LncRNAs: genetic and epigenetic effects in plants

Elif Karlika,c, Sule Arıb and Nermin Gozukirmiziβ,c

αDepartment of Biotechnology Institute of Graduate Studies in Science and Engineering, Istanbul University, Istanbul, Turkey; βDepartment of Molecular Biology and Genetics Faculty of Science, Istanbul University, Istanbul, Turkey; γDepartment of Molecular Biology and Genetics Faculty of Science, İstinye University, Istanbul, Turkey

ABSTRACT
Long non-coding RNAs (lncRNAs) transcribed from the eukaryotic genome play important roles in essential biological processes, transcriptional and post-transcriptional gene regulation. LncRNAs act both in the nucleus and in the cytoplasm, mostly in association with chromatin in the nucleus. LncRNAs appear to be important regulators of gene expression, gene regulation and genome stability. This review outlines the major types of plant lncRNAs, their genetic and epigenetic effects with a focus on plant lncRNA instances, and discusses the recent advances in our understanding of their mechanism of action.

ARTICLE HISTORY
Received 28 July 2018
Accepted 7 February 2019

KEYWORDS
Long non-coding RNAs; epigenetic regulation; plant biotechnology

Introduction

One of the greatest challenges in modern day genomics is to fully reveal the functions of genes and genetic elements. The recent approaches, including whole-genome and RNA-sequencing (RNA-seq) studies and many biochemical techniques such as native RNA immunoprecipitation (nRIP), cross-linking and immunoprecipitation (CLIP), RNA immunoprecipitation, the RNA pull-down assay, chromatin isolation by RNA purification (ChIRP), capture hybridization analysis of RNA targets (CHART) and RNA antisense purification (RAP), have revealed that the transcription pattern in eukaryotes is much more complex than previously appreciated [1–9]. Especially, the human genome is pervasively transcribed, but less than 2% of these outputs are putatively functional RNAs that encode for proteins [10]. They mostly remain uncharacterized and make up the ‘dark matter’ of the genome. Long non-coding RNAs (lncRNAs) have been revealed to emerge from these dark regions of the genome [11]. For a long time, these transcripts remained unknown and considered to be functionless due to their lack of any protein-coding potential and presented as transcriptional noise (reviewed in [12]). However, over the past ten years, thousands of novel non-coding RNAs (ncRNAs), which involve no open reading frame (ORF), have been identified in many organisms, including human, animals and plants by using computational and experimental approaches, indicating that they are actually functional molecules and key players in many biological processes [11–15].

ncRNAs are grouped into two classes: housekeeping ncRNAs and regulatory ncRNAs. Housekeeping ncRNAs, including ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs), are usually expressed constitutively. The regulatory ncRNAs are classified into two sub-classes: (a) small RNAs (sRNAs), including miRNAs and siRNAs, and (b) lncRNAs, including long intronic ncRNAs and long intergenic ncRNAs [16–18]. Additionally, circRNAs, which have been recently discovered as a new class of ncRNAs, are classified into five groups: (1) exonic circRNAs, (2) intronic circRNAs, (3) UTR circRNAs, (4) intergenic circRNAs and (5) other circRNAs derived from two or more genes [19].

In plants, lncRNAs are involved in diverse biological processes such as phosphate homeostasis, flowering, photomorphogenesis, stress response [20] and fertility [21]. As previously reviewed in greater detail [20], they also play important roles to: (1) process into shorter ncRNAs [22], (2) act as decoys for the target mimics of miRNAs [23, 24], (3) repress histone modifying activities and direct epigenetic silencing through interaction with specific chromatin domains [25–28], (4) function as molecular cargo for protein re-localization [29, 30] and (5) regulate of post-translational processes...
via protein modifications and protein–protein interactions [31].

lncRNAs are categorized according to various properties, including transcript length, sequence and structure conservation, genomic location, functions exerted on DNA or RNA [32, 33]. Based on length, the typical threshold value is 200 nt. lncRNAs shorter than 200 nt are termed small ncRNAs, and those longer than 200 bases are classified as lncRNAs [34]. However, the most common criterion for discrimination of lncRNAs is their genomic location. lncRNAs are grouped into six categories [32]: (a) Sense lncRNAs, which are transcribed from the sense strand of protein-coding genes and overlapping transcripts, comprise a coding gene within an intron on the same strand. (b) Intron transcripts lying within the introns of a coding gene, which do not have exon–exon overlapping, are defined as sense intronic lncRNAs. (c) Antisense lncRNAs are transcribed from the antisense strand and they intersect any exon of a protein-coding locus on the opposite strand. (d) Long intervening non-coding RNAs (lincRNAs) with a length >200 bp are also called as long ‘intergenic’ non-coding RNAs which do not overlap protein-coding exons and reside within the genomic interval between two genes [35]. (e) Bidirectional lncRNAs are expressed within 1 kb of promoters in the opposite direction from the neighbouring protein-coding gene. (f) Enhancer lncRNA (elncRNA or eRNA), which are generally considered as <2 kb, are transcribed from enhancer regions of the genome and might contribute to enhancer function [36, 37].

