUGUU) that increases its stability (Hwang et al. 2007). The motif is unique and not found at the 3′-end of other mammalian miRNAs. This observation suggests that there may be a correlation between different miRNA functions and their intracellular localization and that there might be mechanisms controlling the intracellular distribution of some miRNAs. The presence of mature miRNAs (not only the primary miRNA transcripts) in the nucleus explains how specific miRNAs can regulate transcription of some genes at the DNA level. For example, miR-551b-3p binds a complementary sequence on the STAT3 promoter, recruiting RNA polymerase II and the TWIST1 transcription factor to activate STAT3 transcription, therefore directly up-regulating STAT3 expression. This interaction promotes resistance to apoptosis and increases the survival and proliferation of cancer cells in vitro and in vivo (Chaluvally-Raghavan et al. 2016).

Low-specificity versus high-specificity IEs

Most IncRNAs interact with proteins located in the same subcellular compartment (Chen et al. 2016; Carlevaro-Fita and Johnson 2019). Generally, the same IncRNA was reported to interact with various proteins in different cell types; likewise, the same protein was identified to interact with several IncRNAs. Some of these interactions are ubiquitous (e.g., the interaction between the PRCC complex, specifically the EZH2 enzyme, and several IncRNAs) (Cifuentes-Rojas et al. 2014), while others are highly specific for complex, specifically the EZH2 enzyme, and several lncRNAs. The presence of mature miRNAs (not only the primary miRNA transcripts) in the nucleus explains how specific miRNAs can regulate transcription of some genes at the DNA level. For example, miR-551b-3p binds a complementary sequence on the STAT3 promoter, recruiting RNA polymerase II and the TWIST1 transcription factor to activate STAT3 transcription, therefore directly up-regulating STAT3 expression. This interaction promotes resistance to apoptosis and increases the survival and proliferation of cancer cells in vitro and in vivo (Chaluvally-Raghavan et al. 2016).

Identification of protein-binding sites (IEs) within IncRNAs

Chromatin remodeling enzymes such as PRC1, PRC2, and EZH2 (Margozen and Reinberg 2011) and many other proteins associate with ncRNAs, but it is unclear where and how they interact with them. For example, the class of RBP is composed by a mixture of proteins directly binding to RNAs and proteins binding to these RBPs instead of RNAs. Consequently, it is not known whether these ncRNA-RBP complexes interact directly with specific fragments (IEs) or conserved structural domains (SEs) specifically, perhaps by recognizing conserved folded structures, or whether they bind to the full RNA without specificity for a given region (Margozen and Reinberg 2011; Davidovich et al. 2013). RNA secondary structure provides a better understanding of how RBPs and epigenetic enzymes interact with IncRNAs. Advances in UV cross-linking and immunoprecipitation (CLIP) methods (Lee and Ule 2018), as well as several databases, provide faster and more efficient methods to identify proteins likely to be associated with a given RNA in cells (McHugh et al. 2015). For example, the server RAPID (Bellucci et al. 2011) predicts RNA interaction propensities for known RBPs. Protein partners of a given ncRNA can be identified within the ENCODE enhanced CLIP database (eCLIP) as well. In a recent study (A. Jones, G Pisignano, G Varani et al, in prep.), we showed that most changes in in cell SHAPE (the Selective 2′ Hydroxyl Acylation analyzed by Primer Extension method) reactivity profile, compared with in vitro SHAPE, occurred at or near binding sites of proteins identified by eCLIP. By defining the secondary structure and independently folded domains, it is easier to address the question of where these enzyme complexes and proteins bind to ncRNA.

