Genomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Classification and experimental identification of plant long non-coding RNAs

Muneeza I. Rai[a,1], Maheen Alam[a,1], David A. Lightfoot[b], Priyatansh Gurha[c], Ahmed J. Afzal[a,b,*]

[a] Department of Biology, SBA School of Science and Engineering, Lahore University of Management Sciences, Lahore, Pakistan
[b] Department of Plant, Soil and Agricultural Systems, Southern Illinois University Carbondale, USA
[c] Center for Cardiovascular Genetics, Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, United States

ARTICLE INFO

Keywords:
LncRNA
Non-coding RNA
Gene regulation
RNA-seq

ABSTRACT

Our understanding of gene regulation is constantly evolving. It is now clear that the majority of cellular transcripts are non-coding RNAs. The spectrum of non-coding RNAs is diverse and includes short (< 200 nt) and long non-coding RNAs (lncRNAs) (> 200 nt). LncRNAs regulate gene expression through diverse mechanisms. In this review, we describe the emerging roles of lncRNA mediated plant gene regulation. We discuss the current classification of lncRNAs and their role in genome organization and gene regulation. We also highlight a subset of lncRNAs that are epigenetic regulators of plant gene expression. Lastly, we provide an overview of emerging techniques and databases that are employed for the identification and characterization of plant lncRNAs.

1. Introduction

The C-value paradox posits that there is an enormous variation in genome size among species and this variation does not significantly correlate with developmental complexity [43]. The vast majority of the non-coding genome was declared as "Junk DNA" due to the presence of simple repetitive sequences, transposons and pseudogenes [16,43]. However, numerous studies using high-throughput sequencing technologies have now determined that almost 90% of the eukaryotic genome gets transcribed [3,43] which is in sharp contrast to the approximately 2% that gets translated. The remaining transcriptome comprises of non-coding RNAs (ncRNAs), which have little to no coding ability [109]. In contrast to the C-value paradox, the non-coding RNA content within a genome is proposed to correlate with the developmental and physiological complexity of an organism [63,70]. The ncRNAs consist of a diverse range of transcripts, which vary in size, ranging from 20 to 30 nucleotides (nts) for small ncRNAs to > 200 nts for long ncRNAs or lncRNAs [28]. The location and function of several ncRNAs, such as ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snoRNAs and snRNAs), transfer RNAs (tRNAs) as well as regulatory RNAs such as microRNAs (miRNA) and small interfering RNA (siRNA) have been well characterized in eukaryotes [8,56,68,122]. With the development of advanced sequencing techniques and computational methods, thousands of lncRNAs have recently been identified in a wide array of plants [33,48,55,58,112]. However, limited information is available on the molecular mechanisms governing the function of lncRNAs [56]. Recent studies have highlighted the roles played by plant lncRNAs in key cellular processes including transcription, epigenetic modification and post-transcriptional modifications [28]. This short report provides an overview on the current knowledge of plant lncRNAs, their mode of identification, classification and functions.

2. LncRNA classification

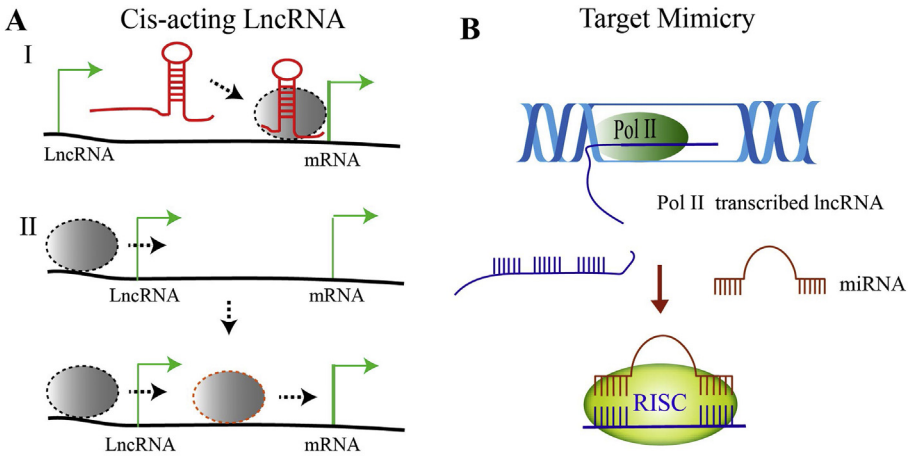
Long noncoding RNAs are classified on the basis of their genomic location and function (Fig. 1). In this review, we will discuss these two major classification schemes and further highlight the importance of a subclass of plant specific lncRNAs that play a role in RNA dependent DNA methylation (RdDM).