The majority of lncRNAs are transcribed by RNA polymerase II (Pol II) in eukaryotes such as mRNAs which are assumed as stable ‘typical’ or classic lncRNAs [38]. Studies demonstrated that, unlike mRNAs, lncRNAs are primarily found to be within the nucleus. Also, most of them can be in very low quantity in specific patterns and in diverse tissue types [34, 39]. However, some lncRNAs are transcribed by RNA polymerase III (Pol III), which was previously believed to transcribe only tRNA and SS rRNA [40]. Additionally, lncRNAs are transcribed by RNA polymerase IV (Pol IV) and V (Pol V) in plants [41, 42]. The lncRNAs transcribed by plant-specific Pol V are involved in the process of RNA-directed DNA methylation (RdDM), which is a plant-specific de novo DNA methylation mechanism that requires long lncRNA. However, the siRNAs, which are combined into ARGONAUTE (AGO) (binding to chromatin), pair with lncRNAs transcribed by Pol V, which facilitates the recruitment of AGO to define the particular genomic loci [43, 44].

**Historical development of IncRNA research**

In the 1990s and early 2000s, before the advent of gene knockdown techniques became widely available, the functions of IncRNAs were studied by using classical gene targeting strategies and genetic approaches. Xist and Tsix are two overlapping antisense-transcribed IncRNAs, which are involved in X chromosome inactivation in female mammals [45–47]. Additionally, in *Drosophila*, the functionally redundant roX1 and roX2, which regulate sex chromosome dosage compensation, were initially demonstrated by researchers in the late 1990s [48]. Moreover, Gt12, Airm and Kcnq1ot1 lncRNAs have been shown to be transcribed from imprinted genomic loci [49–51], and these studies suggest the concept of chromatin associating IncRNAs which regulate epigenetic gene expression, serving as the ‘Rosetta Stone’ for recent analyses of IncRNAs, as reviewed by Nakagawa [52]. The advent of the gene knockdown technique and high-throughput genome-wide gene expression analyses facilitated the study of candidate IncRNAs with differential expression patterns in particular tissues or cells.

*GmENOD40* (soybean), *MtENOD40* (*Medicago truncatula*), *TPS11* (tomato), *OsPI1* (rice), *LDMAR* (rice) lncRNAs, which are associated with diverse biological processes, were identified and their physiological functions were initially demonstrated in plants [53–56]. For instance, *ENOD40* IncRNAs, which are considered as ‘riboregulators’ residing in the cytoplasm associated with the growth control and differentiation, have been identified in different plant species such as *Glycine max*, *Medicago truncatula*, *Medicago sativa*, etc. [53]. An increasing number of plant IncRNA studies have contributed insight into the diverse biological roles of IncRNAs in different plant species. This review focuses on the genetic and epigenetic effects of IncRNAs.

**Genetic effects of IncRNAs**

The functions of IncRNAs are often deduced from their subcellular localization, i.e. in the nucleus, the nucleolus, the cytoplasm or in some cellular compartments. lncRNAs located in the nucleus and/or the cytoplasm were confirmed by our previous study via fluorescence *in situ* hybridization (FISH) in barley as shown in Figure 1 [unpublished data]. Other studies indicate that some IncRNAs may act as ribozymes or riboswitches at the RNA level [57]. However, IncRNAs mostly function as ribonucleoprotein particles (RNPs). In turn, RNPs may contain a different number of RNA partners, including snRNA, mRNA, telomerase RNA, etc. [58]. Many DNA-binding proteins involved in
eukaryotic transcription are also able to bind lncRNAs in vitro and in vivo [59, 60].

LncRNAs function in either cis or trans position by complementary sequence or homology with RNAs or DNA, and/or structure, by forming molecular frames and scaffolds for assembly of macromolecular complexes. Additionally, lncRNAs perform significant roles in gene regulation at different levels such as regulating DNA-binding activity by phosphorylation. Some lncRNAs regulate gene expression via diverse post-transcriptional mechanisms, although the regulation of gene expression is mostly controlled at the transcriptional level [61]. For instance, lncRNAs can act as decoy which inhibits the influence of regulatory proteins. In Arabidopsis, some lncRNAs have been demonstrated to interact with miRNAs as competitors by mimicking their targets. This type of lncRNAs serve as miRNA target mimics for inhibition of miRNA activity. The best known lncRNA as decoy for miRNA target mimicry in Arabidopsis is The Induced by Phosphate Starvation1 (IPS1) lncRNA, which competes with miR399 and binds miR399 to inhibit the degradation of PHO2 (PHOSPHATE2) [62]. Another lncRNA which is defined as a decoy of Arabidopsis lncRNA is Alternative splicing competitor lncRNA (ASCO-lncRNA), which binds the alternative splicing regulators and nuclear speckle RNA-binding proteins to regulate plant development by changing the alternative splicing patterns [63]. Recently, Cho and Paszkowski [64] discovered the retrotransposon-derived transcript MIKKI, which comprises multiple introns and has low coding potential. The fourth intron of MIKKI originates from an independent family of retrotransposons; thus, the splicing of this intron produces a binding site for miR171 at the exon–exon junction [64]. However, the cleavage of MIKKI mRNAs does not occur by miR171, so the miR171-binding site of MIKKI is not perfectly complementary to its cognate miRNA, there being two mismatches in the cleavage positions. The cleavage activity of miRNA is attenuated by such mismatches, which are considered a landmark of miRNA target mimicry [62, 64–66]. Some MIKKI knock-out rice plants lacking the target mimicking sequence show higher miR171 levels, whereas over-expression of MIKKI is associated with down-regulation of miR171 in rice roots [64, 67].