SEs in ncRNAs

Functional SNPs provide evidence that SEs are important in gene regulation

Although SEs in ncRNAs are poorly defined and even less studied, there is growing evidence supporting their functional importance (Table 1; Blythe et al. 2016; Zampetaki et al. 2018). For instance, because the ncRNA XIST has a conserved A-repeat element, it adopts in vivo an inter-repeat structure essential for mediating its gene silencing function (Pintacuda et al. 2017). It was recently reported that the cancer-risk-associated rs6983267 SNP located at the 8q24 amplion induces changes in the secondary structure of the overlapping IncRNA CCAT2 that leads to allele-specific re-programming of cellular energy metabolism (Redis et al. 2016). Secondary structure predictions using the RNAfold webserver anticipated major local structural changes induced near the putative upstream binding sequence of NUDT21 (Cleavage Factor I subunit, CFM25) (Masamha et al. 2014), by a single-nucleotide variation between the G and T alleles. Such changes may plausibly translate
into distinct secondary and perhaps tertiary folds and could explain binding of the G and T alleles with different affinities. In this model, formation of the RNA:protein complex is controlled by the rs6983267 SNP through its effect on the secondary structure of CCAT2. The consequence of the allele-specific interaction between CCAT2, the splicing complex CFIM, and the metabolic enzyme glutaminase (GLS) pre-mRNA appears to be the preferential splicing to the GAC isoform (Redis et al. 2016), the more catalytically active of the two GLS isoforms (termed KGA and GAC, respectively) and therefore more effective at replenishing intermediates of the TCA cycle (Fig. 1D; Cassago et al. 2012; Le et al. 2012). No other IncRNA has been so far proven to interact with two subunits (CFIM25 and CFIM68) of this complex, providing an example of a high-specificity interaction: The G and the T alleles bind to the NUDT21(CFIm25) and CPSF6 (CFIm68) subunits of the CFIm cleavage factor with different affinities, due to the presence of a high-specificity ID containing the rs6983267 site. The CCAT2–CFIM–GLS regulation axis is altered in about two-thirds of colorectal cancers (Redis et al. 2016).

Further supporting the importance of SNPs in the function of IncRNAs, it was reported that lnc13, harboring a celiac disease–associated haplotype block, forms a complex with the RBP HNRNPD and the histone deacetylase HDAC1 that represses the promoters of proinflammatory genes (Castellanos-Rubio et al. 2016). Of note, the lnc13 interaction with HNRNPD depends on the genotype of the SNP rs917997: This interaction is stronger with the “wild-type” CC genotype than with the disease-associated genotype TT, and the IncRNA structure was predicted to be different based on the SNP genotype (Castellanos-Rubio et al. 2016). Furthermore, in basal cell carcinomas, mutations in an enhancer RNA impaired its activity and reduced the host actin-related gene ACTRT1 expression, ultimately leading to aberrant activation of Hedgehog signaling (Bal et al. 2017).

SEs entrap miRNAs within IncRNA 3-dimensional structures

Recently, a new type of RNA:RNA interaction was described involving a T-UCR. The transcript of uc.339 harbors three IEs for miR-339-3p, miR-663b-3p, and miR-95-5p; of note, the uc.339 interaction with these miRNAs leads to up-regulation of their common target CCNE2, the cyclin E2 that plays a role in cell cycle G1/S transition and promotes lung carcinogenesis (Vannini et al. 2017), overriding a canonical competing endogenous RNA (ceRNA)-type of interaction (Poliseno et al. 2010; Cesana et al. 2011) in which overexpression of the IncRNA or of the miRNAs down-regulates the interacting counterpart. Conversely, while up-regulation of uc.339 reduces the levels of miR-339, miR-663b, and miR-95, modulation (both up- and down-) of these three miRNAs does not significantly affect the expression levels of the uc.339 transcript. This system introduces “entrapping,” distinct from ceRNA, because the structural context is important: The interaction between uc.339 and the three miRNAs is modulated by distinct 4-nt sequences in the IncRNA transcript, located outside of the IEs and called trapping related elements (TREs). These sequences represent a particular type of SE. The disruption of the TREs impairs the binding of miR-339, miR-663b, and miR-95 to uc.339 (Vannini et al. 2017). TREs are likely to be present in other IncRNAs and miRNA. For example, the levels of the SNH6-003 IncRNA that interacts with miR-26a/b did not change when Huh7 cells were treated with either miR-26a or miR-26b mimics or inhibitors (Cao et al. 2017). Therefore, their existence needs to be taken into consideration in every RNA:RNA interaction prediction, because mutations or SNPs in TREs might modulate intermolecular interactions even when the RNA:RNA match (the IE) is intact.