2.1. Classification based on function

Functional analysis of eukaryotic lncRNAs has revealed that they act as scaffolds, guides, and targets to regulate gene expression through epigenetic modification and other post-transcriptional gene regulatory mechanisms (Table 1) [6,28,56,68,99]. Based on the functional roles the lncRNAs can be classified as:

* Corresponding author at: Department of Biology, SBA School of Science and Engineering, Lahore University of Management Sciences, Lahore, Pakistan.
E-mail address: ahmed.afzal@lums.edu.pk (A.J. Afzal).
[1] Equal contributors.

https://doi.org/10.1016/j.ygeno.2018.04.014
Received 8 February 2018; Received in revised form 13 April 2018; Accepted 17 April 2018



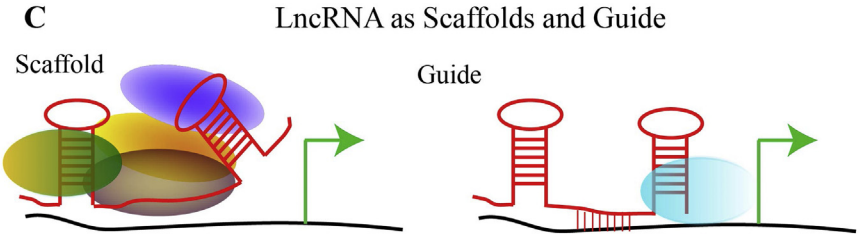


Fig. 1. Schematic model showing the classification of lncRNAs based on regulatory functions. A) Cis-acting lncRNAs modulate expression of the target gene at the same locus by (I) the recruitment of proteins or (II) through the act of getting transcribed, which leads to the opening of local chromatin for downstream mRNA transcription. B) As miRNA target mimics, lncRNAs (purple) sequester the miRNA (red) to inhibit the silencing of target mRNA. C) Scaffold lncRNAs: either acts as decoy (left panel) to sequester a protein or act as guide RNA to recruit protein/s to mediate gene expression (Right panel). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
List of plant long noncoding RNAs identified on the basis of their biological function and mode of action.

SPECIES	Name	Mechanism of action	Biological function	References
Arabidopsis thaliana (Arabidopsis)	IPS1	Target mimicry: interacts with an miRNA, ath-mir399	Regulates phosphate homeostasis	[22]
	COLDAIR	Epigenetic silencing: modification of histones	Regulates flowering	[27]
	COOLAIR	Epigenetic silencing: Promoter interference	Regulates flowering	[88]
	HID1	Stage specific: Associates with chromatin and represses transcription of phytohormone interaction factor 3 (PIF3)	Regulates photomorphogenesis	[103]
	asHSFB2a	Hijacks nuclear Antisense transcription regulators	Regulates the gametophytic and vegetative development	[111]
	ASCO-lncRNA	Regulators of alternate splicing	Regulates the development of lateral roots	[5]
	APOLO	Causes the formation of Chromatin loop and regulates the expression of its neighboring genes	Regulated plant development through auxin signaling	[2]
	ELENA1	Increases the expression of genes involved in plant immunity, such as PR1 and PR2, B 1,3glucanase and salicylic acid genes	Upregulated in disease caused by *Pseudomonas syringae*	[61]
	Drought induced RNA (DRIR)	Regulates stress responses such as stomata closure	Upregulated in drought and salt stress	[72]
	CDF5 LONG NONCODING RNA (FLORE)	Natural anti-sense transcript of CDF5, upregulates flowering	Photomorphogenesis	[26]
Glycine max (soybean)	GmENOD40	Involved in the development of nodules	Regulates the re-localization of proteins from nucleus to cytoplasm	[117]
Hordeum vulgare (Barley)	HvCesA6 lnc-NAT	Acts as a precursor for siRNA	Involved in the synthesis of cell wall	[25]
Medicago truncatula (Barrel medic)	MtENOD40	Involved in the development of nodules	Regulates the re-localization of proteins	[83]
Oryza sativa (Rice)	OsPI1	Unknown	Regulates phosphate homeostasis	[106]
	OsENOD40	Involved in the development of nodules	Unknown	[42]
	LDMAR (P/TMS12-1)	Acts by causing promoter methylation	Male fertility regulation	[17]
	Cis-NATPHO1;2	Acts by enhancing translation of *PHOSPHATE1;2*	Regulates phosphate homeostasis	[30]
	lncRNAs: 2224 lincRNAs: 1624 lncNATs: 600	Tissue and stage specific: Anthers, pistils, seeds five days after pollination, and shoots 14 days after germination	Regulates sexual reproduction	[128]
Petunia hybrid (Petunia)	SHO lnc-NAT	Degrades dsRNA	Regulates the synthesis of local cytokines	[138]
Populus trichocarpa	incRNA20 lincRNA2752	target mimic of ptc-miR476 target mimic of ptc-miR169	Upregulated in abiotic stress (drought)	[80]
Solanum lycopersicum (Tomato)	TPS11	Unknown	Regulates phosphate homeostasis	[53]
	Slylnc0195 slylnc1077	target mimic of miR166 target mimic of miR399	Upregulated in infections caused by tomato yellow leaf curl virus	[97,101]