LncRNAs can directly regulate Pol II by promoting the phosphorylation of transcription factors (TFs). The phosphorylation of TFs is controlled by regulating their DNA-binding activity via lncRNAs [68]. Transcription initiation and elongation are regulated through several eukaryotic lncRNAs by control of RNA Pol II pausing, functioning via transcriptional interference and as scaffolds recruiting chromatin remodelers (reviewed in [61]). They play important roles and affect the chromatin topology and nuclear organization [69]. Wang et al. [70] demonstrated that trans-acting lncRNA HID1 interacts with the chromatin of PIF3, which encodes a TF, and represses its transcription in Arabidopsis. Recently, CDF5 LONG NONCODING RNA (FLORE), a circadian-regulated lncRNA, which is a natural antisense transcript (NAT) of CDF5, has been identified and its functional characterization were carried out by Henriques et al. [71]. In Arabidopsis, FLORE and CDF5 antiphase expression reflects mutual inhibition in a similar way to frq/qrf, which is an lncRNA encoding central oscillator component that is proposed as modulator of core clock function in fungi. Moreover, Henriques et al. [71] showed the CDF5/FLORE NAT pair forms an additional circadian regulatory module.

Figure 1. FISH analysis of AK363461 and AK370506 barley lncRNAs. DAPI-stained chromatin in the nucleoplasm is indicated with blue fluorescence. Regions hybridized with TRITC-labelled probes are indicated as red fluorescence. (a) Display of AK363461 with four signals; (b) Display of AK370506 with one signal. Scale bar = 2 μm.
with conserved (mutual inhibition) and unique (function in \textit{trans}) features. The unique features of CDF5/FLORE NAT pair provide ability to fine-tune its own circadian oscillation, and consequently, adjusting the onset of flowering to favourable environmental conditions.

Additionally, elncRNAs, which are clearly distinguishable from the canonical lncRNAs, perform important roles in gene activation, and they function as \textit{cis}-acting scaffolds to recruit co-activator complexes. ElncRNAs regulate the chromatin topology by mediating the chromosome looping between enhancer and promoter regions (Figure 2) [72, 73]. Some studies have demonstrated that at least some elncRNAs play key roles in target gene expression. For example, the knockdown of several elncRNAs resulted in decreased expression of nearby target genes [74–76]. In a plasmid-based reporter system, the artificial tethering of elncRNAs upstream of a minimal promoter enhanced the reporter gene expression, and these results were found to be consistent with the proposed role of elncRNAs in transcriptional activation. However, the critical determinants of the activating function of elncRNAs, including sequence- or strand-specificity are not clear [74, 77].

**Epigenetic effects of lncRNAs**

LncRNAs in plants can function during tissue development, sexual reproduction and in response to external stimuli such as drought, salinity, heat stress and infections [9, 78–81]. Transcriptional and posttranscriptional regulation of gene expression by lncRNAs is only recently beginning to be recognized, together with revealing the molecular mechanisms underlying this type of regulation. Studies indicated that lncRNAs can play important roles in gene regulation at different levels (reviewed in [61]).

As we mentioned above, some lncRNAs can act as decoys to prevent the influence of regulatory proteins to DNA or RNA by mimicking their targets. The mechanism is generally used for the inhibition of miRNA activity called ‘target mimicry’. In Arabidopsis, lncRNAs including \textit{IPS1} and alternative splicing competitor ASCO-lncRNA have been demonstrated to interact with miRNAs as competitors which serve as miRNA target mimics. \textit{IPS1} lncRNA interferes with binding of ath-miR399 as a decoy for PHO2; it binds to mir399 pairing with a three-nucleotide bulge and miR399 targets the PHO2 mRNA for degradation. However, ASCO-lncRNA comprises an alternative splicing regulatory module to bind the regulators of alternative splicing, nuclear speckle RNA binding proteins (NSRs). ASCO-lncRNA hijacks nuclear alternative splicing regulators to alter the alternative splicing patterns to generate alternative splice isoforms during development (Figure 2) [63].

Chromatin-bound lncRNAs are considered as scaffolds that are able to associate with chromatin-modifying enzymes involved in the recruitment of chromatin regulatory proteins to particular genomic DNA locations. LncRNAs scaffold function for the cooperative assembly of chromatin-modifying complexes to recruit them in either an sRNA dependent or an sRNA-independent manner. The findings indicate that lncRNA scaffolds recruit chromatin-modifying complexes independently of sRNAs. However, how protein complexes recognize, join or target lncRNAs still remains unclear. In plants, lncRNAs interact with chromatin-modifying complexes by lncRNA-mediated chromatin modifications (lncR2Epi) regulation pathway [82]. COLD ASSISTED INTRONIC NONCODING RNA (COLDAIR) and COOLAIR are involved in expression of FLOWERING LOCUS C (FLC) in cold-stressed Arabidopsis. These are the best known plant lncRNAs associated with the lncR2Epi regulation pathway [26, 83]. FLC is a MADS-box TF which represses flowering under cold temperature [84]. On the one hand, COLDAIR physically associates with PRC2 to promote H3K27me3 accumulation at FLC [25]. On the other hand, COOLAIR mediates the repression by reducing in H3K36me3 or H3K4me2 at FLC during vernalization [28, 82, 85].

LncRNA-consisting complexes can promote either selective repression or activation of genes based on the nature of the chromatin complex [86–98]. For instance, several lncRNAs recruit histone H3K4 methyltransferases to promote activation of gene expression [90–92], while others bind DNA methyltransferases, including DNMT1 and DNMT3b, to repress transcription by promoting DNA methylation [93–95]. As previously reviewed [98], many lncRNAs interact with the chromatin complex PRC2, illustrating their functional role in epigenetic regulation of gene expression; it catalyzes the generation of the H3K27me3 silencing mark [91, 92, 96–98]. Recently, Neumann et al. [99] demonstrated that the interaction between the lncRNA antisense transcript of GATA6 (GATA6-AS) and the epigenetic regulator LOXL2 regulate the endothelial gene expression by altering histone methylation. The study showed that silencing of GATA6-AS reduced H3K4me3 methylation of periostin and cyclooxygenase-2, indicating that GATA6-AS acts as a negative regulator of nuclear LOXL2 function and epigenetically
regulates endothelial gene expression via interaction with LOXL2 [99]. This mechanism is still not fully understood.