RNA structure is highly hierarchical

Studies of RNA structure are facilitated by its highly hierarchical nature: Secondary structure forms first, creating local stem–loops and helices that fold locally and coalesce into higher-ordered structural domains. Tertiary and higher-order interactions generally form only after the secondary structure is established, with only rare cases of secondary structure rearrangements induced by tertiary structure formation (Wu and Tinoco 1998). This hierarchical folding generates architectures composed of structured domains connected to each other perhaps flexibly, which interact with specific protein complexes, other RNAs, or lipids and bring them in proximity to each other and with target chromatin (Guttman and Rinn 2012). In fact, primary-sequence elements that influence the processing of the primary miRNAs (in fact long ncRNAs) have started to be identified (Chang et al. 2015; Roden et al. 2017). Thus, it is reasonable to hypothesize that IncRNAs contain highly structured regions or actual RNA domains that we might call “structural domains,” that are evolutionarily conserved at the secondary and perhaps tertiary structure and would control function (Ulitsky and Bartel 2013). Such domains might harbor one or multiple SEs or both SEs and IEs (Fig. 2A). RBPs can interact with an IE or with a SE or can even recognize a structural domain or a NIN. These individual structural domains, each composed of SEs and IEs, can be assembled in various configurations according to a ncRNA universal grammar (Fig. 2B and Box 1).

Simple sequence motifs often suffice to confer specific recognition features to a ncRNA; for example, N6-methyladenosine alters RNA structure to regulate binding to low-complexity regions in RBPs (Liu et al. 2017b). However, other functional properties of the ncRNA rely on secondary and 3D structure. RNAs belonging to different functional classes possess different degrees of secondary structure, as remarked by high-throughput technologies that couple chemical modification with deep sequencing (Berkhout and van Wamel 2000; Lu et al. 2011, 2016; Chu et al. 2015; Spitalle et al. 2015). Recent studies suggest that the extent of base-pairing in IncRNAs is comparable to that of the ribosome (Batye et al. 2004; Yonath 2005), and reported the observation of distinct structural domains arranged in a manner that approaches the complexity of RNA enzymes (Novikova et al. 2012; Somarowthu et al. 2015; Hawkes et al. 2016; Liu et al. 2017a). However, relatively few IncRNAs have been characterized at the molecular level, besides the longtime-known cases of H19 (Hurst and Smith 1999; Juan et al. 2000) and Xist (Wutz et al. 2002; Pintacuda et al. 2017) and other more recently characterized examples, such as HOTAIR (Somarowthu et al. 2015), COOLAIR (Hawkes et al. 2016), or lincRNA-p21 (Chillon and Pyle 2016) and MALAT1 (Brown et al. 2014). Of note, this conclusion does not collide with the functional implications of structures and their evolutionary significance (Rivas et al. 2017) because, in cells, RBPs such as hnRNPs (Dreyfuss et al. 1993) might keep these RNAs unfolded. Nevertheless, although no correlation can be found between large stretches of predicted secondary structure in a particular ncRNA and its evolutionary conservation (Ulitsky et al. 2011; Ulitsky and Bartel 2013; Ulitsky 2016), RNA structure is likely to be more conserved than sequence. The frequently intimidating size of IncRNAs (several in the range of tens of kilobases) makes the combinatorial problem of finding the true folded structure difficult.
because of sampling (search space is vast), convergence (whether the absolute minimum in the free energy of folding is found), and flexibility (multiple structures might be equally consistent with the experimental data, perhaps because of conformational heterogeneity in the transcript). In addition, folding in cells and in vitro might very well differ because of kinetic constraints on RNA cotranscriptional folding and the presence of RBPs (Leamy et al. 2016). Advances in techniques such as SHAPE (Mustoe et al. 2018) and dimethyl sulfate (DMS) mapping (Novikova et al. 2012; Kwok et al. 2013), psoralen-cross linking (Lu et al. 2016), and high-throughput ligation followed by deep sequencing (Ramani et al. 2015) provide increasingly rapid access to RNA secondary structure (Carlson et al. 2018). In contrast, RNA 3D organization remains very challenging to study and requires methods that are highly specialized and technically demanding and provide very low throughput, such as small angle X-ray scattering (SAXS) (Rambo and Tainer 2013; Bai et al. 2014) or x-ray crystallography and NMR. More specific information can be found in several recent publications (Lin et al. 2019; Qian et al. 2019).

