Published in final edited form as:
Nat Rev Genet. 2014 June ; 15(6): 423–437. doi:10.1038/nrg3722.

The rise of regulatory RNA

K.V. Morris[1,2] and **J.S. Mattick**[1,3,4]

[1]School of Biotechnology and Biomedical Sciences, University of New South Wales, Sydney, NSW 2052, Australia

[2]Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA

[3]Garvan Institute of Medical Research, 384 Victoria St. Darlinghurst NSW 2010, Australia

[4]St Vincent's Clinical School, University of New South Wales, Sydney, NSW 2052, Australia

Abstract

Discoveries over the last decade portend a paradigm shift in molecular biology. Evidence suggests that RNA is not only functional as a messenger between DNA and protein but also in the regulation of genome organization and gene expression, which is increasingly elaborated in complex organisms. Regulatory RNAs appear to operate at many levels, but in particular to play an important role in the epigenetic processes that control differentiation and development. These discoveries suggest a central role for RNA in human evolution and ontogeny. Here we survey the emergence of the previously unsuspected world of regulatory RNAs from an historical perspective.

Introduction

RNA has long been at the centre of molecular biology and was likely the primordial molecule of life, encompassing both informational and catalytic functions. It is thought that its informational functions were subsequently devolved to the more stable and easily replicable DNA, and its catalytic functions to the more chemically versatile polypeptides[1]. The idea that the contemporary role of RNA is to function as the intermediary between the two had its roots in the early 1940s with the entry of chemists into biology, notably Beadle and Tatum[2], whose work underpinned the "one gene-one enzyme" hypothesis, which later morphed into the more familiar phrase "one gene-one protein", and gained currency despite the prescient misgivings of experienced geneticists, notably McClintock[3]. The concept that genes encoded (solely) the functional components of cells (the 'enzymes') itself had deeper roots in the mechanical zeitgeist of the era, being decades before the widespread understanding of the use of digital information for systems control.

Although the one gene-one protein hypothesis has long been abandoned, due to the discovery of alternative splicing in the 1970s, the protein-centric view of molecular biology has persisted, aided by phenotypic and ascertainment bias towards protein-coding mutations in genetic studies and by the assumption that regulatory mutations affected cis-acting

The authors declare no conflict of interest.

regulatory protein binding sites[4]. However, this view is challenged by the discovery of nuclear introns and the phenomenon of RNA interference (RNAi), as well as by the advent of high throughput sequencing, which led to the identification of large numbers and different types of large and small RNAs, whose functions are still under exploration.

Here we examine the history and chart the shift in thinking that is still underway about the role of RNA in cell and developmental biology, especially in animals. The emerging evidence suggests that there may be more genes encoding regulatory RNAs than encoding proteins in the human genome, and that the amount and type of gene regulation in complex organisms has been substantially misunderstood for most of the past 50 years.

Early ideas for the role of RNA

RNA, the central dogma and gene regulation

After the elucidation of the double-helical structure of DNA in 1953[5], the following years were preoccupied with deciphering the 'genetic code' and establishing the mechanistic pathway between gene and protein: the identification of a transitory template (mRNA), an adaptor (tRNA) and the 'ribosome' factory comprised of ribosomal proteins and RNA (rRNA) for the translation of the code into a polypeptide. In 1958, Crick published the celebrated 'central dogma' to describe the flow of genetic information (DNA → RNA → protein), which has proved remarkably accurate and durable, including the prediction of reverse transcription[6]. Nonetheless, in conceptual terms, RNA was tacitly consigned to be the template (and infrastructural platform – ribosomal and transfer RNAs) for protein synthesis or, at least, has been interpreted in this way by most people since that time.

The link between rRNA (which is highly expressed in virtually all cells) and the structures termed ribosomes as the platform for protein synthesis was established in the mid 1950s[7]. The roles of tRNA and mRNA were experimentally confirmed in 1958[8] and 1961[9], respectively, the latter the same year that Jacob and Monod published their classic paper on the *lac* operon of *Escherichia coli*[10], the first locus to be characterized at the molecular genetic level. These studies confirmed that (at least some, but presumed most) genes encoded proteins, and supported the emerging idea that gene expression was controlled by regulating the transcription of the gene, as indicated by the locus encoding the *lac* repressor – the repressor-operator model. At the time Jacob and Monod did not know the chemical identity of the repressor, speculating *en passant* that it "may be a polyribonucleotide" (i.e., RNA)[10]. However, Gilbert later showed that the repressor was a polypeptide that allosterically bound the substrate lactose, and the brief idea faded[11].

These studies reinforced and extended the conception that proteins are not only 'enzymes' but also the primary analogue components and control factors that comprise the cellular machinery. This in turn has led to the prevailing 'transcription factor' paradigm of gene regulation, including the derived assumption that combinatoric interactions would provide factorially 'explosive' regulatory combinations[12], more than enough to supervise human ontogeny. However, this assumption has not been substantiated theoretically or mechanistically, and the observed scaling of regulatory genes and the extent of the regulatory challenge in programming human developmental architecture appears to be quite

different from these expectations[13]. In this context it is noteworthy that the genome-wide association studies have shown that most haplotype blocks influencing complex diseases fall outside the known boundaries of protein-coding genes[14].

Small nuclear / spliceosomal RNAs and small nucleolar RNAs

Following the discovery and functional description of tRNAs and rRNAs, other new classes of common small nuclear RNAs were identified by biochemical fractionation[15]. Many of the small RNAs were found to be part of ribonucleoprotein complexes (RNPs) (reviewed in [16]). One class, the small nuclear RNAs (snRNAs, Figure 1), were later found to be central co-factors in RNA splicing (see below)[17], hence their newer designation as 'spliceosomal' RNAs. The snRNAs U1, U2, U4, U5 and U6 participate in a number of RNA-RNA and RNA-protein interactions in the assembly and function of canonical spliceosomes, recognizing the 5′-end splice site (U1) and the branch point (U2), followed by the recruitment of U4, U5 and U6, which displace U1 and interact with U2 (via U6) as well as the 5′ and 3′ splice sites (via U5)[18]. A set of less abundant snRNAs (U11, U12, u4atac and U6atac) along with U5 are found in a variant 'minor' spliceosome, termed U12-type[19].

Other small RNAs were found to be localized to nucleoli and to guide the methylation (the box CD subclass) and pseudouridylation (the box H/ACA subclass) of ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), and snRNAs[20–22] (Figure 1). The chemical modifications of rRNAs, tRNAs and snRNAs proved to be essential in ribosomal and cellular function, in particular tRNA and mRNA maturation and pre-mRNA splicing (U2). Notably, the disruption of snoRNAs was found to lead to a loss of processing of the 18S, 5.8S, and 25S/28S rRNAs[20]. Although some snoRNAs are subject to parental imprinting and/or differentially expressed, for example in the brain[23,24], and appear to target a wider range of RNAs including mRNAs[25], suggesting a regulatory role, and there are related small RNAs (scaRNAs) in other subnuclear structures called Cajal bodies (which also process telomerase RNA)[26], none of these early studies suggested anything other than that the role of RNAs was limited to protein synthesis.

The emergence of heterogeneous nuclear RNAs

The first hint that RNA may have additional roles in complex organisms was the discovery of 'heterogenous' nuclear RNA (hnRNA)[27] and the observation that the complexity of this population, as determined by denaturation-renaturation hybridization kinetics, was much greater in the nucleus than the cytoplasm. The existence of hnRNA and the concomitant discovery of the large amount of 'repetitive' sequences (different classes of retrotransposed sequences with similar composition that occupy large fractions of plant and animal genomes) led Britten & Davidson to speculate in 1969 that animal cells might contain extensive RNA-based regulatory networks[28–30]. While this hypothesis attracted a great deal of interest at the time, it also quickly lapsed, with the proponents not re-visiting it even after the subsequent discovery of introns (see below), instead focussing on regulatory networks controlled by transcription factors[31,32] or the significance of transposons in protein evolution[33].

The discovery of introns

The discovery of introns in 1977[34,35] was perhaps the biggest surprise in the history of molecular biology[36], as no one expected that the genes of higher organisms would be mosaics of coding and noncoding sequences, all of which are transcribed. However, the prevailing conception of the flow of genetic information was not overly disturbed as the removal of the intervening sequences ('introns') and the reconstruction of a mature mRNA by RNA splicing preserved the conceptual status quo. That is, genes still made proteins. In parallel, it was assumed that the excised intronic RNAs were simply degraded, although the technology of the time was too primitive to confirm this. In any case, introns were immediately and universally dismissed as genomic debris, and their presence rationalized as evolutionary remnants involved in the prebiotic modular assembly of protein-coding RNAs that have lingered (and been expanded by transposition) in complex organisms[37]. This was consistent, superficially at least, with the implications of the C-value enigma that eukaryotes contained varying amounts of DNA baggage, and the accompanying conclusion that retrotransposon sequences, often pejoratively referred to 'repeats', which occupy much of the genomic real estate in plants and animals, are largely selfish DNA that are parasitic co-travellers[38,39].

RNA as a catalyst

A few years later, Cech, Altman and colleagues showed that RNA itself was capable of enzymatic catalysis ('ribozymes')[40,41], which provided evidence in support of the 'RNA early' hypothesis, and showed that RNA catalysis exists and has persisted in particular contexts, notably at the core of RNA splicing[42] and mRNA translation[43]. This reinforced the mechanical conception of molecular biology, and the role of RNA as the platform for protein synthesis, but did not give any hint of RNA as a widespread regulatory factor, although that possibility is perfectly feasible. Indeed there is increasing evidence that catalytic RNA exists in animal and plant cells, in introns, UTRs and elsewhere, and may play a variety of roles including, for example, in regulating post-transcriptional cleavage reactions[44,45].

The small RNA revolution

The discovery of microRNAs

In 1993 Ambros and colleagues showed the first evidence for small (~22 nt) regulatory RNAs, by the discovery of the genetic loci *lin-4* and *let-7*, which regulate the timing of *Caenorhabditis elegans* development[46,47]. Although *let-7* is highly conserved from nematodes to humans[48], very few miRNAs have been discovered genetically[49,50], and these RNAs remained interesting idiosyncrasies until the discovery of RNAi (see below), which led to the targeted cloning after size selection of many more[51–53] and the demonstration that these 'microRNAs' (miRNAs) act, at least in part, by imperfect base pairing with (usually) the 3′UTRs of target mRNAs to inhibit their translation and accelerate their degradation[54].

Currently, there are large numbers of evolutionarily widespread miRNAs in the databases[55], almost all of which had evaded prior detection by genetic screens (but many subsequently validated by reverse genetics). While many miRNAs can be identified by conservation, it is

also evident that many are tissue- and lineage-specific[56,57], and that there may be many more to be discovered.

It has also become evident that many if not most protein-coding transcripts are targets for miRNA regulation[58,59], that miRNAs can, in some cases, regulate large numbers of target mRNAs[60], and reciprocally that many mRNAs contain target sites for many miRNAs[61], although the implied regulatory logic of this complex multiplex arrangement has not been explained. The targets of miRNAs are usually thought to be mRNAs, but may also include other RNAs[62]. Biologically, miRNAs have been shown to regulate many physiological, developmental and disease processes, including, for example, pluripotency[63], epithelial-mesenchymal transition and metastasis[64], testis differentiation[65], diabetes[66], and neural plasticity and memory[67], among others[68].

The RNA interference pathway

MicroRNAs are just one facet of the phenomenon of 'RNA interference' (RNAi), which causes silencing of gene expression after the introduction of sense-antisense RNA pairs, discovered in 1998 in plants[69] and *C. elegans*[70]. These discoveries were presaged by the curious phenomenon of transgene silencing, mainly in plants[71,72], linked to antisense RNA and small RNA-directed DNA methylation of transgenes, indicating transcriptional as well as post-transcriptional silencing[73,74]. Mechanistic analysis of these silencing mechanisms showed that exogenous double-stranded RNA was processed into short fragments (short interfering RNAs or siRNAs) with a similar size to miRNAs, suggesting that miRNAs may represent a similar endogenous system.

This was confirmed and led to the elucidation of natural double-stranded precursors in stem-loop structures[75] and the identification of key genes and enzymes involved in their biogenesis and function, notably Drosha[76], Dicer[77] and multiple Argonaute (Ago) proteins[78], the latter of which were already known to play central roles in differentiation and development,[79] but are now known to also be involved in defence against RNA viruses in many organisms[80]. Drosha and Exportin 5 are involved in the cleavage and export of double-stranded RNA (dsRNA) precursors from the nucleus to the cytoplasm[76], where they are further processed by Dicer to small (21–24 nt) dsRNA moieties, one strand of which is loaded into Ago component of the RNA-induced silencing complex (RISC), which also contains other proteins[77]. The RISC is guided by the small RNA strand to complementary RNA targets, which are subsequently silenced by translational repression and/or RNA destabilization[81,82] (Figure 2).