As previously outlined by Chekanova [61], plant IncRNAs can also participate in epigenetic silencing via the RdDM silencing pathway, which generally
functions by silencing of repetitive sequences. This pathway typically requires the plant-specific RNA polymerases Pol IV and Pol V [41] with some involvement of Pol II [100]. Recent findings indicate that RdDM is a plant-specific de novo DNA methylation mechanism that requires IncRNA as scaffold to define target genomic loci [100]. Matzke and Mosher [101] reviewed that some IncRNAs transcribed by Pol IV produce 24-nt siRNAs, while a group of IncRNAs generated by Pol V act as scaffold RNAs recognized via the siRNA-AGO complex through sequence complementarity. The illustration of complexity of siRNA biogenesis in plants centred that Pol IV produces most siRNAs, while Pol V, and to a lesser extent Pol II, generate the templates for siRNAs [102–105]. Based on their very low abundance or stability, discrimination of IncRNAs generated by Pol IV and Pol V is difficult. To date, several Pol V-transcribed scaffold IncRNAs have only been shown to be non-polyadenylated and either tri-phosphorylated or capped at the 5’ ends [41, 61]. In Arabidopsis, Li et al. [106] identified that P4RNAs are Pol IV/RDR2-dependent transcripts which are non-polyadenylated, correspond to both DNA strands and are derived from thousands of loci. Interestingly, the findings demonstrated that the 5’ ends of P4RNAs carry a monophosphate instead of a 5’ triphosphate, or a cap structure [61, 106]. Recently, Au et al. [107] obtained sequences of RdDM-associated IncRNAs by utilizing nuclear RNA immunoprecipitation against ARGONAUTE 4 (AGO4), a key component of RdDM binding specifically with the IncRNA. They compared these IncRNAs with gene expression data of RdDM mutants and identified novel RdDM target genes. Interestingly, a large proportion of these target genes were normally activated by RdDM. These RdDM-activated genes were found to be more enriched for gene body IncRNA than the RdDM-repressed genes, suggesting that RdDM, or AGO4, may play a role in maintaining or activating gene expression by directing gene body chromatin modification preventing cryptic transcription.

In our laboratory, we performed physiological, molecular and fluorescence in situ hybridization (FISH) studies to reveal the effects of IncRNA on barley under salt stress conditions during the germination period. Especially, expression analysis demonstrated that the expression levels of AK363461 and AK370506 barley IncRNAs, additionally, the expression levels of CNT0018772 (maize) and CNT0031477 (rice) IncRNAs, were altered under 150 mmol/L salinity stress [81, 108]. Salinity also affected the expression levels of sense IncRNA AK370814 of AK372815 barley gene, which is related to the vitamin B6 salvage pathway. However, it was observed that the relative expression levels of sense IncRNA AK370814 tended to be up-regulated, although the relative expression levels of the AK372815 gene tended to be down-regulated under salt stress conditions, indicating a relationship between the protein-coding gene and its own sense IncRNA, through alternative splicing [109]. In FISH analysis, AK363461 and AK370506 probes were labelled by tetramethylrhodamine-dUTP (TRITC). The results showed that the AK363461 and AK370506 probes revealed four signals and one signal under a confocal microscope, respectively. To our knowledge, this was the first report to demonstrate barley IncRNA localisations at the prophase stage in the nucleus and also to present a protocol for visualization of barley IncRNAs [108].

Bioinformatics approaches

The improvement of next generation sequencing (NGS) technologies, and the large number of sequencing projects have been generating enormous volumes of biological data [110]. Experimental (NGS technologies) and computational screenings generate fragments of transcripts; then the transcript sequences are mapped to a reference genome and the identified transcribed units are defined as the ‘novel’ RNAs [111–113]. To distinguish between coding and non-coding sequences, the most important and also popular criterion is similarity to known coding sequences or statistics of codon frequencies for coding potential [114].

Some computational techniques involving machine learning techniques, e.g. CONC (coding or non-coding) [115], CPC (coding potential calculator) [116] and PORTRAIT [117], have been designed to distinguish ncRNAs from protein-coding transcripts (PCTs). CONC performs slowly on analyzing large ncRNA datasets, while CPC is suitable for known protein-coding transcripts. However, CPC may tend to classify novel PCTs into ncRNAs not recorded in the protein databases [111, 118]. The most common methods use pairwise comparisons (CPC [116] and PORTRAIT [117]) or multiple alignments (PhyloCSF [119] and RNAcode [120]). Coding Potential Assessment, which is also another popular approach, uses an alignment-free logistic regression model to determine the coding potential of RNAs [121]. To classify IncRNAs by machine learning techniques, there are CNCI [122], PLEC [123], LncRScan-SVM [124], IncRNA-MFDL [125], IncRNA-ID [126] and IncRNApred [127]. In human and mice, iSeeRNA [128], linc-SF [129] and DeepLNC [130] use
machine learning techniques to categorize IncRNAs from transcriptome sequencing data [85]. iSeeRNA is a computational pipeline to identify IncRNAs on the basis of support vector machine (SVM) algorithm. However, DeepLNC utilizes a different approach to identify novel IncRNAs using a deep neural network (DNN)-based classification model [85, 115, 122]. In plants, some studies have used both experimental and computational techniques to find and characterize IncRNAs [16, 85, 131].