2.1.1. Cis acting lncRNAs

Long non coding RNAs regulating expression of the genes present on the same chromosomal location are cis-acting, that is, both the regulatory RNA and the target gene are transcribed from the same locus. These cis acting lncRNAs include various promoter and enhancer associated RNAs as well as intronic and antisense lncRNA. In certain cis

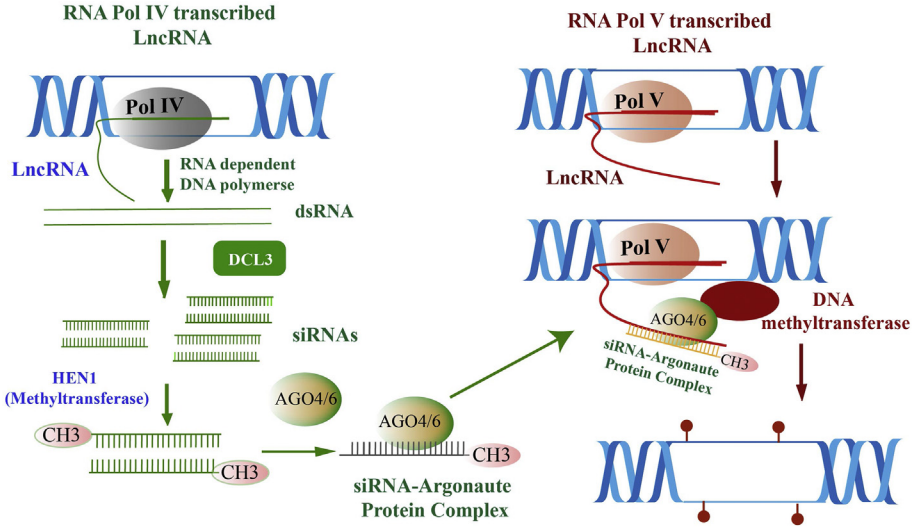


Fig. 2. A simplified model depicting the role of lncRNAs in Pol IV/V mediated RNA dependent DNA methylation (RdDM). Pol IV (grey oval) transcription leads to the generation of ssRNAs that gets copied to dsRNAs, which are further processed into 24 nt siRNAs (green). These siRNAs are stabilized by methylation of their 3′ ends and subsequently loaded onto the Argonaute complex for DNA methylation. The targeting of siRNA-Argonaute complex to DNA (to be methylated) is mediated by Pol V (brown oval) transcribed lncRNAs through direct base pairing between siRNA and lncRNA. Recruitment of DNA methyltransferase leads to the methylation at the cytosine residues (red circles) of the DNA. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

acting lncRNAs, the lncRNA transcription process can affect the expression of nearby genes mainly through RNAP II recruitment and through changes in chromatin organization of the locus.

The prime example of cis-regulation in placental mammals is Xist lncRNA that is involved in the induction of X chromosome inactivation. This lncRNA is expressed from one X chromosome and brings about the silencing of the entire chromosome.

Another example of a cis-acting lncRNA involved in regulation of plant development is the antisense Delay of Germination 1 (*as-DOG1)* lncRNA [20]. Seed dormancy in *Arabidopsis* is regulated by the *DOG1* gene and the release of dormancy marks the first developmental stage in the plants life cycle. This is regulated by the as-DOG1 lncRNA which begins from the proximal transcription termination site (TTS) of the DOG1 gene and ends at the DOG1 TSS. This lncRNA has both a 5′ cap and a polyadenylation tail and acts as a suppressor of seed dormancy.