While still under discussion, the current view is that siRNAs (and 'short-hairpin RNAs', shRNAs), which seem to naturally occur more commonly in plants, act primarily by perfect base-pairing and by Ago-mediated cleavage of complementary target RNAs (and hence are used widely as experimental tools and potential therapeutic agents[83]), whereas miRNAs have incomplete homology with their targets and act primarily at the translational level[81,82] (Figure 2).

MiRNAs and siRNAs are thought to act post-transcriptionally and cytoplasmically, but the existence of Ago in the nucleus[84–87] and the role of the RNAi pathway in epigenetic

modulation[88] suggests that the system is more complex and multifaceted than expected, with (for example) demonstrations that miRNA isoforms are developmentally regulated[89] that the target 'seed' sequence is only one factor in target recognition[90,91], and that miRNAs can also act to impose transcriptional silencing[92] (Figure 2). There is also increasing evidence of intersecting pathways, such as RNA editing and modification, in these networks[93–96].

Piwi-associated small RNAs

While most Ago proteins are expressed ubiquitously and associate with miRNAs and siRNAs, there is a subclade of Argonaute proteins, termed Piwi, that is required for germ cell development[97–100]. Piwi and Piwi-like proteins associate with a distinctive class of small RNAs (26–30 nt; piRNAs), which act to epigenetically and post-transcriptionally silence transposons in germ cells[101–110]. Piwi is found predominantly in the nucleus[111], co-localizes in an RNA-dependent manner with Polycomb group proteins[112], and appears to be expressed in other tissues, including the brain[113], suggesting a role beyond genome protection in epigenetic processes[114,115].

Other classes of small RNAs in eukaryotes

The molecular genetics, biochemistry and structural biology of the RNAi system are still being unravelled, but indicate an ancient, widespread and multilaterally adapted system that controls many cellular processes, whose dimensions are still being explored. These include potentially lineage-specific variations, such as the '21U' RNAs in *C. elegans*[116]. Surprisingly, it appears that all snoRNAs from fission yeast to human produce at least 3 different subclasses of small RNAs[117], one of which has the same size and functions as a miRNA[118], and another that is similar in size to piRNA[117]. There are also intriguing and recurring reports of fragments of tRNAs produced in tissue-specific patterns[119] and associated with Ago proteins[120].

More recently, deep sequencing of small RNA populations has revealed the existence of another class of small RNAs in animals but not plants, which are 17–18 nt in length and associated with transcription initiation ('tiRNAs')[121] and splice sites ('spliRNAs')[122] (Figure 2). The origin and function of these RNAs is uncertain, but preliminary evidence suggests that they may play a role in nucleosome positioning[123] and/or be involved in other levels of chromatin organization[124]. There are also other reports of less distinct classes of promoter-associated RNAs called PASRs[125], TSSa-RNAs[126] and PROMPTS[127], some of which may play a role in RNA-directed transcriptional gene silencing[128].

Regulatory RNAs in prokaryotes

Many small regulatory RNAs (sRNAs) have been identified in bacteria, which regulate a wide variety of adaptive responses. Bacterial sRNAs generally function by simple antisense mechanisms to regulate translation or stability of target mRNAs by altering their secondary structure to expose or sequester cis-acting sites[129,130]. Studies in bacteria have also identified *cis*-acting regulatory RNA sequences ('riboswitches'), which act allosterically by binding metabolites to regulate gene expression[131,132], and which almost certainly exist as part of the RNA regulatory landscape in all kingdoms of life.

Very recently, the prokaryotic kingdom has once again surprised us with the sophistication of its molecular machinery. Many bacterial and most archaeal genomes possess loci comprised of regularly spaced repeats interspersed by other DNA sequences derived from viruses[133–136]. These loci, now termed CRISPRs (**C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeats) act as an innate immune system by incorporating fragments of viral DNA between the repeats, which are then transcribed and processed to produce small guide RNAs (linked to their effector complexes via the repeat sequence) that target and destroy viral DNA[137–140] or RNA[141]). This system has recently been adapted for RNA programmable sequence-specific genome manipulation in eukaryotes, including mammals[142–145], with extraordinary versatility including targeted gene excision and fusion, and modified CRISPRs capable of recruiting silencing and activating proteins to target loci[146–150]. Moreover, the biological arms race continues, with bacteriophages encoding their own CRISPR system to evade host innate immunity[151].

Long noncoding RNAs

The eukaryotic transcriptome and long non-coding RNAs

Noting that the density and size of introns (and, as it turned out later, intergenic sequences) expanded with developmental complexity, Mattick posited in 1994 that introns had evolved to express an expanding repertoire of *trans*-acting regulatory RNAs, that some genes subsequently evolved to express only (intronic or exonic) regulatory RNAs, and that this RNA-based regulatory system was the essential prerequisite for the emergence of developmentally complex organisms[152]. Subsequently, the application of genome tiling array technology and deep sequencing to the characterization of the transcriptome showed that there are tens of thousands of loci in mammals that express large transcripts that do not encode proteins, located intergenic, intronic and antisense to protein-coding genes. The initial findings[153–155] were confirmed in 2005[156–159] and extended by the ENCODE project[160–162], all of which showed that the vast majority (at least 80%) of the human and mouse genomes are differentially transcribed in one context or another, with other studies reporting similar findings in all organisms examined. Indeed, it seems most intergenic (and by definition intronic) sequences are differentially transcribed, and therefore that the extent of the transcriptome expands with developmental complexity[163].

Using more focussed deep sequencing methodologies, it has become evident that the full repertoire of the protein-coding and non-protein-coding transcriptome is still vastly under-sampled[164]. In addition, many transcripts are not polyadenylated, and represent a largely different sequence class[156,165], some of which appear relevant to development (e.g., OCT4[166,167]). Moreover, 95% of human transcription initiation sites are not associated with mRNAs, but rather mainly with non-polyadenylated noncoding transcription[168]. These non-polyadenylated transcripts are as yet largely uncharacterized because of the historical use of polyA tails to remove the overwhelming rRNA contamination in RNA preparations, which is being alleviated by more sophisticated approaches, including cap trapping[169], oligonucleotide subtraction[170] and array capture[164,171].

Defining long noncoding RNAs

Long noncoding RNAs (lncRNAs) are operationally defined as any non-protein-coding RNA >200 nt in length, which corresponds to a convenient cut-off in biochemical fractionation and excludes all known classes of small RNAs[172]. Transcripts are judged to be 'noncoding' if they lack a long open reading frame (traditionally >100 codons) and/or do not show codon conservation, although with initially limited genomic and transcriptomic data for comparison, there was considerable uncertainty. However, recent analyses provide strong evidence that most annotated lncRNAs do not encode proteins, although some specify small proteins that had previously fallen under the bioinformatic radar[173–175].

These noncoding transcripts can be parsed into intronic, antisense or intergenic ('lincRNA') subsets in experimental studies and databases[159,176,177], partly because of mechanistic expectations[178] coupled with a desire to reduce ambiguity and overlap with protein-coding loci in functional analyses[179–181]. However, there is no evidence of any intrinsic difference between RNAs that are intronic, intergenic, antisense or overlap with protein-coding genes, for example in their interaction with chromatin-activating or -repressive complexes (Figure 1 and below), although subclasses will inevitably exist and be defined, some of which may have be biased in relation to genomic origin.

Exploration of lncRNA functions

The unexpected discovery of large numbers of noncoding transcripts in eukaryotes, some of which span tens or hundreds of kilobases[182], has led to debates about their functionality (see e.g. [183,184]), particularly since many have relatively low evolutionary conservation and low levels of expression, leading some to posit that they represent 'transcriptional noise' and/or redundant transcripts with no biological significance. This is a possibility, at least in part. However, it is clear that lncRNAs actually show a wide range of evolutionary conservation, from those that are ultraconserved[185] to those that are primate-specific[186–188], which can be explained as the result of different structure-function constraints and lineage-specific adaptive radiation[189]. Indeed there is now considerable evidence that lack of primary sequence conservation in lncRNAs does not indicate lack of function[190,191], and that many lncRNAs show evidence of structural conservation[192,193].

Moreover, those loci expressing lncRNAs show all of the hallmarks of *bona fide* genes[4], including conservation of promoters[169], indicative chromatin structure[194] and regulation by conventional morphogens and transcription factors[195]. LncRNAs have been found to have a similar range of cellular half-lives as mRNAs[196] and to be differentially expressed in a tissue-specific manner[158,197], with higher resolution in the brain[198]. The latter showed that while the level of expression of many lncRNAs superficially appears to be lower than mRNAs in whole tissues, lncRNAs are highly expressed and easily detectable in particular cells[198], and it appears that they have, on average, higher cell-specificity than proteins[165,199], consistent with their proposed role in architectural (as opposed to 'cell-type') regulation, where each cell has a unique positional identity in precisely sculpted organs, bones and muscles[200].

Many lncRNAs are alternatively spliced[201], further evidence of the precision of their expression and hard to reconcile with the suggestion that they are simply transcriptional noise. It should also be noted that some functionally validated lncRNAs can have isoforms that encode proteins[202] and reciprocally that some (perhaps many) mRNAs may also have intrinsic functions as *trans*-acting regulatory RNAs[203–205], that in some contexts 3′UTRs can be separately expressed and convey genetic functions in *trans*[204], and that both may be further processed to produce subsidiary species[206].

LncRNAs have been shown to be dynamically expressed in a range of differentiating systems, including embryonal stem cell[207], muscle[208] T-cell[209], breast[210,211], erythroid[211] and neuronal differentiation[212–214], as well as in cancer and other diseases (see e.g.[210,215–222]), at least partly controlled by conventional transcription factors[195,213].

The validation of lncRNA function has to date mainly relied on knockdown of candidate lncRNAs. It has proven surprisingly easy to knockdown lncRNA expression by si/shRNA-mediated approaches, and thereby to detect phenotypic changes in cultured cells, where most analyses have been carried out. By 2009, ~50 lncRNAs had been shown to be functional[4] and hundreds more are now published or en route to publication, a large enough sample to draw the conclusion that these transcripts are generally functional.

Roles in development and differentiation

Some lncRNAs play a role in general or differentiation-specific cell biological processes. These include: Template RNAs that guide chromosomal rearrangements in ciliates[223]; Terra RNAs, involved in telomere biology[224]; 7S RNA, an essential component of the signal recognition particle involved in protein export[225]; 7SK, a highly expressed structured RNA that is the scaffold to assemble a multimeric protein complex containing SR splicing proteins and P-TEFb, a cyclin-dependent kinase required for transcriptional elongation by RNA polymerase II and other factors [226]; Neat1, an essential component of paraspeckles, enigmatic subnuclear organelles that appear in differentiated but not stem cells in mammals[227,228]; the nuclear-localized MALAT-1, which regulates alternative splicing[229] and cell cycle progression[230]; Gomafu, which is expressed in an unknown subnuclear structure, possibly a specialized spliceosome, in a subset of neurons[231], and has recently been implicated in schizophrenia[232]; and others of unknown function associated with bipolar structures in the nuclei of Purkinje cells[198].

Not surprisingly, given their expression patterns, most functionally analyzed lncRNAs appear to play roles in the regulation of differentiation and development[233]. These include, based on studies in cell culture, the regulation of apoptosis and metastatic processes[211,218,220,221,234], retinal and erythroid development[211,235], breast development[210,236], and epidermal differentiation[237], among many others.

Antisense knockdown of lncRNAs in zebrafish and deletion of sequences specifying lncRNAs in mouse have shown that some confer visible developmental defects[181,191,238,239], although others do not, including knockouts of the widely expressed Neat1 required for paraspeckle function[240] and some of the most highly conserved sequences in the mammalian genome[241]. This suggests that more sophisticated phenotypic

screens may be required, especially in relation to cognitive function, since most mammalian lncRNAs are expressed in the brain[198] and many are mammal- or primate-specific[242,243]. A good example is the retrotransposon-derived lncRNA BC1, which is widely expressed in the brain but whose knockout causes no visible anatomical abnormality, but leads to behavioural changes that would be lethal in the wild[244].

Epigenetic roles of noncoding RNAs

Consistent with their roles in differentiation and development, a range of genetic and biochemical evidence suggests that a major function of lncRNAs and many small RNAs is the regulation of epigenetic processes[245,246], likely by guiding chromatin-modifying enzymes to their sites of action and/or acting as scaffolding for chromosomal organization[179,246–249] (Figure 3).