Conclusions

Compelling evidence supports the involvement of IncRNAs’ control in the epigenetic state of particular genes, participation in transcriptional regulation; involvement in alternative splicing and contribution of sub-nuclear compartments. However, our knowledge of the regulatory and functional roles of IncRNAs is still in its infancy, research being hindered by their diversity, low expression level and sequence divergence. To better understand their genetic and epigenetic functions, researchers are employing the latest advancements in computational technologies. Our recent studies demonstrated that IncRNAs were also associated with the responses to salt stress in barley, as well as drought in Arabidopsis, heat and biotic stress in wheat. It is hoped that accumulating data about the expression patterns, tissue distribution, subcellular localization, abundance and splicing patterns of IncRNAs under particular conditions will help to unravel the functional roles of IncRNAs on the genetic and epigenetic level.

Funding

This work was supported by Scientific Research Projects Coordination Unit of Istanbul University under grant number FDK-2016-23086.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Jacquier A. The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. Nat Rev Genet. 2009;10:833–844.

[2] Chu C, Qu K, Zhong FL, et al. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell. 2011;44: 667–678.

[3] Chu C, Quinn J, Chang HY. Chromatin isolation by RNA purification (ChIRP). J Vis Exp. 2012;61:e3912.

[4] Gong C, Popp MW, Maquat LE. Biochemical analysis of long non-coding RNA-containing ribonucleoprotein complexes. Methods. 2012;58:88–93.

[5] Darnell R. CLIP (cross-linking and immunoprecipitation) identification of RNAs bound by a specific protein. Cold Spring Harb Protoc. 2012;2012:1146–1160.

[6] Sánchez-León N, Arteaga-Vázquez M, Alvarez-Mejía C, et al. Transcriptional analysis of the Arabidopsis ovule by massively parallel signature sequencing. J Exp Bot. 2012;63:3829–3842.

[7] Feng Y, Hu X, Zhang Y, et al. Methods for the study of long noncoding RNA in cancer cell signaling. Berlin: Springer; 2014.

[8] Engreitz J, Lander ES, Guttman M. RNA antisense purification (RAP) for mapping RNA interactions with chromatin. Berlin: Springer; 2015.

[9] Uszczynska-Ratajczak B, Lagarde J, Frankish A, et al. Towards a complete map of the human long non-coding RNA transcriptome. Nat Rev Genet. 2018;19: 535–548.

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

[11] Kung JT, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. Genetics. 2013;193: 651–669.

[12] Fok ET, Scholefield J, Fanucci S, et al. The emerging molecular biology toolbox for the study of long non-coding RNA biology. Epigenomics. 2017;9: 1317–1327.

[13] Ravasi T, Suzuki H, Pang KC, et al. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. Genome Res. 2005;16:11–19.

[14] Guttman M, Amit I, Garber M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009;458: 223–227.

[15] Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009;136: 629–641.

[16] Liu J, Jung C, Xu J, et al. Genomewide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. Plant Cell. 2012;24:4333–4345.

[17] Matsui A, Nguyen AH, Nakaminami K, et al. Arabidopsis non-coding RNA regulation in abiotic stress responses. Int J Mol Sci. 2013;14:22642–22654.

[18] Santosh B, Varshney A, Yadava PK. Non-coding RNAs: biological functions and applications. Cell Biochem Funct. 2015;33:14–22.

[19] Ye CY, Chen L, Liu C, et al. Widespread noncoding circular RNAs in plants. New Phytol. 2015;208:88–95. [cited 2018 Aug 22]

[20] Yuan J, Li J, Yang Y, et al. Stress-responsive regulation of long non-coding RNA polyadenylation in Oryza sativa. Plant J. 2018;93:814–827.

[21] Liu D, Mewalal R, Hu R, et al. New technologies accelerate the exploration of non-coding RNAs in horticultural plants. Hortic Res. 2017;4:17031. [cited 2018 Aug 22]
[22] Ben Amor B, Wirth S, Merchan F, et al. Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. Genome Res. 2008;19:57–69.

[23] Aung K, Lin SL, Wu CC, et al. pho2, a phosphate overaccumulator, is caused by a nonsense mutation in a microRNA399 target gene. Plant Physiol. 2006;141:1000–1011.

[24] Pant BD, Buhtz A, Kehr J, et al. MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. Plant J. 2008;53:731–738.

[25] Sheldon CC, Rouse DT, Finnegan EJ, et al. The molecular basis of vernalization: the central role of FLOWERING LOCUS C (FLC). Proc Natl Acad Sci USA. 2000;97:3753–3758.

[26] Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science. 2011;331:76–79.

[27] Zhu QH, Wang MB. Molecular functions of long non-coding RNAs. Mol Cell. 2011;43:904–915.

[28] Marquardt S, Raitskin O, Wu Z, et al. Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. Mol Cell. 2014;54:156–165.

[29] Campalans A, Kondorosi A, Crespi M. Enod40, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in Medicago truncatula. Plant Cell. 2004;16:1047–1059.

[30] Dey M, Complainville A, Charon C, et al. Phytohormonal responses in enod40-overexpressing plants of Medicago truncatula and rice. Physiol Plant. 2004;120:132–139.