The NIN and the functional roles of IEs and SEs

RNA–RNA interactions through IEs are now widely recognized and conceptually assembled under the nonsynonymous concepts of “ceRNA” and “miRNA sponge,” discussed in several reviews (Tay et al. 2014; Thomson and Dinger 2016). We propose here a general concept of interacter ncRNAs, that we name the noncoding RNA interactor network (NIN), based on the new paradigm that a ncRNA can interact not only with RNAs (as exemplified by ceRNAs or sponges) but also with DNAs, proteins, and lipids, adding complexity and functional versatility to the regulatory networks. Each NIN is composed of one or more ncRNAs, each having IEs and SEs that participate in direct physical interactions (through IEs) or modulate these interactions (through SEs) with multiple partners (Fig. 2A). Further increasing the intricacy of NINs is the finding that some of the IncRNAs code for proteins that also have functional effects. For example, the protein encoded by the WRAP53 gene is overexpressed in a variety of cancers and promotes cellular transformation, while its down-regulation activates the proapoptotic mitochondrial pathway (Mahmoudi et al. 2011). The IncRNA WRAP53 is also a natural TP53 antisense transcript, functionally archetypal of another class of regulatory IncRNAs. WRAP53 transcript regulates endogenous TP53 mRNA levels and increases translation of the TP53 protein by targeting the 5′ untranslated region of TP53 mRNA, thus sensitizing the cells to TP53-dependent apoptosis (Mahmoudi et al. 2009).

Examples of NINs continuously expand (Marín-Béjar et al. 2017). In addition to the miR-328-PCBP2-mRNA of CEBP alpha-3′ UTR of PIM1 kinase mRNA (Eiring et al. 2010) network presented above (through two different IEs: a repeat-rich IE for PCBP2 protein and an antisense complementary IE used for PIM1 recognition), another recent example is provided by the IncRNA BCAR4 (Xing et al. 2014), which promotes cancer cell metastasis by coordinating a noncanonical Hedgehog signal transduction pathway. At the molecular level, BCAR4 is part of a protein complex containing two RBPs (SNIP1 and PPP1R10 [known also as PNUTS]), one kinase (CIT), and one transcription factor (GLI2) coordinating transcription in response to signaling activation. BCAR4 contains at least two IEs, one at position 235–288 for the interaction with SNIP1 and a second at position 991–1044 for the interaction with PPP1R10. In response to the CCL21 cytokine, BCAR4 binds to SNIP1 and PPP1R10 and releases the SNIP1’s inhibition of EP300-dependent histone acetylation, subsequently enabling PPP1R10 recruitment by the IncRNA to bind to H3K18ac and relief of the inhibition of RNA Pol II by PP1 phosphatase activation. This mechanism activates a noncanonical Hedgehog/GLI2 transcriptional program that promotes cell migration (Xing et al. 2014).

Exploiting ncRNA structures for therapeutic developments

The exploding awareness of the role of noncoding genes in normal human physiology and disease such as cancer (Ling et al. 2013; Shah et al. 2016) is prompting renewed interest in the discovery of both oligonucleotide analogs and more drug-like chemicals aiming at targeting RNA in a sequence and/or structure-specific manner (Matsui and Corey 2017). Exploiting RNA as a drug target has been strongly advocated in antiviral and antibacterial research for >25 yr (for review, see Gallego and Varani 2001; Thomas and Hergenröther 2008; Abooul-ela 2010; Guan and Disney 2012) but not yet reduced to practice. Small molecules or peptidic or antisense molecules targeting ncRNAs may provide an alternative approach to oligonucleotides (Gumireddy et al. 2008; Chirayil et al. 2009; Zhang et al. 2010; Velagapudi and Disney 2014; Shortridge et al. 2017; Monroig-Bosque et al. 2018), and several lines of evidence suggest that the challenges, while significant, are not insurmountable (Blount and Breake 2006; Wilson 2014). RNA can be readily selected to recognize small molecules with greater sensitivity to fine details of the chemistry than antibodies (Warner et al. 2014). It can be envisioned that targeting the IEs directly (by sequence complementarity) or indirectly (by affecting the production of the ncRNA harboring the IEs) by using IE-blocker small molecules (IEBs), will have a double effect on the ncRNAs (long or short) that harbor the IEs and also on the interactor molecules (RNAs, DNAs, proteins, or lipids) by making them available to interact with other partners. The same is also true for targeting SEs by using SE-lockers (SELs), which lock the ncRNA structure in a stable conformation by locking the SEs. For instance, IEB can be considered an oligonucleotide antisense or a protein able to “hide” the IE from the interaction with its effector, therefore functioning as an “inverse agonist.” A SEL could be an antisense oligonucleotide or an integrating transposon that disrupts the SE sequence affecting the secondary and tertiary structure of the IncRNA. Therefore, IEBs and SELs include different types of molecules and should not be considered synonyms of “agonist/antagonist” or “ligand/ inhibitors” sensu strictu. Changing the conformation of an IncRNA will indirectly disrupt the interaction through IE(s) and will change the interactor spectrum of proteins, RNAs, DNAs, or lipids as well and affect their poorly understood functions in perhaps unpredictable ways. Of note, by not inducing the degradation of the IncRNA, but only “freezing” the conformation, the SELs will not modify the abundance of the ncRNA but only affect downstream signaling pathways. The effects of SELs could be of higher functional magnitude compared with IEBs, because affecting the structure of a ncRNA would influence a wide spectrum of interactors molecules and not only those involved in a specific interaction (Fig. 2). However, at the moment, these are only hypotheses that await testing and validation in cellular and small animal models.