RNAs were first shown to induce transcriptional gene silencing in plants[74,250], fungi[251] and human cells[88], with an intimate involvement of small RNAs and the RNA interference pathway in the epigenetic processes involved[251–253], consistent with the observation that small RNAs interact with Polycomb[254] and that Ago proteins occur in the nucleus[86,87] (Figure 3).

In parallel, dating back to 1990, antisense RNAs have also been shown to affect gene expression, again initially in plants[73] and later in animals[159,166,255–257]. Some lncRNAs, similar to small ncRNAs[258], have been shown to control alternative splicing[259,260]. Other naturally occurring lncRNAs had been shown to control epigenetic processes *in vivo*, notably in X chromosome dosage compensation[261–265] and parental imprinting in mammals[266–268] and vernalization in plants[269], with subsequent studies showing that intergenic and antisense RNAs bind to polycomb chromatin repressive complexes[194,270–272], to trithorax chromatin activating complexes and activated forms of histones[207], and to DNA methyltransferases[201,273,274]. These observations were writ large in 2009 when it was shown that approximately 20% of ~3,300 lncRNAs examined were bound by PRC2, with others bound by other chromatin-modifying complexes, that siRNA-mediated knockdown of lncRNAs associated with PRC2 led to changes in gene expression, with the up-regulated genes being enriched for those normally silenced by PRC2[179], and that Polycomb binds RNA with high affinity but low specificity[275], consistent with the idea that many RNAs are Polycomb interactors.

One of the notable lncRNAs to emerge from the studies of Rinn, Chang and colleagues, HOTAIR, is derived from the *HOXC* locus and regulates *HOXD* in *trans*[194], is involved in cancer metastasis[220] and, when inactivated, results in homeotic transformation *in vivo*[276], Chang and colleagues also showed that lncRNAs could act as scaffolds for the assembly of histone modification complexes[277], with the widespread alternative splicing of these RNAs suggesting that the cargo and/or target specificity can be varied in a context-dependent and differentiation-specific manner.

LncRNAs may also be involved in orchestrating the highly dynamic spatial structure of chromatin during differentiation and development[164,278], which would explain their often highly cell-specific expression patterns[200]. Developmental enhancers, as well as polycomb-

and trithorax-response elements, are transcribed in the cells in which they are active[203,279–284], and are likely not only scaffolds for the recruitment of epigenetic regulators[285] but also the physical mediators of the complex genetic phenomena of transvection and transinduction[245]. Moreover, many lncRNAs display the properties of enhancers[180]. These RNAs may well guide the physical looping that occurs between enhancers, target promoters and exons, with precise positioning of nucleosomes[286–290], to control transcription and alternative splicing[248,291,292], again modulated by alternative splicing. Indeed the emerging picture is of a chromatin and transcriptional landscape that is exquisitely and precisely controlled in 4 dimensions by a suite of regulatory RNAs that assemble relatively generic (albeit often cell- or differentiation state-specific) enzyme complexes and isoforms to their sites of action in a context-dependent manner[249].

A substantial proportion of lncRNAs reside within, or are dynamically shuttled, to the cytoplasm indicating roles in other cellular processes, including the regulation of protein localization[293], mRNA translation[294] and mRNA stability[295].

RNA editing, modification, retrotransposition and inheritance

Regulatory RNAs may also be influenced by environmental signals and be transmitted between cells and generations, which has important implications for understanding gene-environment interactions and evolution. There is evidence that plasticity has been superimposed on RNA-directed epigenetic networks by the expansion of RNA editing, especially during cognitive evolution[296,297], and by retrotransposon utilization and mobility[114,298–301], which harks back to the insights of McClintock and Britten & Davidson. The raw material for evolution is gene duplication and transposition, the latter having the advantage of being able to mobilize functional cassettes in regulatory networks[302], which appears to be the main driver of adaptive radiation[245,303]. Indeed many lncRNAs may have originated from retrotransposons and the evolution of mRNAs and lncRNAs may have been accelerated by retrotransposition of functional modules[304–308].

Moreover, apart from snoRNA-directed modifications, there are well over 100 other documented modifications of RNA[309,310], including cytosine and adenosine methylation which have known physiological and cognitive effects[311–314], indicating a new additional layer of RNA informational code and epitranscriptomics, an exciting field that is just beginning to emerge[315,316].

There is evidence for systemic transmission of RNA[317,318] and RNA-mediated epigenetic inheritance in plants and animals[319–323]. There is also the intriguing possibility of RNA-directed DNA recoding, which may place RNA at the centre not only of gene regulation in the developmental ontogeny of higher organisms, but also of both hard- and soft-wired somatic and germline evolution[324–326].

Conclusions and outlook

The past two decades have seen an explosion in our understanding of the previously hidden and unanticipated world of RNA regulation. Indeed, in retrospect, it appears that we may have fundamentally misunderstood the nature of the genetic programming in complex

organisms because of the assumption that most genetic information is transacted by proteins. This maybe largely true in simpler organisms, but is turning out not to be the case in more complex organisms, whose genomes appear to be progressively dominated by regulatory RNAs that orchestrate the epigenetic trajectories of differentiation and development.

The picture that emerges is of an extraordinarily complex transcriptional landscape in mammals and other multicellular organisms, comprised of overlapping, intergenic and intronic sense and antisense small and large RNAs with interlaced exons[327,328], whose promoters, splicing patterns, polyadenylation sites and regional repertoire varies in different cells and developmental contexts (see below). Since there appear to be few distinct boundaries to genes in humans, it seems better to change the focus of analysis to the transcript, with genetic loci redefined as fuzzy transcription clusters[165,328,329] albeit semantically anchored or related to an enclosed or nearby protein-coding locus. However, this can only be stretched so far, and non-protein-coding loci raise problems for existing schema of human genome nomenclature.

Indeed even the notion of a (simple) protein-coding sequence needs to be reassessed. It is becoming evident not only that mRNAs can have multiple functions[205], but also that protein-coding sequences themselves can have other embedded functions, as suggested by constraints on synonymous codon usage[330,331], including regulatory functions as epigenetic modulators[203], tissue-specific enhancers[331,332] and transcription factor binding sites[333]. The possibility, if not likelihood, is that there is a very complex functional and evolutionary interplay between the protein-coding and regulatory functions of RNAs[200], and that some lncRNAs may have evolved, at least in part, from protein-coding genes, as appears to have occurred with Xist, by duplication or pseudogenization followed by the emergence of paralogous regulatory and/or coding functions[201,334]. Conversely, it appears that new protein-coding capacity may also appear in lncRNAs[174].

The sheer number and diversity of RNAs juxtaposed with their extraordinarily complex molecular functions (Figure 3) in regulating epigenetic processes, subcellular organelles, protein-coding and non-coding gene transcription, translation, RNA turnover, chromosomal organization and integrity, and genome defence, among others, suggests that we have a long way to go to understand the structure and functions of what is surely a highly interconnected system. There are literally tens of thousands, if not more, of individual noncoding RNAs whose roles in cell and developmental biology, as well as brain function, remain to be determined. Moreover, many if not most regulatory RNAs, especially in complex organisms, remain to be identified, including new classes such as the circular RNAs and others that may function as miRNA ‘sponges’[62,335–340], which will require targeted deep sequencing of small and large RNAs that are derived from different genomic locations in different cells, using targeted techniques such as RNA CaptureSeq[164,171].

RNA is not a linear molecule, but can rather fold into complex and allosterically responsive 3-dimensional structures that can both recruit generic effector proteins and guide the resulting complexes sequence-specifically to other RNAs and DNA, via duplex or triplex formation. There are many important questions. These include the identification of functional domains in RNA and their interacting partners, so that we can predict and parse

RNA functional interactions in the same way that is already done by recognition of well-characterized motifs and domains in proteins. One way to do this, already underway in many laboratories, is to combine immunoprecipitation of different types of RNA binding proteins (chromatin-modifying proteins, transcription factors, and RNA transport proteins, among others) with deep sequencing of the associated RNAs (RIP-Seq) followed by analysis of primary and predicted secondary structures, and ultimately by biochemical validation and characterization.

Determination of the structure of RNAs, RNA-RNA, RNA-DNA and ribonucleoprotein complexes will be a rapidly growing field, requiring the development of new technologies, such as RNA footprinting with high-throughput sequencing[341], as well in vivo studies using RNA-based genetic techniques like CRISPR-mediated mutation[143]. Other objectives include determination of whether small RNA pathways are used in viral defence in humans[80], the functions of ti/spliRNAs and snoRNA-derived small RNAs, the roles of piRNAs in retrotransposon dynamics and the remodelling of the genome by retrotransposons in the brain[114], the mechanisms and extent of RNA-mediated trans-generational epigenetic inheritance[342], the locations of RNA binding sites (RNA-DNA duplexes and RNA-DNA:DNA triplexes) in DNA, the cross-talk between different types of regulatory RNAs, the logic and hierarchy of RNA- and protein-mediated regulation of gene expression, and the extent, mechanisms and information content of RNA-mediated communication between cells, both within[318] and between organisms (‘social RNA’)[343]. Indeed it appears that RNA is the computational engine of cell biology, developmental biology, brain function and perhaps even evolution itself[325]. The complexity and interconnectedness of these systems should not be cause for concern but rather the motivation for exploring the vast unknown universe of RNA regulation, without which we will not understand biology.

Acknowledgments

This work was supported by NIH PO1 AI099783-01 and ARC Future Fellowship FT130100572 (KVM) and by NHMRC Australia Fellowship 631688 (JSM).

Glossary

Antisense RNA	A single stranded RNA that is complimentary to a messenger RNA or a gene
ENCODE	Encyclopaedia of DNA elements is a consortium of international collaborators involved in building a comprehensive list of functional elements in the human genome[160,161]
hnRNA	Heterogenous nuclear RNA. Similar to messenger RNA, or pre-mRNAs, but retained predominantly in the nucleus
Intron	A term first coined by Walter Gilbert to describe those nucleotide regions in RNA that are removed, by being spliced out, to produce messenger RNAs[34,35]

lncRNA	Long non-coding RNAs, first used to differentiate between smaller forms of non-coding RNA, e.g. greater than 200bp in size[358]
piRNA	Piwi associated RNAs. Small RNAs associated with the Piwi protein complex and emanating from transposable like elements[100]
Pseudogene	Relics of genes that have lost their protein coding potential but remain transcribed and integrated within the genome[359]
PTGS	Post-transcriptional gene silencing. Silencing a gene at the messenger RNA or translational level, after transcription has occurred[210]
RNA-directed DNA methylation	An epigenetic process whereby processed double stranded small RNAs (21–24bp) guide the methylation of homologous DNA loci[52]
siRNA	Small interfering RNAs, double stranded RNAs that can be used to suppress homology containing transcripts in a transcriptional and post-transcriptional manner[352]
Transposons	Mobile genetic elements[262], with evolutionary links to retroviruses
tiRNAs	Transcription initiation RNAs are small RNAs associated with promoters with peak density ~10–30 nucleotides downstream of transcriptional start sites[121]. Similar RNAs are derived from splice sites (spliRNAs)[122]
TGS	Transcriptional gene silencing. The regulation of a gene at the transcriptional level
UTRs	Untranslated regions, referring to either side, 5′ (leader sequence) or 3′ (trailer sequence) of a coding sequence on a strand of messenger RNA

Literature cited

1. Gilbert W. Origin of life: The RNA world. Nature. 1986; 319:618.
2. Beadle GW, Tatum EL. Genetic control of biochemical reactions in Neurospora. Proc Natl Acad Sci U S A. 1941; 27:499–506. [PubMed: 16588492]
3. Comfort, NC. The Tangled Field: Barbara McClintock's Search for the Patterns of Genetic Control. Harvard University Press; Cambridge, Massachussetts: 2003.
4. Mattick JS. The genetic signatures of noncoding RNAs. PLoS Genet. 2009; 5:e1000459. [PubMed: 19390609]
5. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953; 171:737–8. [PubMed: 13054692]
6. Crick FH. On protein synthesis. Symp Soc Exp Biol. 1958; 12:138–63. [PubMed: 13580867]
7. Palade GE. A small particulate component of the cytoplasm. J Biophys Biochem Cytol. 1955; 1:59–68. [PubMed: 14381428]
8. Hoagland MB, Stephenson ML, Scott JF, Hecht LI, Zamecnik PC. A soluble ribonucleic acid intermediate in protein synthesis. J Biol Chem. 1958; 231:241–57. [PubMed: 13538965]
9. Brenner S, Jacob F, Meselson M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature. 1961; 190:576–581. [PubMed: 20446365]

10. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol. 1961; 3:318–356. [PubMed: 13718526]
11. Gilbert W, Muller-Hill B. Isolation of the *lac* repressor. Proc Natl Acad Sci U S A. 1966; 56:1891–8. [PubMed: 16591435]
12. Levine M, Tjian R. Transcription regulation and animal diversity. Nature. 2003; 424:147–51. [PubMed: 12853946]
13. Mattick JS, Gagen MJ. Accelerating networks. Science. 2005; 307:856–8. [PubMed: 15705831]
14. Freedman ML, et al. Principles for the post-GWAS functional characterization of cancer risk loci. Nat Genet. 2011; 43:513–8. [PubMed: 21614091]
15. Weinberg RA, Penman S. Small molecular weight monodisperse nuclear RNA. J Mol Biol. 1968; 38:289–304. [PubMed: 5718554]
16. Dreyfuss G, Philipson L, Mattaj IW. Ribonucleoprotein particles in cellular processes. J Cell Biol. 1988; 106:1419–25. [PubMed: 2836428]
17. Butcher SE, Brow DA. Towards understanding the catalytic core structure of the spliceosome. Biochem Soc Trans. 2005; 33:447–9. [PubMed: 15916538]
18. Wang Z, Burge CB. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. RNA. 2008; 14:802–13. [PubMed: 18369186]
19. Pessa HK, et al. Minor spliceosome components are predominantly localized in the nucleus. Proc Natl Acad Sci U S A. 2008; 105:8655–60. [PubMed: 18559850]
20. Maxwell ES, Fournier MJ. The small nucleolar RNAs. Annu Rev Biochem. 1995; 64:897–934. [PubMed: 7574504]
21. Henras AK, Dez C, Henry Y. RNA structure and function in C/D and H/ACA s(no)RNPs. Curr Opin Struct Biol. 2004; 14:335–43. [PubMed: 15193314]
22. Meier UT. The many facets of H/ACA ribonucleoproteins. Chromosoma. 2005; 114:1–14. [PubMed: 15770508]
23. Cavaille J, Seitz H, Paulsen M, Ferguson-Smith AC, Bachellerie JP. Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. Hum Mol Genet. 2002; 11:1527–38. [PubMed: 12045206]
24. Rogelj B, Hartmann CE, Yeo CH, Hunt SP, Giese KP. Contextual fear conditioning regulates the expression of brain-specific small nucleolar RNAs in hippocampus. Eur J Neurosci. 2003; 18:3089–96. [PubMed: 14656304]
25. Bachellerie JP, Cavaille J, Huttenhofer A. The expanding snoRNA world. Biochimie. 2002; 84:775–90. [PubMed: 12457565]
26. Jady BE, Bertrand E, Kiss T. Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. J Cell Biol. 2004; 164:647–52. [PubMed: 14981093]
27. Warner JR, Soeiro R, Birnboim HC, Girard M, Darnell JE. Rapidly labeled HeLa cell nuclear RNA. I. Identification by zone sedimentation of a heterogeneous fraction separate from ribosomal precursor RNA. J Mol Biol. 1966; 19:349–61. [PubMed: 5969070]
28. Britten RJ, Davidson EH. Gene regulation for higher cells: a theory. Science. 1969; 165:349–57. [PubMed: 5789433]
29. Britten RJ, Davidson EH. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. Q Rev Biol. 1971; 66:111–38. [PubMed: 5160087]
30. Davidson EH, Klein WH, Britten RJ. Sequence organization in animal DNA and a speculation on hnRNA as a coordinate regulatory transcript. Dev Biol. 1977; 55:69–84. [PubMed: 832773]
31. Howard ML, Davidson EH. cis-Regulatory control circuits in development. Dev Biol. 2004; 271:109–18. [PubMed: 15196954]
32. Davidson, EH. The Regulatory Genome: Gene Regulatory Networks In Development And Evolution. Academic Press; 2006.
33. Britten R. Transposable elements have contributed to thousands of human proteins. Proc Natl Acad Sci U S A. 2006; 103:1798–803. [PubMed: 16443682]
34. Berget SM, Moore C, Sharp PA. Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci U S A. 1977; 74:3171–5. [PubMed: 269380]

35. Chow LT, Gelinas RE, Broker TR, Roberts RJ. An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. Cell. 1977; 12:1–8. [PubMed: 902310]
36. Williamson B. DNA insertions and gene structure. Nature. 1977; 270:295–7.
37. Gilbert W, Marchionni M, McKnight G. On the antiquity of introns. Cell. 1986; 46:151–154. [PubMed: 2424613]
38. Doolittle WF, Sapienza C. Selfish genes, the phenotype paradigm and genome evolution. Nature. 1980; 284:601–3. [PubMed: 6245369]
39. Orgel LE, Crick FH. Selfish DNA: the ultimate parasite. Nature. 1980; 284:604–7. [PubMed: 7366731]
40. Kruger K, et al. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell. 1982; 31:147–57. [PubMed: 6297745]
41. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell. 1983; 35:849–57. [PubMed: 6197186]
42. Fica SM, et al. RNA catalyses nuclear pre-mRNA splicing. Nature. 2013; 503:229–34. [PubMed: 24196718]
43. Steitz TA, Moore PB. RNA, the first macromolecular catalyst: the ribosome is a ribozyme. Trends Biochem Sci. 2003; 28:411–418. [PubMed: 12932729]
44. Webb CH, Riccitelli NJ, Ruminski DJ, Luptak A. Widespread occurrence of self-cleaving ribozymes. Science. 2009; 326:953. [PubMed: 19965505]
45. de la Pena M, Garcia-Robles I. Intronic hammerhead ribozymes are ultraconserved in the human genome. EMBO Rep. 2010; 11:711–6. [PubMed: 20651741]
46. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993; 75:843–54. [PubMed: 8252621]
47. Reinhart BJ, et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature. 2000; 403:901–6. [PubMed: 10706289]
48. Pasquinelli AE, et al. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature. 2000; 408:86–9. [PubMed: 11081512]
49. Brennecke J, Hipfner DR, Stark A, Russell RB, Cohen SM. bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. Cell. 2003; 113:25–36. [PubMed: 12679032]
50. Johnston RJ, Hobert O. A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. Nature. 2003; 426:845–849. [PubMed: 14685240]
51. Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science. 2001; 294:858–862. [PubMed: 11679671]
52. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science. 2001; 294:853–858. [PubMed: 11679670]
53. Lee RC, Ambros V. An extensive class of small RNAs in *Caenorhabditis elegans*. Science. 2001; 294:862–864. [PubMed: 11679672]
54. Williams TM, et al. The regulation and evolution of a genetic switch controlling sexually dimorphic traits in Drosophila. Cell. 2008; 134:610–23. [PubMed: 18724934]
55. Kaya KD, Karakulah G, Yakicier CM, Acar AC, Konu O. mESAdb: microRNA expression and sequence analysis database. Nucleic Acids Res. 2011; 39:D170–80. [PubMed: 21177657]
56. Berezikov E, et al. Diversity of microRNAs in human and chimpanzee brain. Nat Genet. 2006; 38:1375–7. [PubMed: 17072315]
57. Heimberg AM, Sempere LF, Moy VN, Donoghue PC, Peterson KJ. MicroRNAs and the advent of vertebrate morphological complexity. Proc Natl Acad Sci U S A. 2008; 105:2946–50. [PubMed: 18287013]
58. John B, et al. Human microRNA targets. PLoS Biol. 2004; 2:e363. [PubMed: 15502875]
59. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005; 120:15–20. [PubMed: 15652477]
60. Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem. 2010; 79:351–79. [PubMed: 20533884]

61. Schnall-Levin M, et al. Unusually effective microRNA targeting within repeat-rich coding regions of mammalian mRNAs. Genome Res. 2011; 21:1395–403. [PubMed: 21685129]
62. Hansen TB, et al. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. EMBO J. 2011; 30:4414–22. [PubMed: 21964070]
63. Leonardo TR, Schultheisz HL, Loring JF, Laurent LC. The functions of microRNAs in pluripotency and reprogramming. Nat Cell Biol. 2012; 14:1114–21. [PubMed: 23131918]
64. Bracken CP, Gregory PA, Khew-Goodall Y, Goodall GJ. The role of microRNAs in metastasis and epithelial-mesenchymal transition. Cell Mol Life Sci. 2009; 66:1682–99. [PubMed: 19153653]
65. Rakoczy J, et al. MicroRNAs-140–5p/140–3p modulate Leydig cell numbers in the developing mouse testis. Biol Reprod. 2013; 88:143. [PubMed: 23616593]
66. Fernandez-Valverde SL, Taft RJ, Mattick JS. MicroRNAs in beta-cell biology, insulin resistance, diabetes and its complications. Diabetes. 2011; 60:1825–31. [PubMed: 21709277]
67. Bredy TW, Lin Q, Wei W, Baker-Andresen D, Mattick JS. MicroRNA regulation of neural plasticity and memory. Neurobiol Learn Mem. 2011; 96:89–94. [PubMed: 21524708]
68. Park CY, Choi YS, McManus MT. Analysis of microRNA knockouts in mice. Hum Mol Genet. 2010; 19:R169–75. [PubMed: 20805106]
69. Waterhouse PM, Graham MW, Wang MB. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc Natl Acad Sci U S A. 1998; 95:13959–64. [PubMed: 9811908]
70. Fire A, et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature. 1998; 391:806–11. [PubMed: 9486653]
71. Napoli C, Lemieux C, Jorgensen R. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell. 1990; 2:279–289. [PubMed: 12354959]
72. Matzke MA, Primig M, Trnovsky J, Matzke AJM. Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. The EMBO Journal. 1989; 8:643–649. [PubMed: 16453872]
73. van der Krol AR, Mur LA, de Lange P, Mol JN, Stuitje AR. Inhibition of flower pigmentation by antisense CHS genes: promoter and minimal sequence requirements for the antisense effect. Plant Mol Biol. 1990; 14:457–66. [PubMed: 2102827]
74. Wassenegger M, Heimes S, Riedel L, Sanger HL. RNA-directed de novo methylation of genomic sequences in plants. Cell. 1994; 76:567–76. [PubMed: 8313476]
75. Basyuk E, Suavet F, Doglio A, Bordonne R, Bertrand E. Human let-7 stem-loop precursors harbor features of RNase III cleavage products. Nucleic Acids Res. 2003; 31:6593–7. [PubMed: 14602919]
76. Lee Y, et al. The nuclear RNase III Drosha initiates microRNA processing. Nature. 2003; 425:415–9. [PubMed: 14508493]
77. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001; 409:363–366. [PubMed: 11201747]
78. Doi N, et al. Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. Curr Biol. 2003; 13:41–6. [PubMed: 12526743]
79. Peters L, Meister G. Argonaute proteins: mediators of RNA silencing. Mol Cell. 2007; 26:611–23. [PubMed: 17560368]
80. Maillard PV, et al. Antiviral RNA interference in mammalian cells. Science. 2013; 342:235–8. [PubMed: 24115438]
81. Zeng Y, Yi R, Cullen BR. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci U S A. 2003; 100:9779–84. [PubMed: 12902540]
82. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature. 2010; 466:835–40. [PubMed: 20703300]
83. Ramachandran PV, Ignacimuthu S. RNA interference--a silent but an efficient therapeutic tool. Appl Biochem Biotechnol. 2013; 169:1774–89. [PubMed: 23340870]

84. Ahlenstiel CL, et al. Direct evidence of nuclear Argonaute distribution during transcriptional silencing links the actin cytoskeleton to nuclear RNAi machinery in human cells. Nucleic Acids Res. 2012; 40:1579–95. [PubMed: 22064859]

85. Ameyar-Zazoua M, et al. Argonaute proteins couple chromatin silencing to alternative splicing. Nat Struct Mol Biol. 2012; 19:998–1004. [PubMed: 22961379]

86. Rudel S, Flatley A, Weinmann L, Kremmer E, Meister G. A multifunctional human Argonaute2-specific monoclonal antibody. RNA. 2008; 14:1244–53. [PubMed: 18430891]

87. Kim DH, Villeneuve LM, Morris KV, Rossi JJ. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat Struct Mol Biol. 2006; 13:793–7. [PubMed: 16936726]

88. Morris KV, Chan SW, Jacobsen SE, Looney DJ. Small interfering RNA-induced transcriptional gene silencing in human cells. Science. 2004; 305:1289–92. [PubMed: 15297624]

89. Fernandez-Valverde SL, Taft RJ, Mattick JS. Dynamic isomiR regulation in Drosophila development. RNA. 2010; 16:1881–8. [PubMed: 20805289]

90. Didiano D, Hobert O. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. Nat Struct Mol Biol. 2006; 13:849–51. [PubMed: 16921378]