[31] Liu J, Wang H, Chua NH. Long noncoding RNA transcriptome of plants. Plant Biotechnol J. 2015;13:319–328.

[32] Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. RNA Biol. 2013;10:925–933.

[33] St Laurent G, Shitokalo D, Tackett MR, et al. Intronic RNAs constitute the major fraction of the non-coding RNA transcriptome of plants. Plant Physiol. 2012;156:13504. [cited 2018 Aug 22]

[34] Kapranov P, Cheng J, Dike S, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science. 2007;316:1484–1488.

[35] Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012;22:1760–1774.

[36] Atkinson SR, Marguerat S, Bahler J. Exploring long non-coding RNAs through sequencing. Semin Cell Dev Biol. 2012;23:200–205.

[37] Devaux Y, Zangrando J, Schroen B, et al. Long non-coding RNAs in cardiac development and ageing. Nat Rev Cardiol. 2015;12:415–425.

[38] Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. Cell. 2013;154:26–46.

[39] Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell. 2011;43:904–914.

[40] Dieci G, Fiorino G, Castelnuovo M, et al. The expanding RNA polymerase III transcriptome. Trends Genet. 2007;23:614–622.

[41] Wierzbicki AT, Haag JR, Pikaard CS. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell. 2008;135:635–648.

[42] Wierzbicki AT, Ream TS, Haag JR, et al. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. Nat Genet. 2009;41:630–634.

[43] Holoch D, Moazed D. RNA-mediated epigenetic regulation of gene expression. Nat Rev Genet. 2015;16:71–84.

[44] Böhmderger G, Sethuraman S, Rowley MJ, et al. Long non-coding RNA produced by RNA polymerase V determines boundaries of heterochromatin. eLife. 2016;5:e19092.

[45] Penny GD, Kay GF, Sheardown SA, et al. Requirement for Xist in X chromosome inactivation. Nature. 1996;379:131–137.

[46] Marahrens Y, Loring J, Jaenisch R. Role of the Xist gene in X chromosome choosing. Cell. 1998;92:657–664.

[47] Lee JT, Lu N. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. Cell. 1999;99:47–57.

[48] Franke A, Baker B. The rox1 and rox2 RNAs are essential components of the compensosome, which mediates dosage compensation in Drosophila. Mol Cell. 1999;4:117–122.

[49] Schuster-Gossler K, Simon-Chazottes D, Guenet JL, et al. Gtl2lacZ, an insertional mutation on mouse chromosome 12 with parental origin dependent phenotype. Mamm Genome. 1996;7:20–24.

[50] Sleutels F, Zwart R, Barlow DP. The non-coding air RNA is required for silencing autosomal imprinted genes. Nature. 2002;415:810–813.

[51] Mancini-Dinardo D, Steele SJ, Levorse JM, et al. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. Genes Dev. 2006;20:1268–1282.

[52] Nakagawa S. Lessons from reverse-genetic studies of IncRNAs. Biochim Biophys Acta. 2016;1859:177–183.

[53] Yang WC, Katinakis P, Hendriks P, et al. Characterization of GmENOD40, a gene showing novel patterns of cell-specific expression during soybean nodule development. Plant J. 1993;3:573–585.

[54] Crespi MD, Jurkевич E, Poirier M, et al. enod40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. Embo J. 1994;13:5099–5112.

[55] Liu C, Muchhal US, Raghothama KG. Differential expression of TPS11, a phosphate starvation-induced gene in tomato. Plant Mol Biol. 1997;33:867–874.

[56] Wasaki J, Yonetani R, Shinano T, et al. Expression of COOLAIR on FLC transcription. Mol Cell. 2014;54:1000–1011.

[57] Cech TR, Steitz JA. The noncoding RNA revolution—trashing old rules to forge new ones. Cell. 2014;157:933–944.

[58] Law MJ, Rice AJ, Lin P, et al. The role of RNA structure in the interaction of U1A protein with U1 hairpin II RNA. RNA. 2006;12:1168–1178.
Kim TK, Hemberg M, Gray JM. Enhancer RNAs: a class of long noncoding RNAs synthesized at enhancers. Cold Spring Harb Perspect Biol. 2015;7: a018622.

[59] Hudson WH, Ortlund EA. The structure, function and evolution of proteins that bind DNA and RNA. Nat Rev Mol Cell Biol. 2014;15:749–760.

[60] Long Y, Wang X, Youmans DT, et al. How do IncRNAs regulate transcription? Sci Adv. 2017;3: eaa02110. [cited 2018 Nov 10]

[61] Chekanova JA. Long non-coding RNAs and their functions in plants. Curr Opin Plant Biol. 2015;27: 207–216.

[62] Franco-Zorrilla JM, Valli A, Todesco M, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet. 2007;39:1033–1037.

[63] Bardou F, Ariel F, Simpson CG, et al. Long noncoding RNA modulates alternative splicing regulators in Arabidopsis. Dev Cell. 2014;30:166–176.

[64] Cho J, Paszkowski J. Regulation of rice root development by a retrotransposon acting as a microRNA sponge. Elife. 2017;6:e30038. http://10.7554/elife.30038.e001

[65] Yan J, Gu Y, Jia X, et al. Effective small RNA destruction by the expression of a short tandem target mimic in Arabidopsis. Plant Cell. 2012;24:415–427.

[66] Reichel M, Li Y, Li J, et al. Inhibiting plant microRNA activity: molecular SPONGEs, target MIMICs and STTMs all display variable efficacies against target microRNAs. Plant Biotechnol J. 2015;13:915–926.