Concluding perspectives

The number of IncRNAs has long surpassed the number of protein-coding genes, and their functional versatility expands
continuously. A simplifying framework, like “universal grammar” (Chomsky 2005; Dąbrowska 2015; Chomsky 2017) can help us think about how the functional language of ncRNAs is organized and frame the identification of new therapeutic targets within these functional transcripts. Potential therapeutic agents could interact with regions harboring IEs or SEs (the “letters”) by IEs or the structural domains (the “words”) by SEs or affect globally the 3D structure of an ncRNA (the “phrase”) by SEs as well. The ultimate goal is to affect the function of an entire NIN (a “sentence” composed of various phrases producing a common “meaning”) that influences one or multiple cancer hallmarks (the “meanings”). For example, agents (small molecules, oligonucleotides, or peptides) that lock the ncRNA structure in a stable conformation by locking the SEs (the SEs) and indirectly disrupt the interaction through IE(s) with specific DNA elements, other oncogenic RNAs, or proteins would represent a new approach to the therapy of cancer or any disease involving alterations of ncRNAs (Fig. 2). Some advances have already been made: The MALAT1 triplex-helix SE involved in stability was proven to be selectivity targetable with small molecules (Donlic et al. 2018), but this field is just in its infancy. Building on the specific model based on the interaction of TERT RNA and telomerase (Zappulla and Cech 2004), in 2011, Wang and Chang proposed a general model of IncRNAs function as “molecular scaffold” (Wang and Chang 2011). While embracing the specific types of interactions anticipated by those two models, the paradigm we propose here essentially extends beyond these models by clearly distinguishing between sequences that interact with effectors (IEs) and sequences that affect the secondary/tertiary structure of the IncRNA (SEs). The field of ncRNAs has matured well beyond descriptive expression–pheno­typic correlations toward the more laborious, time-consuming, but ultimately most rewarding stage of establishing structure–function–pheno­type correlations. Understanding the structure–function relationship of ncRNAs using their domain architecture as an organizing principle will facilitate the sorting of the very large and growing class of ncRNAs within functional categories and provide a foundation for more comprehensive mechanistic studies, ultimately leading to their exploitation for clinical purposes. The ultimate goal is to affect the function of an entire NIN to modulate normal cellular states and pathological conditions.

Acknowledgments

Due to the vast expansion of the covered topic in the literature, we apologize to the colleagues whose work was not cited. We thank Chunlai Li and Mihnea Dragomir for the comments on the manuscript and Enrique Fuentes-Mattei for the drawings of the Figures. We apologize to the colleagues whose work was not cited. We thank G.A.C. is the Felix L. Haas endowed professor in Basic Science and Enrique Fuentes-Mattei for the drawings of the Figures. We apologize to the colleagues whose work was not cited. We thank for the Visual Impaired, SINF StranCan, King Gustaf V Jubilee Foundation, Stockholm Cancer Society, Stockholm County, and Karolinska Institute. M.F. is supported by the NIH/NCI grants R01CA215753 and R01CA219024 and the Pablove Foundation Accelerator Award.