2.1.2. LncRNAs as target mimics

Unlike cis-acting lncRNAs, trans-acting lncRNAs bind miRNAs and block the interaction between miRNAs and their specific target mRNAs. This process is termed endogenous target mimicry (eTM) (Fig. 1). In eTM, the target mimics act as decoys by binding to specific miRNAs in order to block their interaction with actual targets. Target mimics have an insertion at the miRNA cleavage site, which creates a mismatch of at least three nucleotides. This mismatch is at the position of miRNA cleavage site to avoid any detection by the miRNA dicing proteins thereby resulting in a more stable sequestration of miRNA [22,110]. Ever since the first reported eTM, several putative and functionally characterized eTMs have been identified within the intergenic or non-coding genes in a number of plant species [22,56,110]. Several eTMs have been reported to regulate plant immunity by suppressing the function of the corresponding miRNA thereby up-regulating specific defense related genes [6]. For instance, in tomato, the lncRNAs slylnc0195 and 1077 are up-regulated in response to infection caused by Tomato Yellow Leaf Curl Virus (TYLCV). Slylnc0195 and 1077 act as target mimics of miR-166 and miR-399 respectively [68]. Slylnc0195 mediated mimicry of miR-166 was associated with an increase in expression of a number of miR-166 predicted targets.

Apart from their biological significance, eTMs are also being used to characterize the function of different plant miRNAs in transgenic plants [91,110]. As miRNA target mimics, lncRNAs can modulate miRNA mediated gene silencing. Target mimics can be designed for the miRNA gene family of interest for simultaneous silencing of all miRNA family members [22]. Recent miRNA functional studies have shown that the target mimics can effectively abolish miRNA activity in both transient assays as well as in stable transgenic lines [91,115,124].

2.1.3. LncRNAs as scaffold and guides

Many lncRNAs carry out their cellular functions by interacting with proteins to form RNA-protein complexes. In some cases the lncRNA can pair with the DNA element and thereby act as a guide for these RNA-protein complexes while in others the lncRNA facilitates recruitment or formation of a macromolecular complex thereby acting as a scaffold. These interactions are mediated by specific RNA sequence motifs or domains that are generally composed of secondary and tertiary structures formed by RNA folding within a lncRNA [49,86]. In Arabidopsis the trans-acting lncRNA COLDAIR acts as a guide RNA to recruit chromatin complexes to mediate gene expression. Likewise, the APOLO lincRNA acts as scaffolding RNA that participates in chromatin loop formation [2,27].

2.1.4. LncRNAs as mediators of RNA dependent DNA methylation

RNA dependent DNA methylation is an epigenetic silencing mechanism for genes, transposons and repetitive elements [108] that is also proposed to regulate gene expression by repressing chromosomal interactions and chromatin contact [77]. Plants have evolved two additional non-redundant polymerases (Pol IV and V) that specifically transcribe lncRNAs that regulate gene expression by RNA mediated DNA methylation (RdDM) [38].

The plant specific multi-subunit RNA Polymerase IV and V and their lncRNA products are the major players in the RdDM pathway (Fig. 2). In the first step of the pathway, RNA Pol IV transcribed lncRNA is diced into 24 nucleotide small interfering RNAs (siRNAs), which get methylated and bind to Argonaute proteins (AGO4/6 in Arabidopsis). This results in the formation of an Argonaute-siRNA complex. Subsequently, RNA Pol V transcribed lncRNAs act as scaffolds and recruit the Argonaute-siRNA complex to the chromatin target site. After the subsequent interaction with Argonaute proteins, Pol V transcribed lncRNAs recruit methyl transferases to cytosine residues at regions surrounding the lncRNA site, ensuing gene silencing (Fig. 2) [67].

Plants employ POL IV and V for both steps in the RdDM pathway that is siRNA biogenesis and scaffold lncRNAs transcription respectively (Fig. 2). An important point to note here is that some lncRNAs are transcribed by both Pol II and Pol IV/V [2]. Pol II has also been also proposed to have a role in RdDM silencing as a weak mutant of Pol II subunit shows decreased methylation rate [84,130].

RdDM mediated gene regulation is not the only distinguishing characteristic of Pol IV/V transcribed lncRNAs. There are additional distinct features of Pol IV/V transcribed lncRNAs. For instance Pol IV/V transcribed lncRNAs are not processed like mRNAs for 3′ polyadenylation as they lack the specific subunits for interaction with the processing enzymes [107]. Pol IV/V recruitment requires specific

chromatin modifications such as histone H3K9 dimethylation (H3K9me2), HDA6 dependent histone deacetylation and maintenance methylation by MET1 and CMT3 [7,107]. Identification of Pol IV/V transcribed lncRNAs specifically requires more sensitive techniques as they are expressed at extremely low levels owing to silencing at the chromatin level. Recently, over 20,000 Pol IV dependent lncRNAs were identified using a Pol IV mutant [51].