[67] Cho J. Transposon-derived non-coding RNAs and their function in plants. Front Plant Sci. 2018;9:600. [cited 2018 Nov 10]

[68] Wang P, Xue Y, Han Y, et al. The STAT3-binding long noncoding RNA Lnc-DC controls human dendritic cell differentiation. Science. 2014;344:310–313.

[69] Bonasio R, Shiekhattar R. Regulation of transcription by long noncoding RNAs. Annu Rev Genet. 2014;48: 433–455.

[70] Wang Y, Fan X, Lin F, et al. Arabidopsis noncoding RNA mediates control of photomorphogenesis by red light. Proc Natl Acad Sci USA. 2014;111:10359–10364.

[71] Henriques R, Wang H, Liu J, et al. The antiphase regulatory module comprising CDF5 and its antisense RNA FLORE links the circadian clock to photo-periodic flowering. New Phytol. 2017;216:854–867.

[72] Lai F, Orom UA, Cesaroni M, et al. Activating RNAs associate with mediator to enhance chromatin architecture and transcription. Nature. 2013;494:497–501.

[73] Pefanis E, Wang J, Rothschild G, et al. RNA exosome-regulated long non-coding RNA transcription controls superenhancer activity. Cell. 2015;161:774–789.

[74] Lam MT, Cho H, Lesch HP, et al. Rev-Erbα represses macrophage gene expression by inhibiting enhancer-directed transcription. Nature. 2013;498:511–515.

[75] Li W, Notani D, Ma Q, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 2013;498:516–520.

[76] Melo CA, Drost J, Wijchers PJ, et al. eRNAs are required for p53-dependent enhancer activity and gene transcription. Mol Cell. 2013;49:524–535.

[77] Kim TK, Hemberg M, Gray JM. Enhancer RNAs: a class of long noncoding RNAs synthesized at enhancers. Cold Spring Harb Perspect Biol. 2015;7: a018622.

[78] Xin M, Wang Y, Yao Y, et al. Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. BMC Plant Biol. 2011;11:61.

[79] Kim ED, Sung S. Long noncoding RNA: unveiling hidden layer of gene regulatory networks. Trends Plant Sci. 2012;17:16–21.

[80] Zhang YC, Liao JY, Li ZY, et al. Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. Genome Biol. 2014;15:512.

[81] Karlk E, Gözküürmüzı N. Evaluation of barley LncRNA expression analysis in salinity stress. Russ J Genet. 2018;54:198–204.

[82] Wang J, Meng X, Yuan C, et al. The roles of cross-talk epigenetic patterns in Arabidopsis thaliana. Brief Funct Genomics. 2016;15:278–287.

[83] Csorba T, Questa Ji, Sun Q, et al. Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalisation. Proc Natl Acad Sci USA. 2014;111:16160–16165.

[84] Michaels SD, Amasio RM. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell. 1999;11:949–956.

[85] Wang J, Meng X, Dobrovolskaya OB, et al. Non-coding RNAs and their roles in stress response in plants. Genomics Proteomics Bioinf. 2016;15:21–26.

[86] Marchese FP, Huarte M. Long non-coding RNAs and chromatin modifiers: their place in the epigenetic code. Epigenetics. 2014;9:21–26.

[87] Morlando M, Ballarino M, Fatica A, et al. The role of long noncoding RNAs in the epigenetic control of gene expression. ChemMedChem. 2014;9:505–510.

[88] Gendrel AV, Heard E. Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation. Annu Rev Cell Dev Biol. 2014;30:561–580.

[89] Meller VH, Joshi SS, Deshpande N. Modulation of chromatin by noncoding RNA. Annu Rev Genet. 2015;49:673–695.

[90] Wang KC, Yang YW, Liu B, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature. 2014;2013;24:206–214.

[91] Cabianca DS, Casa V, Bodega B, et al. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. Cell. 2012;149:819–831.

[92] Grote P, Wittler L, Hendrix D, et al. The tissue-specific IncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. Dev Cell. 2013;24:206–214.

[93] Mohammad F, Mondal T, Guseva N, et al. Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. Development. 2010;137:2493–2499.

[94] Schmitz KM, Mayer C, Postepksa A, et al. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. Genes Dev. 2010;24:2264–2269.

[95] Johansson P, Ackley A, Vidasdottir L, et al. A pseudo-gene long-noncoding-RNA network regulates PTEN
transcription and translation in human cells. Nat Struct Mol Biol. 2013;20:440–446. [cited 2018 Aug 22]

[96] Davidovich C, Cech TR. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. RNA. 2015;21:2007–2022.

[97] Kotzin JJ, Spencer SP, McCright SJ, et al. The long non-coding RNA Morrbid regulates Bim and short-lived myeloid cell lifespan. Nature. 2016;537:239–243.

[98] Marchese FP, Raimondi I, Huarte M. The multidimensional mechanisms of long noncoding RNA function. Genome Biol. 2017;18:206.

[99] Neumann P, Jaé N, Knau A, et al. The IncRNA GATA6-AS epigenetically regulates endothelial gene expression via interaction with LOXL2. Nat Commun. 2018;9:237.

[100] Zheng B, Wang Z, Li S, et al. Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. Genes Dev. 2009;23:2850–2860.

[101] Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet. 2014;15:394.