References

Aboul-ela F. 2010. Strategies for the design of RNA-binding small molecules. Future Med Chem 2: 93–119. doi:10.2217/fmc.09.49
Ambros V. 2003. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell 113: 673–676. doi:10.1016/S0092-8674(03)00428-8
Ayueung VC, Ulitsky I, McGeary SE, Bartel DP. 2013. Beyond secondary structure: Primary-sequence determinants license pri-miRNA hairpins for processing. Cell 152: 844–858. doi:10.1016/j.cell.2013.01.031
Bai Y, Tambe A, Zhou K, Doudna JA. 2014. RNA-guided assembly of Rev-BRE nuclear export complexes. eLife 2: e03656. doi:10.7554/eLife.03656
Bal E, Park HS, Belaid-Choucair E, Kayserili H, Naville M, Madrange M, Chitticiarau E, Hadji-Rabia S, Cagnard N, Kuonen F, et al. 2017. Mutations in AC2RT1 and its enhancer RNA elements lead to aberrant activation of Hedgehog signaling in inherited and sporadic basal cell carcinomas. Nat Med 23: 1226–1233. doi:10.1038/nm.4368
Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281–297. doi:10.1016/S0092-8674(04)00045-5
Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. Cell 138: 216–233. doi:10.1016/j.cell.2009.01.002
Bartel DP. 2018. Metazoan microRNAs. Cell 173: 20–51. doi:10.1016/j.cell.2018.03.006
Batey RT, Gilbert SD, Montange RK. 2004. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. Nature 432: 411–415. doi:10.1038/nature03037
Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D. 2004. Ultraconserved elements in the human genome. Science 304: 1321–1325. doi:10.1126/science.1098119
Bellucci M, Agostini F, Masin M, Tartaglia GG. 2011. Predicting protein associations with long noncoding RNAs. Nucleic Acids Res 39: 1610–1620. doi:10.1093/nar/gkr1029
Berkhout B, van Wamel JL. 2000. The leader of the HIV-1 RNA genome motif is not essential for viral infectivity. J Virol 74: 3128–3131. doi:10.1128/JVI.74.7.3128-3131.2000
Blomlof KF, Breaker RR. 2006. Riboswitches as antibacterial drug targets. Nat Biotechnol 24: 1558–1564. doi:10.1038/nbt1268
Blythe AJ, Fox AH, Bond CS. 2016. The ins and outs of IncRNA structure: how, why and what comes next? Biochim Biophys Acta 1859: 46–58. doi:10.1016/j.bbadg.2015.08.009
Brown JA, Bulkley D, Wang J, Valenstein ML, Yario TA, Steitz TA, Steitz JA. 2004. The universal grammar of noncoding RNA networks. Science 304: 1558–1564. doi:10.1126/science.1098119
Brown JA, Bulkley D, Wang J, Valenstein ML, Yario TA, Steitz TA, Steitz JA. 2014. Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. Nat Struct Mol Biol 21: 633–640. doi:10.1038/nsmb.2844
Bullrich F, Fuji H, Calin G, Tabuchi H, Negishi M, Pekarsky Y, Rassenti L, Alder H, Reed JC, Keating MJ, et al. 2001. Characterization of the 13q14 tumor suppressor locus in CLL: identification of ALT1, an alternative splice variant of the LEU2 gene. Cancer Res 61: 6640–6648.
Calin GA, Croce CM. 2006. Genomics of chronic lymphocytic leukemia microRNAs as new players with clinical significance. Semin Oncol 33: 167–173. doi:10.1053/j.semioncol.2006.01.010
Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Alder H, Rattan S, Keating M, Rai K, et al. 2002. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci USA 99: 15524–15529. doi:10.1073/pnas.242606799
Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visoni R, Sever NI, Fabbri M, et al. 2005. A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 353: 1793–1801. doi:10.1056/NEJMc050995
Calin GA, Liu CG, Ferracin M, Hyslop T, Spizzo R, Sevignani C, Fabbri M, Cimmino A, Lee EJ, Wojcik SE, et al. 2007. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell 12: 215–229. doi:10.1016/j.ccr.2007.07.027
Cao C, Zhang T, Zhang D, Xie L, Zou X, Lei L, Wu D, Liu L. 2017. The long non-coding RNA, SNHG6-603, functions as a competing endogenous RNA to promote the progression of hepatocellular carcinoma. Oncogene 36: 1112–1122. doi:10.1038/onc.2016.278
Carlevaro-Fita J, Johnson R. 2019. Global positioning system: understanding long noncoding RNAs through subcellular localization. Mol Cell 75: 869–883. doi:10.1016/j.molcel.2019.02.008
Liu N, Zhou KI, Parisien M, Dai Q, Diatchenko L, Pan T. 2017b. Large-scale mapping of IncRNA–protein interactions in roX RNAs act together to mediate X chromosome dosage compensation in Drosophila. Mol Cell 51: 156–173. doi:10.1016/j.molcel.2013.07.001

Iyengar N, Niknafs VS, Malik R, Singhal U, Sahu A, Hosono Y, Barnett TR, Prensner JR, Evans JR, Zhao S, et al. 2015. The landscape of long noncoding RNAs in the human transcriptome. Nat Genet 47: 199–208. doi:10.1038/ng.3192

Juan V, Croft C, Wilson C. 2009. Evidence for evolutionally conserved secondary structure in the H19 tumor suppressor RNA. Nucleic Acids Res 28: 1221–1227. doi:10.1093/nar/28.5.1221