2.2. Classification based on genomic location

2.2.1. Sense and intronic lncRNAs

LncRNAs transcribed from a region overlapping a protein-coding gene that shares the same promoter, are classified as sense lncRNAs and those arising from an intron of a protein-coding gene are called intronic lncRNAs [60]. A well-studied example of sense and intronic lncRNAs in plants is COLDAIR (COLD Assisted Intronic noncoding RNA). COLDAIR regulates silencing of the floral repressor FLC (Flowering Locus C) through the enrichment of PRC2 (Polycomb Repressive Complex 2) at the locus. Physical association of COLDAIR with PRC2 results in the recruitment of PRC2 to the FLC locus. Silencing of the FLC locus upon prolonged cold exposure or vernalization treatment ensues flowering in spring [27].

2.2.2. Antisense lncRNAs

The lncRNAs located in the antisense orientation (in the opposite strand) to a protein-coding gene are classified as antisense lncRNAs or long non-coding natural antisense transcripts (lncNATs). They are expressed either in concordance with the sense strand or in a discordant manner. In the latter case, lncNATs are expressed when the sense strand expression is suppressed [96]. However in both cases, lncNATs mediate the transcriptional or post-transcriptional regulation of the corresponding sense strand transcript [96]. Some of the reported plant lncNATs involved in the regulation of plant morphology and the responses to biotic and abiotic stresses include COOLAIR, cis-NAT*PHO1;2*, SHO lnc-NAT and TL-lncRNA [29,30,57,138]. A group of cis- acting antisense lncRNAs termed COOLAIR direct histone modifications that contribute to the establishment of epigenetic memory required for the regulation of flowering [88,116]. This results in the transition of FLC gene from an actively transcribed to a stably repressed state. Histone methylation at lysine 36 (H3K36me3) plays an antagonistic role to H3K27me3, and in the absence of COOLAIR, the epigenetic repressive mark H3K27me3 is replaced by the active transcription mark H3K36me3 [14,116] suggesting the crucial role of antisense lncRNA COOLAIR in silencing of FLC and the regulation of flowering time.

Another example of cis-acting antisense lncRNA is the TWISTED LEAD (TL) [57] lncRNA present in rice. This lncRNA is also transcribed from the opposite strand of the R2R3 MYBP transcription factor gene locus called OsMYB60. This antisense lncRNA is involved in regulating the development of leaf morphology in rice by mediating the suppression of the sense gene through chromatin modifications. Down-regulation of TL leads to the over-expression of OsMYB60 resulting in twisted leaf blades in transgenic rice plants [57].

Other well characterized plant trans-acting antisense lncRNAs include HIDDEN TREASURE 1 (HID1) and the antisense Heat shock factor (HSF) B2a (Table 1) [20] HID1 was reported to be involved in the regulation of seedling photo-morphogenesis in *Arabidopsis* [103]. This lncRNA promotes photo-morphogenesis in the presence of continuous red light by suppressing the function of a transcription factor, *PHYTOCHROME-INTERACTING FACTOR 3* (*PIF3*), which is a key repressor of photo-morphogenesis in the presence of red light [56]. It was reported that the knockdown of HID1 resulted in the increased expression of PIF3, which in turn led to the development of elongated hypocotyls in the presence of continuous red light [103]. It has also been reported that HID1 is well conserved among different plant species. The HID1 homolog present in rice (OSHID1) was able to rescue the phenotype of mutant *hid1 Arabidopsis* plants [104]. HID1 was the first reported lncRNA to be involved in photo-morphogenesis, proving the importance of lncRNA in seedling development [56].

2.2.3. Intergenic lncRNAs

Long intergenic noncoding RNAs (lincRNAs) are located outside protein-coding genes. Plant lincRNAs are functionally diverse lncRNAs and have remained a major focus of lncRNA research, which has led to the identification of over 2000 plant specific lincRNAs [55,114] in a wide variety of crop species including wheat, maize, rice, sorghum, peach, potato and populus [11,44,48,71,100,112,128]. Although plant lincRNAs are reported to be expressed at low levels in both organ-specific as well as in cell-type specific transcriptome studies [50,55], the advent of high throughput technologies has led to the identification of several plant lincRNAs such as APOLO, ASCO, ENOD40, IPS1, DRIR and LDMAR [2,5,13,17,22,72]. The recently identified ELENA1 (ELF18-Induced Long Noncoding RNA 1) is a lincRNA involved in the regulation of plant innate immunity. As a trans-acting lincRNA, ELENA1 regulates the levels of PR1 (pathogenesis-related 1) to enhance resistance against *Pseudomonas syringae* infection [66,79].