[102] Kong L, Zhang Y, Ye ZQ, et al. PlantRNA Sniffer: a SVM-based workflow to predict long intergenic non-coding RNAs in plants. Non-Coding RNA. 2017;3:11. [cited 2018 Aug 22]

[103] You W, Lorkovic ZJ, Matzke AJM, et al. Interplay among RNA polymerases II, IV and V in RNA-directed DNA methylation at a low copy transgene locus in Arabidopsis thaliana. Plant Mol Biol. 2013;82:85–96.

[104] Sasaki T, Lee TF, Liao WW, et al. Distinct and concurrent pathways of Pol II- and Pol IV-dependent siRNA biogenesis at a repetitive trans-silencer locus in Arabidopsis thaliana. Plant J. 2014;79:127–138.

[105] Li S, Vandivier LE, Tu B, et al. Detection of Pol IV/ RDR2-dependent transcripts at the genomic scale in Arabidopsis reveals features and regulation of siRNA biogenesis. Genome Res. 2015;25:235–245.

[106] Au PCK, Dennis ES, Wang MB. Analysis of Argonaute 4-associated long non-coding RNA in Arabidopsis thaliana sheds novel insights into gene regulation through RNA-directed DNA methylation. Genes. 2017;8:198.

[107] Karlik E, Maraklı S, Gözükürmüz N. Two IncRNAs expression profiles in salt stressed barley (Hordeum vulgare L.) roots. Cytologia. 2018;83:37–43.

[108] Karlik E, Gözükürmüz N. Expression analysis of IncRNA AK370814 involved in the barley vitamin B6 salvage pathway under salinity. Mol Biol Rep. 2018;45:1597–1609.

[109] Neron A, Ear U. Kyrpides N. Genomes onLine Database (GOLD): a mirror of genome projects worldwide. Nucleic Acids Res. 2001;29:126–127.

[110] Trapnell C, Williams BA, Pertea G, et al. Transcript assembly and quantification by RNASeq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511–515.

[111] Kim D, Pertea G, Trapnell C, et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14:R36. [cited 2018 Aug 22]

[112] Dobin A, Davis CA, Schlesinger F, et al. STAR: ultra-fast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

[113] Wang L, Park HJ, Dasari S, et al. CPAT: coding-potential assessment tool using an alignment-free logistic regression model. Nucleic Acids Res. 2013;41:e74. [cited 2018 Aug 22]

[114] Liu J, Gough J, Rost B. Distinguishing protein-coding from non-coding RNAs through support vector machine. PLoS Genet. 2006;2:e29. [cited 2018 Aug 22]

[115] Kong L, Zhang Y, Ye ZQ, et al. PlantRNA Sniffer: a SVM-based workflow to predict long intergenic non-coding RNAs in plants. Non-Coding RNA. 2017;3:11. [cited 2018 Aug 22]

[116] Lin MF, Jungreis I, Kellis M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics. 2011;27:i275–i282.

[117] Wang L, Park HJ, Dasari S, et al. CPAT: coding-potential assessment tool using an alignment-free logistic regression model. Nucleic Acids Res. 2013;41:e74. [cited 2018 Aug 22]

[118] Washietl S, Findeiss S, Müller SA, et al. RNAcode: robust discrimination of coding and noncoding regions in comparative sequence data. RNA. 2011;17:578–594.

[119] Li A, Zhang J, Zhou Z. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. BMC Bioinform. 2014;15:311. [cited 2018 Aug 22]

[120] Sun L, Luo H, Bu D, et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. Nucleic Acids Res. 2013;41: e166. [cited 2018 Aug 22]

[121] Li A, Zhang J, Zhou Z. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. BMC Bioinform. 2014;15:311. [cited 2018 Aug 22]

[122] Sun L, Liu H, Zhang L, et al. LncRNA-SVM: a tool for predicting long non-coding RNAs using support vector machine. PLoS One. 2015;10:e0139654. [cited 2018 Aug 22]

[123] Fan XN, Grativol C, Thiebaut F, et al. PlantRNA Sniffer: a SVM-based workflow to predict long intergenic non-coding RNAs in plants. Non-Coding RNA. 2017;3:11. [cited 2018 Aug 22]

[124] Washietl S, Findeiss S, Müller SA, et al. RNAcode: robust discrimination of coding and noncoding regions in comparative sequence data. RNA. 2011;17:578–594.

[125] Wang L, Park HJ, Dasari S, et al. CPAT: coding-potential assessment tool using an alignment-free logistic regression model. Nucleic Acids Res. 2013;41:e74. [cited 2018 Aug 22]

[126] Fan XN, Zhang SW. lncRNA-MFDL: identification of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. BMC Bioinform. 2014;15:311. [cited 2018 Aug 22]

[127] Washietl S, Findeiss S, Müller SA, et al. RNAcode: robust discrimination of coding and noncoding regions in comparative sequence data. RNA. 2011;17:578–594.
hybrid feature. PLoS One. 2016;11:e0154567 [cited 2018 Aug 22]DOI: 0.1371/journal.pone.0154567

[128] Sun K, Chen X, Jiang P, et al. iSeeRNA: identification of long intergenic non-coding RNA transcripts from transcriptome sequencing data. BMC Genomics. 2013;14:7.

[129] Wang Y, Li Y, Wang Q, et al. Computational identification of human long intergenic non-coding RNAs using a GA-SVM algorithm. Gene. 2014;533:94–99.

[130] Tripathi R, Patel S, Kumari V, et al. DeepLNC, a long non-coding RNA prediction tool using deep neural network. Netw Model Anal Health Inform Bioinform. 2016;5:21.

[131] Wang J, Yu W, Yang Y, et al. Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection. Sci Rep. 2015;5:16946 [cited 2018 Aug 22]16946.