91. Shin C, et al. Expanding the microRNA targeting code: functional sites with centered pairing. Mol Cell. 2010; 38:789–802. [PubMed: 20620952]

92. Kim DH, Saetrom P, Snove O Jr, Rossi JJ. MicroRNA-directed transcriptional gene silencing in mammalian cells. Proc Natl Acad Sci U S A. 2008; 105:16230–5. [PubMed: 18852463]

93. Blow MJ, et al. RNA editing of human microRNAs. Genome Biol. 2006; 7:R27. [PubMed: 16594986]

94. Hundley HA, Bass BL. ADAR editing in double-stranded UTRs and other noncoding RNA sequences. Trends Biochem Sci. 2010; 35:377–83. [PubMed: 20382028]

95. Kawahara Y, Zinshteyn B, Chendrimada TP, Shiekhattar R, Nishikura K. RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. EMBO Rep. 2007; 8:763–9. [PubMed: 17599088]

96. Ota T, et al. Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat Genet. 2004; 36:40–5. [PubMed: 14702039]

97. Lin H, Spradling AC. A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development. 1997; 124:2463–76. [PubMed: 9199372]

98. Cox DN, et al. A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. 1998; 12:3715–27. [PubMed: 9851978]

99. Kuramochi-Miyagawa S, et al. Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development. 2004; 131:839–49. [PubMed: 14736746]

100. Kim JK, et al. Functional genomic analysis of RNA interference in C. elegans. Science. 2005; 308:1164–7. [PubMed: 15790806]

101. Pal-Bhadra M, Bhadra U, Birchler JA. RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in Drosophila. Mol Cell. 2002; 9:315–27. [PubMed: 11864605]

102. Vagin VV, et al. A distinct small RNA pathway silences selfish genetic elements in the germline. Science. 2006; 313:320–4. [PubMed: 16809489]

103. Aravin A, et al. A novel class of small RNAs bind to MILI protein in mouse testes. Nature. 2006; 442:203–7. [PubMed: 16751777]

104. Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature. 2006; 442:199–202. [PubMed: 16751776]

105. Grivna ST, Beyret E, Wang Z, Lin H. A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. 2006; 20:1709–14. [PubMed: 16766680]

106. Watanabe T, et al. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev. 2006; 20:1732–43. [PubMed: 16766679]

107. Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science. 2007; 318:761–4. [PubMed: 17975059]
108. Brennecke J, et al. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell. 2007; 128:1089–103. [PubMed: 17346786]
109. Brennecke J, et al. An epigenetic role for maternally inherited piRNAs in transposon silencing. Science. 2008; 322:1387–92. [PubMed: 19039138]
110. Kuramochi-Miyagawa S, et al. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev. 2008; 22:908–17. [PubMed: 18381894]
111. Cox DN, Chao A, Lin H. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development. 2000; 127:503–14. [PubMed: 10631171]
112. Grimaud C, et al. RNAi components are required for nuclear clustering of Polycomb group response elements. Cell. 2006; 124:957–71. [PubMed: 16530043]
113. Rajasethupathy P, et al. A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. Cell. 2012; 149:693–707. [PubMed: 22541438]
114. Baillie JK, et al. Somatic retrotransposition alters the genetic landscape of the human brain. Nature. 2011; 479:534–7. [PubMed: 22037309]
115. Ross RJ, Weiner MM, Lin H. PIWI proteins and PIWI-interacting RNAs in the soma. Nature. 2014; 505:353–9. [PubMed: 24429634]
116. Ruby JG, et al. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell. 2006; 127:1193–207. [PubMed: 17174894]
117. Taft RJ, et al. Small RNAs derived from snoRNAs. RNA. 2009; 15:1233–40. [PubMed: 19474147]
118. Ender C, et al. A human snoRNA with microRNA-like functions. Mol Cell. 2008; 32:519–28. [PubMed: 19026782]
119. Kawaji H, et al. Hidden layers of human small RNAs. BMC Genomics. 2008; 9:157. [PubMed: 18402656]
120. Haussecker D, et al. Human tRNA-derived small RNAs in the global regulation of RNA silencing. RNA. 2010; 16:673–95. [PubMed: 20181738]
121. Taft RJ, et al. Tiny RNAs associated with transcription start sites in animals. Nat Genet. 2009; 41:572–8. [PubMed: 19377478]
122. Taft RJ, et al. Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. Nat Struct Mol Biol. 2010; 17:1030–4. [PubMed: 20622877]
123. Taft RJ, Kaplan CD, Simons C, Mattick JS. Evolution, biogenesis and function of promoter-associated RNAs. Cell Cycle. 2009; 8:2332–8. [PubMed: 19597344]
124. Taft RJ, Hawkins PG, Mattick JS, Morris KV. The relationship between transcription initiation RNAs and CCCTC-binding factor (CTCF) localization. Epigenetics Chromatin. 2011; 4:13. [PubMed: 21813016]
125. Kapranov P, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science. 2007; 316:1484–8. [PubMed: 17510325]
126. Seila AC, et al. Divergent transcription from active promoters. Science. 2008; 322:1849–51. [PubMed: 19056940]
127. Preker P, et al. RNA exosome depletion reveals transcription upstream of active human promoters. Science. 2008; 322:1851–4. [PubMed: 19056938]
128. Han J, Kim D, Morris KV. Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. Proc Natl Acad Sci U S A. 2007; 104:12422–7. [PubMed: 17640892]
129. Wassarman KM, Zhang A, Storz G. Small RNAs in Escherichia coli. Trends Microbiol. 1999; 7:37–45. [PubMed: 10068996]
130. Gottesman S. Micros for microbes: non-coding regulatory RNAs in bacteria. Trends Genet. 2005; 21:399–404. [PubMed: 15913835]

131. Tucker BJ, Breaker RR. Riboswitches as versatile gene control elements. Curr Opin Struct Biol. 2005; 15:342–348. [PubMed: 15919195]

132. Winkler WC. Riboswitches and the role of noncoding RNAs in bacterial metabolic control. Curr Opin Chem Biol. 2005; 9:594–602. [PubMed: 16226486]

133. Mojica FJ, Diez-Villasenor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol Microbiol. 2000; 36:244–6. [PubMed: 10760181]

134. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol. 2005; 60:174–82. [PubMed: 15791728]

135. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005; 151:2551–61. [PubMed: 16079334]

136. Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology. 2005; 151:653–63. [PubMed: 15758212]

137. Barrangou R, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007; 315:1709–12. [PubMed: 17379808]

138. Brouns SJ, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008; 321:960–4. [PubMed: 18703739]

139. Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science. 2008; 322:1843–5. [PubMed: 19095942]

140. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009; 155:733–40. [PubMed: 19246744]

141. Hale CR, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell. 2009; 139:945–56. [PubMed: 19945378]

142. Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337:816–21. [PubMed: 22745249]

143. Mali P, et al. RNA-guided human genome engineering via Cas9. Science. 2013; 339:823–6. [PubMed: 23287722]

144. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013; 339:819–23. [PubMed: 23287718]

145. Wang H, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013; 153:910–8. [PubMed: 23643243]

146. Gilbert LA, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013; 154:442–51. [PubMed: 23849981]

147. Qi LS, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013; 152:1173–83. [PubMed: 23452860]

148. Perez-Pinera P, et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat Methods. 2013; 10:973–6. [PubMed: 23892895]

149. Cheng AW, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res. 2013; 23:1163–71. [PubMed: 23979020]

150. Hu J, et al. Direct activation of human and mouse Oct4 genes using engineered TALE and Cas9 transcription factors. Nucleic Acids Res. 2014

151. Seed KD, Lazinski DW, Calderwood SB, Camilli A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. Nature. 2013; 494:489–91. [PubMed: 23446421]

152. Mattick JS. Introns: evolution and function. Curr Opin Genet Dev. 1994; 4:823–31. [PubMed: 7888751]

153. Kapranov P, et al. Large-scale transcriptional activity in chromosomes 21 and 22. Science. 2002; 296:916–919. [PubMed: 11988577]

154. Okazaki Y, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature. 2002; 420:563–73. [PubMed: 12466851]

155. Rinn JL, et al. The transcriptional activity of human chromosome 22. Genes Dev. 2003; 17:529–40. [PubMed: 12600945]

156. Cheng J, et al. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science. 2005; 308:1149–54. [PubMed: 15790807]

157. Kapranov P, et al. Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. Genome Res. 2005; 15:987–97. [PubMed: 15998911]

158. Carninci P, et al. The transcriptional landscape of the mammalian genome. Science. 2005; 309:1559–63. [PubMed: 16141072]

159. Katayama S, et al. Antisense transcription in the mammalian transcriptome. Science. 2005; 309:1564–6. [PubMed: 16141073]

160. Birney E, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007; 447:799–816. [PubMed: 17571346]

161. Dunham I, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]

162. Rosenbloom KR, et al. ENCODE whole-genome data in the UCSC Genome Browser: update 2012. Nucleic Acids Res. 2012; 40:D912–7. [PubMed: 22075998]

163. Taft RJ, Pheasant M, Mattick JS. The relationship between non-protein-coding DNA and eukaryotic complexity. Bioessays. 2007; 29:288–99. [PubMed: 17295292]

164. Mercer TR, et al. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. Nat Biotechnol. 2012; 30:99–104. [PubMed: 22081020]

165. Djebali S, et al. Landscape of transcription in human cells. Nature. 2012; 489:101–8. [PubMed: 22955620]

166. Hawkins PG, Morris KV. Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. Transcr. 2010; 1:165–175.

167. Sheik Mohamed J, Gaughwin PM, Lim B, Robson P, Lipovich L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. RNA. 2010; 16:324–37. [PubMed: 20026622]

168. Venters BJ, Pugh BF. Genomic organization of human transcription initiation complexes. Nature. 2013; 502:53–8. [PubMed: 24048476]

169. Carninci P, et al. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet. 2006; 38:626–35. [PubMed: 16645617]

170. Huang R, et al. An RNA-Seq strategy to detect the complete coding and non-coding transcriptome including full-length imprinted macro ncRNAs. PLoS One. 2011; 6:e27288. [PubMed: 22102886]

171. Roberts A, Pachter L. RNA-Seq and find: entering the RNA deep field. Genome Med. 2011; 3:74. [PubMed: 22113004]

172. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009; 10:155–9. [PubMed: 19188922]

173. Frith MC, et al. The abundance of short proteins in the mammalian proteome. PLoS Genet. 2006; 2:e52. [PubMed: 16683031]

174. Gascoigne DK, et al. Pinstripe: a suite of programs for integrating transcriptomic and proteomic datasets identifies novel proteins and improves differentiation of protein-coding and non-coding genes. Bioinformatics. 2012; 28:3042–50. [PubMed: 23044541]

175. Banfai B, et al. Long noncoding RNAs are rarely translated in two human cell lines. Genome Res. 2012; 22:1646–57. [PubMed: 22955977]

176. Dinger ME, et al. NRED: a database of long noncoding RNA expression. Nucleic Acids Res. 2009; 37:D122–6. [PubMed: 18829717]

177. Guttman M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009; 458:223–7. [PubMed: 19182780]

178. Wahlestedt C. Natural antisense and noncoding RNA transcripts as potential drug targets. Drug Discov Today. 2006; 11:503–8. [PubMed: 16713901]

179. Khalil AM, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009; 106:11667–72. [PubMed: 19571010]

180. Ørom UA, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell. 2010; 143:46–58. [PubMed: 20887892]

181. Sauvageau M, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. Elife. 2013; 2:e01749. [PubMed: 24381249]

182. Furuno M, et al. Clusters of internally primed transcripts reveal novel long noncoding RNAs. PLoS Genet. 2006; 2:e37. [PubMed: 16683026]

183. van Bakel H, Nislow C, Blencowe BJ, Hughes TR. Most "dark matter" transcripts are associated with known genes. PLoS Biol. 2010; 8:e1000371. [PubMed: 20502517]

184. Clark MB, et al. The reality of pervasive transcription. PLoS Biol. 2011; 9:e1000625. [PubMed: 21765801]

185. Calin GA, et al. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell. 2007; 12:215–29. [PubMed: 17785203]

186. Tay SK, Blythe J, Lipovich L. Global discovery of primate-specific genes in the human genome. Proc Natl Acad Sci U S A. 2009; 106:12019–24. [PubMed: 19581580]

187. Lipovich L, et al. Developmental changes in the transcriptome of human cerebral cortex tissue: long noncoding RNA transcripts. Cereb Cortex. 2013