Another example of lincRNAs in plants is APOLO (Auxin regulated Promoter Loop RNA) that is transcribed by two polymerases, RNA Pol II/Pol V, and induces the silencing of the neighboring Auxin transport regulator PID (PINOID). APOLO mediates PINOID silencing by chromatin folding, loop formation and promoter methylation [2]. PID is involved in the maintenance of root architecture by modulating Auxin transport through multiple channels including alteration of genome topology and RNA-Dependent DNA Methylation [2]. However, the possibility of regulatory mechanisms other than RdDM have also been proposed for the regulation of APOLO and PID [77].

2.3. In-silico identification

The development of high throughput technologies has aided the process of *in-silico* identification and characterization of plant lncRNAs [6]. However, the difficulty in separating the coding from the non-coding RNAs in the thousands of assembled transcripts makes the identification of lncRNAs very challenging [85]. Therefore the un-annotated plant lncRNAs are initially identified *in-silico* using distinctive features such as the absence of an ORF or start and stop codons [43,119]. Several tools are used for the *in-silico* identification of lncRNAs which include CONC (Coding Or Non-Coding), Phylo Codon Substitution Frequency (CSF), CPC (Coding Potential Calculator), and POTRAIT [4,40,52,54]. Several new tools specifically identify different types of lncRNAs. These include Coding-Non-Coding Index (CNCI), iSeeRNA (for lincRNAs), PLEK, lncRNA-MFDL [27,47,85]. Apart from direct transcriptome analysis, lncRNAs in mammals have been identified through specific chromatin modification signatures, such as the K4 and K36 methylation marks [56]. In case of mammals, regions within the chromatin that are actively transcribed by RNA polymerase II are marked by H3K4 and H3K36 methylation [37]. Since lncRNAs are also synthesized by RNA polymerase II, it was hypothesized that they might also contain the K4-K36 methylation signatures [37]. Initially this approach was applied to mammals, where the presence of such tags helped in the identification of lncRNAs present within the intergenic regions [9,56]. Later on, the approach was also used to successful identify lncRNAs in *A. thaliana* [56,76,127].

2.4. Databases

A large number of lncRNAs have been identified from plant species including *A. thaliana*, *Medicago truncatula* (barrel clover), *Oryza sativa* (Rice), *Petunia hybrida* (Petunia), *Solanum lycopersicum* (tomato), *Triticum aestivum* (wheat), *Zea mays* (corn), etc. ([55,56,98]a; [35]). In order to store this huge amount of information and further help in the mapping and profiling of putative lncRNAs, several databases have

been developed. These include TAIR, Plant NATsDB, lncRNAdb, NONCODE, PlncDB, PNRD, PLNlnRbase, GreeNc and CANTATdb [6,68]. Out of these, TAIR or The *Arabidopsis* Information Resource (contains structure and function annotation of *Arabidopsis* genes) and PlncBD or Plant long noncoding RNA database (contains intergenic transcripts) are specific for Arabidopsis [31,87]. NONCODE is not entirely plant specific and contains information on ncRNAs from different species [129]. The PlantNATsDB and lnRNAdb on the other hand are plant specific but contain limited information. PlantNATsDB contain sequences and information on long noncoding antisense transcripts present in 70 different plant species while lnRNAdb contains information on known functional lncRNAs from 9 plant species [10,73]. In comparison to these plant lncRNA databases, the recently released and updated databases particularly PLNlnRbase, GreeNc and CANTATdb, contain comprehensive details on lncRNAs present in a wide range of plant species [8,28]. The PLNlnRbase contains information on experimentally validated lncRNAs from 43 different plants species, while GreeNc and CANTATdb contain sequences of putative lncRNAs present in 10 and 37 different plant species [69,89,113]. There is also a database known as PeTMbase which is specifically designed for the identification of plant endogenous target mimics [34].

2.5. Expression measurement

As mentioned earlier, the expression level of long non-coding RNAs is low enough to make their identification and differential expression profiling a challenge. Techniques commonly used for the identification of lncRNAs in plants are Microarrays and RNA-Seq. Relative merits and demerits of these techniques are discussed below.

2.5.1. Microarrays

Microarrays are widely used for gene expression analysis and preferred over RNA Seq for expression profiling of known transcripts because of its low cost (about 10 folds) and simplicity of data analysis. Based on probe design, microarrays are subdivided into non-tiling and tiling arrays. Non-tiling arrays (quassi whole genome arrays) have probes representing only annotated genes, exons or splice junctions and can be used to determine expression levels of known transcripts. For tiling arrays, the probes are designed in an overlapping fashion and cover the whole genome or entire length of the desired portion of the genome [45]. Tiling arrays can work for both annotated and un-annotated regions of the genome; hence novel transcripts can be identified by using this method [65]. In tiling arrays, neighboring probes on a plate hybridize to overlapping regions of the genome thereby covering the entire region genomic region of interest. Tiling arrays were used to identify noncoding antisense transcripts for the Arabidopsis FLC locus [88]. Subsequently, many lncRNAs from Arabidopsis and Wheat were identified using tiling microarrays [55,62,74,112]. The major demerits of tiling microarrays, in addition to the need of prior sequence availability are cross hybridization of closely related sequences and reduced sensitivity due to high signal saturation [133]. For this reason, RNA-Seq is the method of choice for lncRNAs identification despite the associated high cost and analytical issues associated with complex data sets.