Kaparova K, Willingham AT, Gingeras TR. 2007. Genome-wide transcription and the implications for genomic organization. Nucleic Acids Res 35: 413–423. doi:10.1093/nrg/2008

Kleaveland B, Shi CY, Stefano J, Bartel DP. 2018. A network of noncoding regulatory RNAs acts in the mammalian brain. Cell 174: 350–362.e17. doi:10.1016/j.cell.2018.05.022

Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, Ambesi-Impacitto A, Caliano A, Migliazza A, Bhagat G, et al. 2010. The DLEU2/mir-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. Cancer Cell 17: 28–40. doi:10.1016/j.ccc.2009.11.019

Krolczewski J, Sobolewska A, Lejnowski D, Collawn JF, Bartoszewski R. 2018. microRNA single nucleotide polymorphism influences on microRNA expression and mRNA target specificity. Gene 640: 66–72. doi:10.1016/j.gene.2017.10.021

Kwok CK, Ding Y, Tang Y, Assmann SM, Bevilacqua PC. 2013. Determination of in vivo RNA structure in low-abundance transcripts. Nat Commun 4: 2971. doi:10.1038/ncomms3997

Le A, Lane AN, Hamaker M, Bose S, Gouw A, Barbi J, Tsukamoto T, Rojas CJ, Slusher BS, Zhang H, et al. 2012. Glucose-independent glutamine metabolism via TCA cycling for proliferation and survival in B cells. Cell 151: 110–121. doi:10.1016/j.cell.2012.12.009

Leamy KA, Assmann SM, Mathews DH, Bevilacqua PC. 2016. Bridging the gap between in vitro and in vivo RNA folding. Q Rev Biophys 49: e10. doi:10.1017/S003358351600007X

Lee FCY, Ule J. 2018. Advances in CLIP technologies for studies of protein–RNA interactions. Mol Cell 69: 354–369. doi:10.1016/j.molcel.2018.01.005

Lee YS, Shibata Y, Malhotra A, Dutta A. 2009. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes Dev 23: 2639–2649. doi:10.1101/gad.183760

Lien EC, Dibble CC, Toker A. 2017. P53 signaling in cancer: beyond AKT. Curr Opin Cell Biol 45: 62–71. doi:10.1016/j_DOI.2017.02.007

Lin A, Hu Q, LLChung, MG, Wang C, Li J, Ye Y, Yao J, Liang K, et al. 2017. The LINKA IncRNA interacts with PtdIns(3,4,5)P3 to hyperactivate AKT and confer resistance to AKT inhibitors. Nat Cell Biol 19: 238–251. doi:10.1038/ncllb.2016.3473

Lin J, Guo H, Hu Y, Zhang X, Zhang H, Zhu H. 2019. Pipelines for cross-species and genome-wide prediction of long noncoding RNA binding. Nat Protoc 14: 795–818. doi:10.1038/s41596-018-0115-5

Ling H, Fabrò M, Calin GA. 2013. MicroRNAs and other non-coding RNAs as targets for anticancer drug development. Nat Rev Drug Discov 12: 847–863. doi:10.1038/nrd4140

Ling H, Vincent K, Pitcher M, Fodde R, Bertinrand-Neagoe I, Slack FJ, Calin GA. 2015. Junk DNA and the long non-coding RNA twist in cancer genome organization. Nature 521: 232–236. doi:10.1038/nature14443

Lu ZP, Zhang QC, Lee B, Flynn RA, Smith MA, Robinson JT, Davidson C, Gooding AR, Goodrich KJ, Mattick JS, et al. 2016. RNA duplex map in living cells reveals higher-order transcript assembly. Structure 24: 1267–1279. doi:10.1016/j.str.2016.04.028