188. Necsulea A, et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. Nature. 2014; 505:635–40. [PubMed: 24463510]

189. Pheasant M, Mattick JS. Raising the estimate of functional human sequences. Genome Res. 2007; 17:1245–53. [PubMed: 17690206]

190. Pang KC, Frith MC, Mattick JS. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. Trends Genet. 2006; 22:1–5. [PubMed: 16290135]

191. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell. 2011; 147:1537–50. [PubMed: 22196729]

192. Smith MA, Gesell T, Stadler PF, Mattick JS. Widespread purifying selection on RNA structure in mammals. Nucleic Acids Res. 2013; 41:8220–36. [PubMed: 23847102]

193. Johnsson P, Lipovich L, Grander D, Morris KV. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. Biochim Biophys Acta. 2014; 1840:1063–71. [PubMed: 24184936]

194. Rinn JL, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007; 129:1311–23. [PubMed: 17604720]

195. Cawley S, et al. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. Cell. 2004; 116:499–509. [PubMed: 14980218]

196. Clark MB, et al. Genome-wide analysis of long noncoding RNA stability. Genome Res. 2012; 22:885–98. [PubMed: 22406755]

197. Ravasi T, et al. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. Genome Res. 2006; 16:11–9. [PubMed: 16344565]

198. Mercer TR, Dinger ME, Sunkin SM, Mehler MF, Mattick JS. Specific expression of long noncoding RNAs in the mouse brain. Proc Natl Acad Sci U S A. 2008; 105:716–21. [PubMed: 18184812]

199. Cabili MN, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 2011; 25:1915–27. [PubMed: 21890647]

200. Mattick JS, Taft RJ, Faulkner GJ. A global view of genomic information--moving beyond the gene and the master regulator. Trends Genet. 2010; 26:21–8. [PubMed: 19944475]

201. Johnsson P, et al. A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. Nat Struct Mol Biol. 2013; 20:440–6. [PubMed: 23435381]

202. Chooniedass-Kothari S, et al. The steroid receptor RNA activator is the first functional RNA encoding a protein. FEBS Lett. 2004; 566:43–7. [PubMed: 15147866]

203. Ashe HL, Monks J, Wijgerde M, Fraser P, Proudfoot NJ. Intergenic transcription and transinduction of the human beta-globin locus. Genes Dev. 1997; 11:2494–509. [PubMed: 9334315]
204. Mercer TR, et al. Expression of distinct RNAs from 3′ untranslated regions. Nucleic Acids Res. 2011; 39:2393–2403. [PubMed: 21075793]
205. Dinger ME, Gascoigne DK, Mattick JS. The evolution of RNAs with multiple functions. Biochimie. 2011; 93:2013–8. [PubMed: 21802485]
206. Mercer TR, et al. Regulated post-transcriptional RNA cleavage diversifies the eukaryotic transcriptome. Genome Res. 2010; 20:1639–50. [PubMed: 21045082]
207. Dinger ME, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res. 2008; 18:1433–45. [PubMed: 18562676]
208. Sunwoo H, et al. MEN ε/β nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res. 2009; 19:347–59. [PubMed: 19106332]
209. Pang KC, et al. Genome-wide identification of long noncoding RNAs in CD8+ T cells. J Immunol. 2009; 182:7738–48. [PubMed: 19494298]
210. Askarian-Amiri ME, et al. SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. RNA. 2011; 17:878–91. [PubMed: 21460236]
211. Hu W, Yuan B, Flygare J, Lodish HF. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. Genes Dev. 2011; 25:2573–8. [PubMed: 22155924]
212. Mercer TR, et al. Long noncoding RNAs in neuronal-glial fate specification and oligodendrocyte lineage maturation. BMC Neurosci. 2010; 11:14. [PubMed: 20137068]
213. Johnson R, et al. Regulation of neural macroRNAs by the transcriptional repressor REST. Rna. 2009; 15:85–96. [PubMed: 19050060]
214. Ng S-Y, Johnson R, Stanton LW. Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. Embo J. 2011 in press.
215. Takeda K, et al. Identification of a novel bone morphogenetic protein-responsive gene that may function as a noncoding RNA. J Biol Chem. 1998; 273:17079–85. [PubMed: 9642273]
216. Bussemakers MJ, et al. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res. 1999; 59:5975–9. [PubMed: 10606244]
217. Pasmant E, et al. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. Cancer Res. 2007; 67:3963–9. [PubMed: 17440112]
218. Wang F, Li X, Xie X, Zhao L, Chen W. UCA1, a non-protein-coding RNA up-regulated in bladder carcinoma and embryo, influencing cell growth and promoting invasion. FEBS Lett. 2008; 582:1919–27. [PubMed: 18501714]
219. Mourtada-Maarabouni M, Pickard MR, Hedge VL, Farzaneh F, Williams GT. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene. 2009; 28:195–208. [PubMed: 18836484]
220. Gupta RA, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010; 464:1071–6. [PubMed: 20393566]
221. Khaitan D, et al. The melanoma-upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion. Cancer Res. 2011; 71:3852–62. [PubMed: 21558391]
222. Kerin T, et al. A noncoding RNA antisense to moesin at 5p14.1 in autism. Sci Transl Med. 2012; 4:128ra40.
223. Nowacki M, et al. RNA-mediated epigenetic programming of a genome-rearrangement pathway. Nature. 2008; 451:153–8. [PubMed: 18046331]
224. Lopez de Silanes I, Stagno d'Alcontres M, Blasco MA. TERRA transcripts are bound by a complex array of RNA-binding proteins. Nat Commun. 2010; 1:33. [PubMed: 20975687]
225. Walter P, Blobel G. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. Nature. 1982; 299:691–8. [PubMed: 6181418]

226. Ji X, et al. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell. 2013; 153:855–68. [PubMed: 23663783]
227. Fox AH, Bond CS, Lamond AI. P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. Mol Biol Cell. 2005; 16:5304–15. [PubMed: 16148043]
228. Mao YS, Sunwoo H, Zhang B, Spector DL. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. Nat Cell Biol. 2011; 13:95–101. [PubMed: 21170033]
229. Tripathi V, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell. 2010; 39:925–38. [PubMed: 20797886]
230. Tripathi V, et al. Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. PLoS Genet. 2013; 9:e1003368. [PubMed: 23555285]
231. Sone M, et al. The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. J Cell Sci. 2007; 120:2498–506. [PubMed: 17623775]
232. Barry G, et al. The long non-coding RNA Gomafu is acutely regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing. Mol Psychiatry. 2013
233. Amaral PP, Mattick JS. Noncoding RNA in development. Mamm Genome. 2008; 19:454–92. [PubMed: 18839252]
234. Mourtada-Maarabouni M, Hedge VL, Kirkham L, Farzaneh F, Williams GT. Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (GAS5). J Cell Sci. 2008; 121:939–46. [PubMed: 18354083]
235. Young TL, Matsuda T, Cepko CL. The noncoding RNA taurine upregulated gene 1 is required for differentiation of the murine retina. Curr Biol. 2005; 15:501–12. [PubMed: 15797018]
236. Ginger MR, et al. A noncoding RNA is a potential marker of cell fate during mammary gland development. Proc Natl Acad Sci U S A. 2006; 103:5781–6. [PubMed: 16574773]
237. Kretz M, et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature. 2013; 493:231–5. [PubMed: 23201690]
238. Gutschner T, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. Cancer Res. 2013; 73:1180–9. [PubMed: 23243023]
239. Li L, et al. Targeted disruption of Hotair leads to homeotic transformation and gene derepression. Cell Rep. 2013; 5:3–12. [PubMed: 24075995]
240. Nakagawa S, Naganuma T, Shioi G, Hirose T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. J Cell Biol. 2011; 193:31–9. [PubMed: 21444682]
241. Ahituv N, et al. Deletion of ultraconserved elements yields viable mice. PLoS Biol. 2007; 5:e234. [PubMed: 17803355]
242. Derrien T, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012; 22:1775–89. [PubMed: 22955988]
243. Necsulea A, et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. Nature. 2014 epub ahead of print. 10.1038/nature12943
244. Lewejohann L, et al. Role of a neuronal small non-messenger RNA: behavioural alterations in BC1 RNA-deleted mice. Behav Brain Res. 2004; 154:273–89. [PubMed: 15302134]
245. Mattick JS, Gagen MJ. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. Mol Biol Evol. 2001; 18:1611–30. [PubMed: 11504843]
246. Mattick JS, Amaral PP, Dinger ME, Mercer TR, Mehler MF. RNA regulation of epigenetic processes. Bioessays. 2009; 31:51–9. [PubMed: 19154003]
247. Koziol MJ, Rinn JL. RNA traffic control of chromatin complexes. Curr Opin Genet Dev. 2010; 20:142–8. [PubMed: 20362426]
248. Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol. 2013; 20:300–7. [PubMed: 23463315]

249. Mercer TR, Mattick JS. Understanding the regulatory and transcriptional complexity of the genome through structure. Genome Res. 2013; 23:1081–8. [PubMed: 23817049]

250. Wassenegger M. RNA-directed DNA methylation. Plant Mol Biol. 2000; 43:203–20. [PubMed: 10999405]

251. Hall IM, et al. Establishment and maintenance of a heterochromatin domain. Science. 2002; 297:2232–2237. [PubMed: 12215653]

252. Volpe TA, et al. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science. 2002; 297:1833–7. [PubMed: 12193640]

253. Verdel A, et al. RNAi-mediated targeting of heterochromatin by the RITS complex. Science. 2004; 303:672–676. [PubMed: 14704433]

254. Kanhere A, et al. Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. Mol Cell. 2010; 38:675–88. [PubMed: 20542000]

255. Imamura T, et al. Non-coding RNA directed DNA demethylation of Sphk1 CpG island. Biochem Biophys Res Commun. 2004; 322:593–600. [PubMed: 15325271]

256. Morris KV, Santoso S, Turner AM, Pastori C, Hawkins PG. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. PLoS Genet. 2008; 4:e1000258. [PubMed: 19008947]

257. Yu W, et al. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. Nature. 2008; 451:202–6. [PubMed: 18185590]

258. Allo M, et al. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. Nat Struct Mol Biol. 2009; 16:717–24. [PubMed: 19543290]

259. Beltran M, et al. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev. 2008; 22:756–69. [PubMed: 18347095]

260. Morrissy AS, Griffith M, Marra MA. Extensive relationship between antisense transcription and alternative splicing in the human genome. Genome Res. 2011; 21:1203–12. [PubMed: 21719572]

261. Brockdorff N, et al. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell. 1992; 71:515–26. [PubMed: 1423610]

262. Brown CJ, et al. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell. 1992; 71:527–42. [PubMed: 1423611]

263. Meller VH, Wu KH, Roman G, Kuroda MI, Davis RL. roX1 RNA paints the X chromosome of male Drosophila and is regulated by the dosage compensation system. Cell. 1997; 88:445–57. [PubMed: 9038336]

264. Lee JT, Davidow LS, Warshawsky D. Tsix, a gene antisense to Xist at the X-inactivation centre. Nature Genet. 1999; 21:400–4. [PubMed: 10192391]

265. Sado T, Wang Z, Sasaki H, Li E. Regulation of imprinted X-chromosome inactivation in mice by Tsix. Development. 2001; 128:1275–86. [PubMed: 11262229]

266. Ripoche MA, Kress C, Poirier F, Dandolo L. Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. Genes Dev. 1997; 11:1596–604. [PubMed: 9203585]

267. Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature. 2002; 415:810–3. [PubMed: 11845212]

268. Thakur N, et al. An antisense RNA regulates the bidirectional silencing property of the Kcnq1 imprinting control region. Mol Cell Biol. 2004; 24:7855–62. [PubMed: 15340049]

269. Swiezewski S, Liu F, Magusin A, Dean C. Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature. 2009; 462:799–802. [PubMed: 20010688]

270. Nagano T, et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science. 2008; 322:1717–20. [PubMed: 18988810]

271. Mohammad F, et al. Kcnq1ot1/Lit1 noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region. Mol Cell Biol. 2008; 28:3713–28. [PubMed: 18299392]

272. Kotake Y, et al. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. Oncogene. 2011; 30:1956–62. [PubMed: 21151178]

273. Mohammad F, Mondal T, Guseva N, Pandey GK, Kanduri C. Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. Development. 2010; 137:2493–9. [PubMed: 20573698]

274. Di Ruscio A, et al. DNMT1-interacting RNAs block gene-specific DNA methylation. Nature. 2013; 503:371–6. [PubMed: 24107992]

275. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. Nat Struct Mol Biol. 2013; 20:1250–7. [PubMed: 24077223]