2.5.2. RNA-sequencing (RNA-Seq)

The development of high-throughput sequencing methods provide a means of quantifying the transcriptome by performing whole transcriptome sequencing or RNA-sequencing. The main steps in a typical RNA-seq workflow entail total RNA extraction, enrichment of RNA, library preparation and sequencing. For RNA enrichment, either polyA RNA selection/enrichment using oligo (dT) primers or the depletion of ribosomal RNA (rRNA depletion) may be carried out. While PolyA RNA-Seq is a cost-effective technology, when applied for lncRNA expression it accounts for approximately 1/3rd of the total lncRNAs within a species. Whole transcriptome sequencing by Ribosomal RNA (rRNA) depletion; a method that retains the complete transcriptional repertoire present in a sample (Both poly-A+ and poly-A− transcripts) but is depleted for ribosomal RNA is routinely used for lncRNA discovery. After enrichment of RNA by oligo dT or rRNA depletion, the RNA is fragmented (200-500 bp) and an adapter is ligated to each fragment. RNA fragments are then reverse transcribed to cDNA. This is followed by high-throughput sequencing, which results in short sequences from both ends (pair-end sequencing) or from one end (single-end sequencing). Following sequencing, the resulting reads are processed to obtain the level of expression and/or the transcriptional structure for each gene. Several modifications to these RNA-seq methods are routinely used to identify the full repertoire of lncRNAs within a plant species. Few of these methods are described below:

2.5.3. Strand specific RNA-seq

Conventional RNA-Seq as described above does not retain information about the strand of DNA that a given RNA is reverse transcribed from. Given the preponderance of anti-sense and other non-canonical RNAs, the originating strand information can significantly enhance the discovery of transcripts in a sequencing experiment. For instance, such information could aid in the identification of sense and antisense transcripts, demarcate boundaries between neighboring genes transcribed from both strands, and determine the precise expression levels of both non-coding and coding overlapping transcripts. Strand Specific RNA-Seq (ssRNA-Seq) is a form of RNA-Seq, which retains the identity of the strand of DNA (sense or antisense). There are various approaches to perform ssRNA-Seq and the most commonly used is the dUTP method (Fig. 3) [46,64]. Here the thymine nucleotide is replaced with uracil during second strand cDNA synthesis. As a result, the original strand is composed of thymine while the complementary strand contains uracil. Strands containing uracil are degraded by the

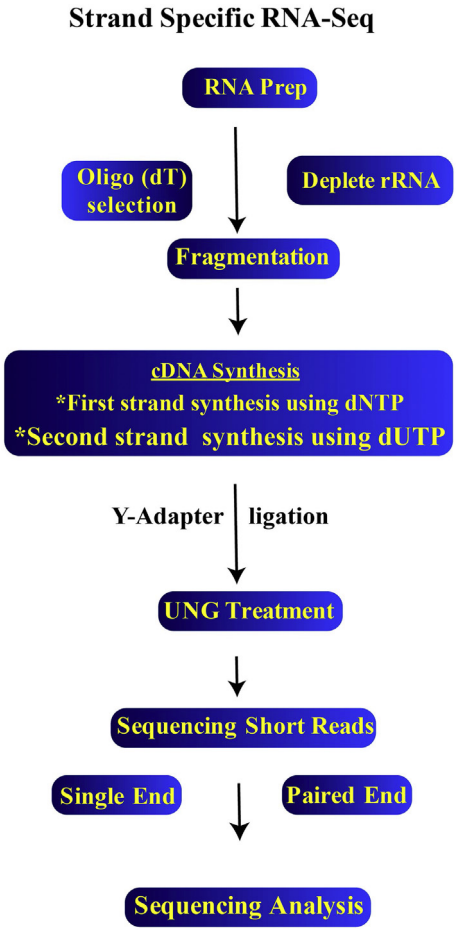


Fig. 3. Schematic diagram showing important steps in strand specific RNA seq.

Table 2
List of techniques used for the identification of lncRNAs in plants.