The universal grammar of noncoding RNA networks
hairpin-containing transcripts and human genetic variation. Genome Res 27: 374–384. doi:10.1101/gr.208900.116
Saridaki Z, Weidhaas JB, Lenz HJ, Laurent-Puig P, Jacobs B, De Schutter J, De Roock W, Salzman DW, Zhang W, Yang D, et al. 2014. A let-7 microRNA-binding site polymorphism in KRAS predicts improved outcome in patients with metastatic colorectal cancer treated with salvage cetuximab/panitumumab monotherapy. Clin Cancer Res 20: 4499–4510. doi:10.1158/1078-0432.CCR-14-0348
Shah MY, Ferrajoli A, Sood AK, Lopez-Berestein G, Calin GA. 2016. microRNA therapeutics in cancer: an emerging concept. Elife 15: e24320. doi:10.1126/sciadv.7b00180
Shah MY, Ferracini M, Pileckzi V, Chen B, Redis R, Fabris L, Jin Z, Yan L, Li Q, Yang J. 2019. Long non-coding RNA XIST regulates miR-106b-5p/P21 axis to suppress tumor progression. Oncogene 38: 5327–5328. doi:10.1038/onc.2016.472
Smith MA, Mattick JS. 2017. Structural and functional annotation of long noncoding RNAs. Methods Mol Biol 1526: 65–85. doi:10.1007/978-1-4939-6613-4_4
Somarowthu S, Legiewicz M, Chillón I, Marcia M, Liu F, Pyle AM. 2015. Ulitsky I. 2017. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. Methods Mol Biol 1508: 353–361. doi:10.1007/978-1-4939-6613-4_4
Sun K, Jia Z, Duan R, Yan Z, Jin Z, Yan L, Li Q, Yang J. 2019. Long non-coding RNA XIST regulates miR-106b-5p/P21 axis to suppress tumor progression in renal cell carcinoma. Biochem Biophys Res Commun 510: 416–420. doi:10.1016/j.bbrc.2019.01.116
Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I. 2008. HOTAIR forms an intricate and modular secondary structure. Nat Struct Mol Biol 15: 353–361. doi:10.1038/nsmb.1426
Tay Y, Rinn J, Pandolfi PP. 2014. The multilayered complexity of cell RNA crossstalk and competition. Nat Cell Biol 16: 1211–1220. doi:10.1038/ncb2986
Thomas JR, Hergenrother PJ. 2008. Targeting RNA with small molecules. Chem Rev 108: 1171–1224. doi:10.1021/cr0681156
Thomson DW, Dinger ME. 2016. Endogenous microRNA sponges: evidence and controversy. Nat Rev Genet 17: 272–283. doi:10.1038/nrg.2016.20
Tsal MC, Manor O, Yan Y, Monomarapasrat N, Wang K, Lan F, Shi Y, Segal E, Chang HY. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. Science 329: 689–693. doi:10.1126/science.1192201
Ulitsky I. 2016. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. Nat Rev Genet 17: 601–614. doi:10.1038/nrg.2016.85
Ulitsky I, Bartel DP. 2013. lincRNAs: genomics, evolution, and mechanisms. Cell 154: 26–46. doi:10.1016/j.cell.2013.06.020
Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 147: 1537–1550. doi:10.1016/j.cell.2011.11.055
Vannini I, Wise PM, Challagundla RK, Plousiou M, Raffini M, Bandini E, Fanini F, Paliga G, Crawford M, Ferracin M, et al. 2017. Transcribed ultraconserved region 339 promotes carcinogenesis by modulating tumor suppressor microRNAs. Nat Commun 8: 1801. doi:10.1038/ncomms16627
Velandapudi SP, Disney MD. 2014. Two-dimensional combinatorial screening enables the bottom-up design of a microRNA-10b inhibitor. Chem Commun (Camb) 50: 3027–3029. doi:10.1039/c3cc0173c
Wang KC, Chang HY. 2011. Molecular mechanisms of long noncoding RNAs. Mol Cell 43: 904–914. doi:10.1016/j.molcel.2011.08.018
Warner KD, Chen MC, Song W, Strack RL, Thorn A, Jaffrey SR, Ferré-D’Amare AR. 2014. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. Nat Struct Mol Biol 21: 658–663. doi:10.1038/nsmb.2865
Washietl S, Kellis M, Garber M. 2014. Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six genomes. Genome Res 24: 616–628. doi:10.1101/gr.165035.113
Wilson DN. 2014. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. Nat Rev Microbiol 12: 85–49. doi:10.1038/nrmicro3155
Worrall C, Suleymanova N, Crudden C, Trocoli Drakensjo I, Candrea E, Nedelev D, Takahashi SI, Girmita L, Girmita A. 2017. Unbalancing p53/Mdm2/IGF-1R axis by Mdm2 activation restrains the IGF-1-dependent invasive phenotype of skin melanoma. Oncogene 36: 3274–3286. doi:10.1038/onc.2016.472
Wu M, Tinoco Jr. 1998. RNA folding causes secondary structure rearrangement of lincRNAs in vertebrate embryonic development despite rapid sequence evolution.