276. Li L, et al. Targeted Disruption of Hotair Leads to Homeotic Transformation and Gene Derepression. Cell Rep. 2013

277. Tsai MC, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science. 2010; 329:689–93. [PubMed: 20616235]

278. Zhang H, et al. Long noncoding RNA-mediated intrachromosomal interactions promote imprinting at the Kcnq1 locus. J Cell Biol. 2014; 204:61–75. [PubMed: 24395636]

279. Sanchez-Herrero E, Akam M. Spatially ordered transcription of regulatory DNA in the bithorax complex of *Drosophila*. Development. 1989; 107:321–9. [PubMed: 2632227]

280. Bae E, Calhoun VC, Levine M, Lewis EB, Drewell RA. Characterization of the intergenic RNA profile at abdominal-A and Abdominal-B in the Drosophila bithorax complex. Proc Natl Acad Sci USA. 2002; 99:16847–16852. [PubMed: 12481037]

281. Jones EA, Flavell RA. Distal enhancer elements transcribe intergenic RNA in the IL-10 family gene cluster. J Immunol. 2005; 175:7437–46. [PubMed: 16301651]

282. Petruk S, et al. Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference. Cell. 2006; 127:1209–21. [PubMed: 17174895]

283. Feng J, et al. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev. 2006; 20:1470–1484. [PubMed: 16705037]

284. Kim TK, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010; 465:182–7. [PubMed: 20393465]

285. Sanchez-Elsner T, Gou D, Kremmer E, Sauer F. Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. Science. 2006; 311:1118–23. [PubMed: 16497925]

286. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res. 2009; 19:1732–41. [PubMed: 19687145]

287. Nahkuri S, Taft RJ, Mattick JS. Nucleosomes are preferentially positioned at exons in somatic and sperm cells. Cell Cycle. 2009; 8:3420–4. [PubMed: 19823040]

288. Schwartz S, Meshorer E, Ast G. Chromatin organization marks exon-intron structure. Nat Struct Mol Biol. 2009; 16:990–5. [PubMed: 19684600]

289. Spies N, Nielsen CB, Padgett RA, Burge CB. Biased chromatin signatures around polyadenylation sites and exons. Mol Cell. 2009; 36:245–54. [PubMed: 19854133]

290. Tilgner H, et al. Nucleosome positioning as a determinant of exon recognition. Nat Struct Mol Biol. 2009; 16:996–1001. [PubMed: 19684599]

291. Luco RF, et al. Regulation of alternative splicing by histone modifications. Science. 2010; 327:996–1000. [PubMed: 20133523]

292. Mercer TR, et al. DNase I-hypersensitive exons colocalize with promoters and distal regulatory elements. Nat Genet. 2013; 45:852–9. [PubMed: 23793028]

293. Willingham AT, et al. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science. 2005; 309:1570–3. [PubMed: 16141075]

294. Carrieri C, et al. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature. 2012; 491:454–7. [PubMed: 23064229]

295. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. Nature. 2011; 470:284–8. [PubMed: 21307942]

296. Mattick JS. RNA as the substrate for epigenome-environment interactions: RNA guidance of epigenetic processes and the expansion of RNA editing in animals underpins development, phenotypic plasticity, learning, and cognition. Bioessays. 2010; 32:548–52. [PubMed: 20544741]

297. Mattick JS. The central role of RNA in human development and cognition. FEBS Lett. 2011; 585:1600–16. [PubMed: 21557942]

298. Muotri AR, et al. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature. 2005; 435:903–10. [PubMed: 15959507]

299. Coufal NG, et al. L1 retrotransposition in human neural progenitor cells. Nature. 2009; 460:1127–31. [PubMed: 19657334]

300. Muotri AR, et al. L1 retrotransposition in neurons is modulated by MeCP2. Nature. 2010; 468:443–6. [PubMed: 21085180]

301. Faulkner GJ, et al. The regulated retrotransposon transcriptome of mammalian cells. Nat Genet. 2009; 41:563–71. [PubMed: 19377475]

302. Feschotte C. Transposable elements and the evolution of regulatory networks. Nat Rev Genet. 2008; 9:397–405. [PubMed: 18368054]

303. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell. 2008; 134:25–36. [PubMed: 18614008]

304. Brosius J. The contribution of RNAs and retroposition to evolutionary novelties. Genetica. 2003; 118:99–116. [PubMed: 12868601]

305. Krull M, Brosius J, Schmitz J. Alu-SINE exonization: en route to protein-coding function. Mol Biol Evol. 2005; 22:1702–11. [PubMed: 15901843]

306. Cordaux R, Udit S, Batzer MA, Feschotte C. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. Proc Natl Acad Sci U S A. 2006; 103:8101–6. [PubMed: 16672366]

307. Kelley D, Rinn J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. Genome Biol. 2012; 13:R107. [PubMed: 23181609]

308. Kapusta A, et al. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. PLoS Genet. 2013; 9:e1003470. [PubMed: 23637635]

309. Czerwoniec A, et al. MODOMICS: a database of RNA modification pathways. 2008 update. Nucleic Acids Res. 2009; 37:D118–21. [PubMed: 18854352]

310. Cantara WA, et al. The RNA Modification Database, RNAMDB: 2011 update. Nucleic Acids Res. 2011; 39:D195–201. [PubMed: 21071406]

311. Motorin Y, Lyko F, Helm M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic Acids Res. 2010; 38:1415–30. [PubMed: 20007150]

312. Abbasi-Moheb L, et al. Mutations in NSUN2 cause autosomal-recessive intellectual disability. Am J Hum Genet. 2012; 90:847–55. [PubMed: 22541559]

313. Meyer KD, et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell. 2012; 149:1635–46. [PubMed: 22608085]

314. Jia G, Fu Y, He C. Reversible RNA adenosine methylation in biological regulation. Trends Genet. 2013; 29:108–15. [PubMed: 23218460]

315. Saletore Y, et al. The birth of the Epitranscriptome: deciphering the function of RNA modifications. Genome Biol. 2012; 13:175. [PubMed: 23113984]

316. Saletore Y, Chen-Kiang S, Mason CE. Novel RNA regulatory mechanisms revealed in the epitranscriptome. RNA Biol. 2013; 10:342–6. [PubMed: 23434792]

317. Brosnan CA, et al. Nuclear gene silencing directs reception of long-distance mRNA silencing in Arabidopsis. Proc Natl Acad Sci U S A. 2007; 104:14741–6. [PubMed: 17785412]

318. Dinger ME, Mercer TR, Mattick JS. RNAs as extracellular signaling molecules. J Mol Endocrinol. 2008; 40:151–9. [PubMed: 18372404]

319. Alleman M, et al. An RNA-dependent RNA polymerase is required for paramutation in maize. Nature. 2006; 442:295–8. [PubMed: 16855589]

320. Rassoulzadegan M, et al. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. Nature. 2006; 441:469–74. [PubMed: 16724059]

321. Chandler VL. Paramutation: from maize to mice. Cell. 2007; 128:641–5. [PubMed: 17320501]
322. Nadeau JH. Transgenerational genetic effects on phenotypic variation and disease risk. Hum Mol Genet. 2009; 18:R202–10. [PubMed: 19808797]
323. Buckley BA, et al. A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature. 2012; 489:447–51. [PubMed: 22810588]
324. Herbert A, Rich A. RNA processing in evolution: The logic of soft-wired genomes. Ann N Y Acad Sci. 1999; 870:119–32. [PubMed: 10415478]
325. Herbert A, Rich A. RNA processing and the evolution of eukaryotes. Nature Genet. 1999; 21:265–9. [PubMed: 10080177]
326. Mattick JS. Has evolution learnt how to learn? EMBO Rep. 2009; 10:665. [PubMed: 19568258]
327. Mattick JS, Makunin IV. Non-coding RNA. Hum Mol Genet. 2006; 15:R17–29. [PubMed: 16651366]
328. Gingeras TR. Origin of phenotypes: genes and transcripts. Genome Res. 2007; 17:682–90. [PubMed: 17567989]
329. Mattick JS. Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. Bioessays. 2003; 25:930–9. [PubMed: 14505360]
330. Chamary JV, Parmley JL, Hurst LD. Hearing silence: non-neutral evolution at synonymous sites in mammals. Nat Rev Genet. 2006; 7:98–108. [PubMed: 16418745]
331. Lin MF, et al. Locating protein-coding sequences under selection for additional, overlapping functions in 29 mammalian genomes. Genome Res. 2011; 21:1916–28. [PubMed: 21994248]
332. Birnbaum RY, et al. Coding exons function as tissue-specific enhancers of nearby genes. Genome Res. 2012; 22:1059–68. [PubMed: 22442009]
333. Stergachis AB, et al. Exonic transcription factor binding directs codon choice and affects protein evolution. Science. 2013; 342:1367–72. [PubMed: 24337295]
334. 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. 2006; 312:1653–5. [PubMed: 16778056]
335. Capel B, et al. Circular transcripts of the testis-determining gene Sry in adult mouse testis. Cell. 1993; 73:1019–30. [PubMed: 7684656]
336. Hansen TB, et al. Natural RNA circles function as efficient microRNA sponges. Nature. 2013; 495:384–8. [PubMed: 23446346]
337. Memczak S, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013; 495:333–8. [PubMed: 23446348]
338. Zhang Y, et al. Circular intronic long noncoding RNAs. Mol Cell. 2013; 51:792–806. [PubMed: 24035497]
339. Wang J, et al. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. Nucleic Acids Res. 2010; 38:5366–83. [PubMed: 20423907]
340. Poliseno L, et al. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature. 2010; 465:1033–8. [PubMed: 20577206]
341. Wan Y, et al. Genome-wide measurement of RNA folding energies. Mol Cell. 2012; 48:169–81. [PubMed: 22981864]
342. Arteaga-Vazquez MA, Chandler VL. Paramutation in maize: RNA mediated trans-generational gene silencing. Curr Opin Genet Dev. 2010; 20:156–63. [PubMed: 20153628]
343. Sarkies P, Miska EA. Molecular biology. Is there social RNA? Science. 2013; 341:467–8. [PubMed: 23908213]
344. Weinberg MS, et al. The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells. RNA. 2006; 12:256–62. [PubMed: 16373483]
345. Janowski BA, et al. Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. Nat Struct Mol Biol. 2006; 13:787–92. [PubMed: 16936728]
346. Ling J, Baibakov B, Pi W, Emerson BM, Tuan D. The HS2 enhancer of the beta-globin locus control region initiates synthesis of non-coding, polyadenylated RNAs independent of a cis-linked globin promoter. J Mol Biol. 2005; 350:883–96. [PubMed: 15979088]

347. Li W, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 2013; 498:516–20. [PubMed: 23728302]

348. Gumireddy K, et al. Identification of a long non-coding RNA-associated RNP complex regulating metastasis at the translational step. EMBO J. 2013; 32:2672–84. [PubMed: 23974796]

349. Watson JD, Crick FH. Genetical implications of the structure of deoxyribonucleic acid. Nature. 1953; 171:964–7. [PubMed: 13063483]

350. Holmes DS, Mayfield JE, Sander G, Bonner J. Chromosomal RNA: its properties. Science. 1972; 177:72–4. [PubMed: 5041779]

351. Brannan CI, Dees EC, Ingram RS, Tilghman SM. The product of the H19 gene may function as an RNA. Mol Cell Biol. 1990; 10:28–36. [PubMed: 1688465]

352. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double stranded RNA in Caenorhabditis elegans. Nature. 1998; 391:806–811. [PubMed: 9486653]

353. Hamilton AJ, Baulcombe DC. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science. 1999; 286:950–2. [PubMed: 10542148]

354. Elbashir SM, et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 2001; 411:494–498. [PubMed: 11373684]

355. Mattick JS. Non-coding RNAs: the architects of eukaryotic complexity. EMBO Rep. 2001; 2:986–91. [PubMed: 11713189]

356. Liu J, et al. Argonaute2 is the catalytic engine of mammalian RNAi. Science. 2004; 305:1437–41. [PubMed: 15284456]

357. Li W, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 2013

358. Shamovsky I, Gershon D. Novel regulatory factors of HSF-1 activation: facts and perspectives regarding their involvement in the age-associated attenuation of the heat shock response. Mech Ageing Dev. 2004; 125:767–75. [PubMed: 15541771]

359. Pink RC, et al. Pseudogenes: pseudo-functional or key regulators in health and disease? RNA. 2011; 17:792–8. [PubMed: 21398401]

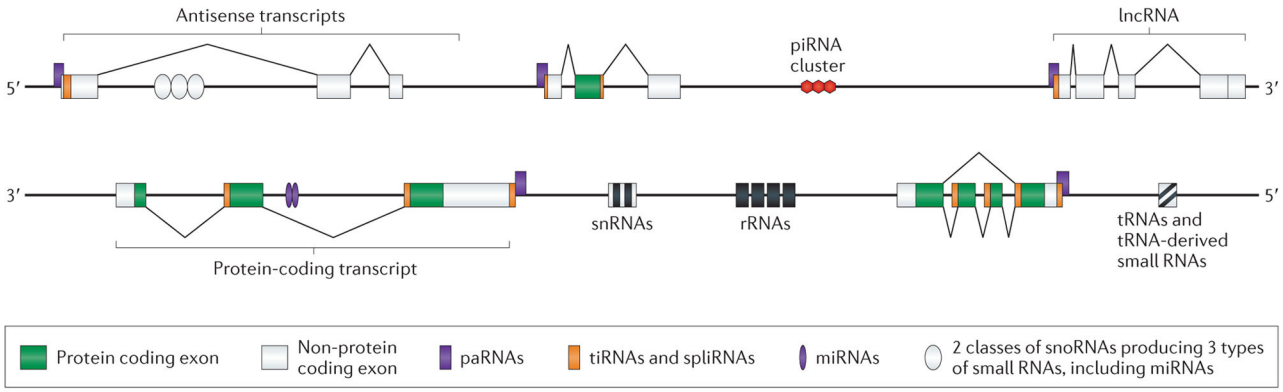

Figure 1. Complex expression of the genome and examples of non-coding RNA expression

Graphical representation of the mammalian transcriptional landscape with genes expressing rRNAs, tRNAs, snRNAs, snoRNAs, various protein coding and non-coding transcripts (mRNAs and lncRNAs), as well as small regulatory RNAs including miRNAs, piRNAs, tiRNAs and spliRNAs, snoRNA-derived small RNAs, and tRNA-derived small RNAs.

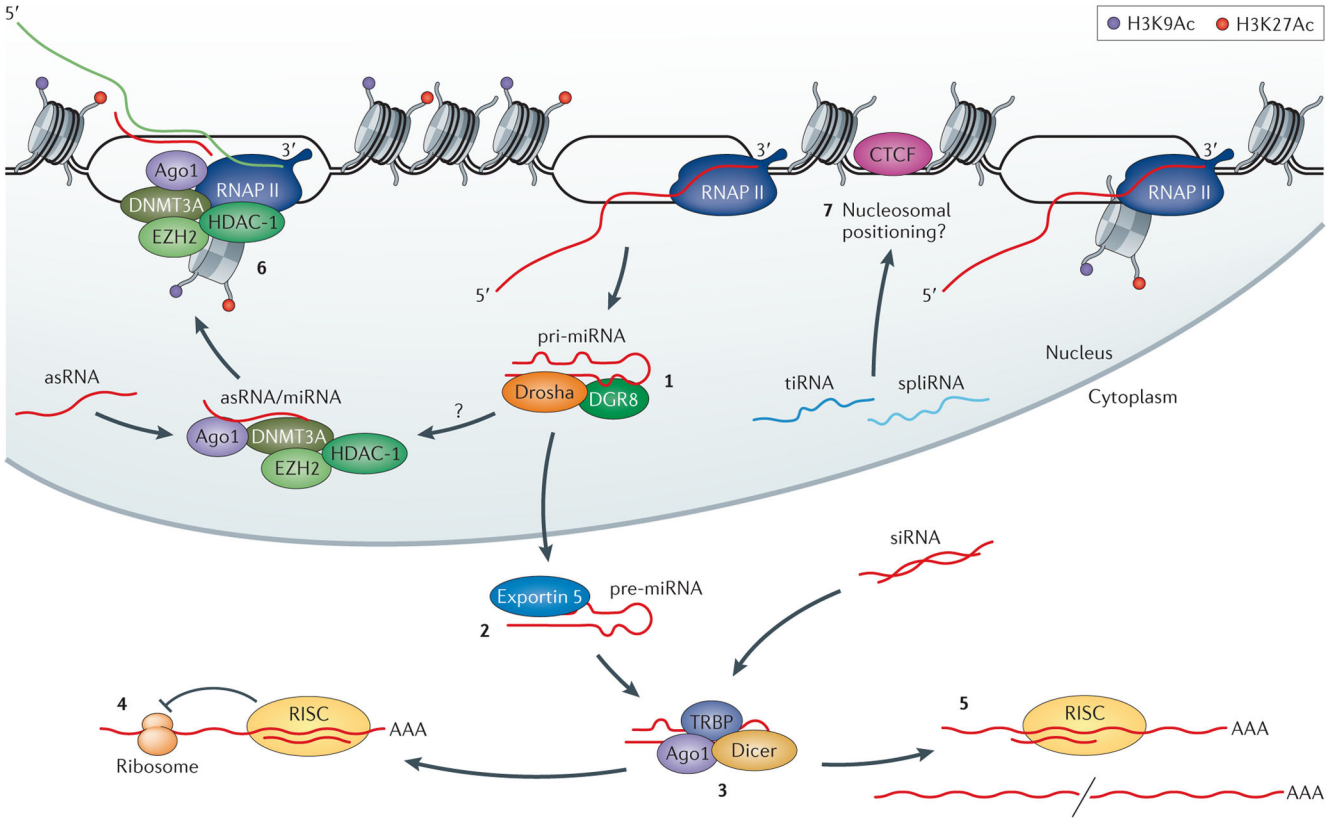

Figure 2. Functional pathways of small regulatory RNAs

(**A**) miRNA precursors are expressed as stem-loop structures[75], which (**B**) interact with Drosha[76] and DGR8, where they are processed then exported from the nucleus by Exportin 5. (**C**) These transcripts are further processed by Dicer to small (21–23 nt) dsRNAs, one strand of which is loaded into AGO component of the RNA-induced silencing complex (RISC). Exogenously introduced siRNAs can also be processed by RISC. Either the endogenous miRNA or exogenously added siRNAs can then to target (**D**) the repression of translation and/or (**E**) cleavage of homology containing transcripts[81,82]. Some small RNAs are functional in the nucleus. (**F**) Exogenously introduced small antisense RNAs (asRNAs) can target epigenetic silencing of targeted loci[88,344,345], a pathway that miRNAs may also utilize in the nucleus[92]. (**G**) tiRNAs and spliRNAs[121,122] are also expressed through an unknown pathway that may involve RNAPII backtracking and TFIIS cleavage[123], with the tiRNAs shown to modulate CTCF chromatin localization and to be associated with nucleosome position[124].

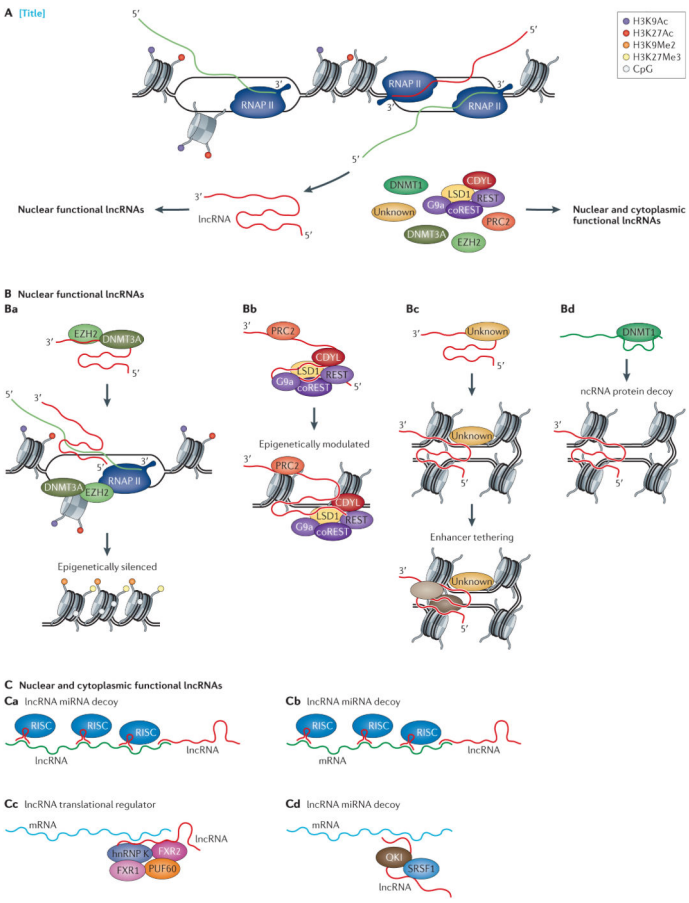

Figure 3. Various roles for lncRNAs in cellular regulation

(**A**) Long non-coding RNAs are expressed from many loci in the genome, sense and antisense, intronic, overlapping and intergenic with respect to nearby protein-coding loci, and function both in *cis* and *trans*. (**B–E**) Some lncRNAs interact with proteins to control the access of chromatin to cellular components and/or guide epigenetic regulatory complexes to target loci resulting in both (**B**) transcriptional suppression[201] and (**C**) activation or suppression (bimodal control)[194]. Proteins involved in chromatin modification such as DNMT3a, EZH2 and PRC2 complexes have been associated with epigenetic targeted lncRNA regulation[194,201,277]. (**D**) Some lncRNAs function to tether distal enhancer elements with their promoters[346,347]. (**E**) LncRNAs can also function by binding proteins to sequester them away from their sites of action (decoy lncRNAs)[274] while other lncRNAs can interact with each other and/or function to sequester small regulatory RNAs such as miRNAs and therefore RISC targeting complexes away from protein-coding mRNAs[201,339,340]. (**G**) LncRNAs can also act as translational inhibitors by binding and sequestering mRNAs away from the translational machinery[348] while other lncRNAs (**H**) appear to regulate splicing[232].

Table 1

Timeline Table:

1941	One gene-one enzyme[2]
1953	Double helical structure of DNA described[349]
1958	Central dogma proposed by Francis Crick[6]
1961	mRNA confirmed as intermediate between protein and DNA[9]
1961	Jacob and Monod speculate that the *lac* repressor is an RNA[10]
1966	Discovery of heterogenous nuclear RNA[27]
1969	Model proposed for RNA acting in intermediate fashion in gene regulation[28]
1972	hnRNAs, chromosomal RNAs shown to be functional without making protein[350]
1977	Intron ncRNA elements defined[34,35]
1982–3	Self Splicing catalytic RNAs[40,41]
1989	Transgene silencing observed in plants[71,72]
1990	Transgene silencing linked to antisense RNA[73]
1990	H19 ncRNA discovered[351]
1992	Xist ncRNA discovered[261, 262]
1993	Lin-4 miRNA discovered[46]
1994	Regulatory RNAs proposed to be central to animal evolution and development[152]
1994	RNA directed DNA methylation observed in plants[74]
1998	RNA interference described in plants[69] and animals[352]
1999	Tsix, antisense transcript to Xist described[264]
1999	Small RNA required for PTGS in plants[353]
2000	Let-7 miRNA discovered[47]
2001	Dicer described involved in RNAi[77]
2001	RNAi (PTGS) found functional in human cells[354]
2001	Regulatory RNA networks proposed to control epigenetic processes[245,355]
2002	First reports of large numbers of noncoding RNAs in animals[153–155]
2002	AIR antisense RNA involved in imprinting[267]
2003	Drosha described in miRNA processing[76]
2004	Small RNA shown to epigenetically control transcription (TGS) in human cells[88]
2004	Argonuate 2 directs catalysis in RNAi in mammals[356]
2005	piRNAs described[100]
2005	Confirmation of large numbers of long noncoding RNAs in mammals[156,158,159]
2005	~70% of sense transcripts have antisense counterparts, some show function[159]
2005	Discovery of the CRISPR system of bacterial RNA-based defence[134–136]
2006	Antisense RNA TGS shown to require DNMT3a, EZH2, HDAC-1[344]
2006	Argonautes 1 and 2 found involved in RNA-directed TGS in human cells[86,87]
2006	ncRNAs involved in trithorax regulation[285].
2007	HOTAIR shown to play a role in development and associate with polycomb[194]
2008	Long antisense RNAs found to epigenetically regulate sense counterparts[256,257]

2008	LncRNAs shown to interact with trithorax and activated chromatin[207]
2008	Hundreds of lncRNAs shown to have specific expression in brain[198]
2009	tiRNAs reported at transcription start sites in mammals[121]
2009	PRC2 found to interact with a large number of lncRNAs[179]
2009	Long antisense RNA shown to direct vernalization in plants[269]
2010	Pseudogene lncRNAs found to regulate protein-coding genes[166,340]
2012	ENCODE reports ~80% of the genome is transcribing ncRNAs[162]
2013	Enhancer RNAs shown in oestrogen-dependent transcriptional activation[357]