Mode of Identification	Species studied
RNA Seq	Tomato [15] Pea [36] Maize [39,126] Sea Buckthorn [123] Wheat [81,125] Sunflower [21] Populus [11,12,80,90] Chinese Cabbage [82,120] Arabidopsis [55,121] Rice [24] Canola [32]
Deep RNA-Seq	Kiwi Fruit [105] Soybean [118] Millet [71] Peach [100]
Strand specific RNA Seq	Populus [102] *Selaginella moellendorffii* [135] Tomato [97,101,132]; [78,131]) Orange [94] Maize [95] Arabidopsis [59,95,96,134] Potato [44] Cotton ([98]; [137]) Rice ([58]; [128]; [95])
Microarray	Wheat [112] Arabidopsis [55]

subsequent treatment with Uracil DNA Glycosylase (UDG) so that only the original strand is sequenced. Comparing the sequence to the reference genome can track the strand used in transcription. Novel lncRNAs have been identified in many plant species using strand specific RNA-Seq. For example, long noncoding natural antisense transcripts (lncNATs) have been identified in a number of plants using this method [58,96]. A detailed catalog of plant lncRNAs is presented in Table 2.B).

2.5.4. RNA immune precipitation- sequencing (RIP-seq)

lncRNAs perform most of their biological function by binding and interacting with proteins. To identify lncRNAs and characterize their function, RNA immunoprecipitation (RIP) is used. RIP uses a protein as bait to pull-down RNAs. RNA-protein complexes are immunoprecipitated using an antibody targeting the protein of interest that was purified under physiological conditions in order to preserve native interactions. Following the capture of RNA-protein complexes, high-throughput sequencing (RIP-seq) is carried out [41]. The identification of the associated RNAs enables a transcriptome-wide view of the protein–RNA/lncRNA regulatory network. As a proof of principle, RIP was used to identify Pol V transcripts throughout the Arabidopsis genome. RIP for Pol V-associated RNAs was performed using an antibody against NRPE1, the largest subunit of Pol V [7]. This was followed by high-throughput sequencing allowing for the identification of POL V associated ncRNAs in Arabidopsis.

2.5.5. Single cell RNA-seq

Analysis in biological systems has routinely been conducted at the level of the organism, organ or tissue. This analysis may often mask unique biological attributes of individual cells/cell-types that comprise the organism, organ or the tissue. Single cell RNA-seq (scRNA-seq) attempts to circumvent this issue. The scRNA-seq protocol involves six prominent methods namely Cell Expression by Linear amplification and Sequencing (CEL-seq), Droplet- sequencing (Drop-seq), Massively parallel single-cell RNA-sequencing (MARS-seq), Single cell RNA barcoding and sequencing (SCRB-seq), Switch mechanism at the 5′ end of RNA template (Smart-seq, and Smart-seq2). All these techniques involve selection of RNA (PolyA) followed by reverse transcription and cDNA amplification. With the exception of Cel-Seq and MARS-seq, which involve an additional *In vitro* transcription step, these methods involve direct PCR amplification of the cDNAs. Methods other than Smart-seq utilize oligos with unique molecular identifiers (UMIs) at the initial cDNA synthesis stage. While Smart-seq2 is capable of detecting the most number of genes across cells and per cell, it suffers from high experimental noise. On the other hand, Drop-seq, CEL-seq2, SCRB-seq and MARS-seq employ UMIs, which result in a better signal/noise ratio. In terms of cost, SCRB-seq, MARS-seq and Smart-seq2 are more cost effective when analyzing fewer cells while Drop-seq is more cost-efficient when analyzing a large number of cells. These single cell RNA seq methods have been used to analyze gene expression patterns from different cell-types from *Zea mays*, *A. thaliana, Medicago truncatula*, rice and *Glycine* max [18,19,23,136]. ScRNA-seq has not been used in plants to study lncRNAs mainly due to the inability to capture and sequence non-polyA RNA and low sequencing coverage. However, given the cell and tissue specific roles of lncRNAs in plants, scRNA-seq could be used for the functional identification of cell specific lncRNAs.

2.5.6. RNA-sequence analysis

Sequencing data from RNA-Seq can be analyzed by aligning raw sequencing data to the referenced genome using software tools such as TopHat, Bowtie, HISAT etc. [92]. Aligned Reads are counted for overlap of reads with genes using HT-seq or other counting software and differentially expressed genes are obtained using edgeR or Desq2 [1,75,93]. However, identification of novel transcripts expressing at low levels, such as lncRNAs, requires more sequencing depth (number of times a sample is sequenced) and high coverage (number of reads obtained after sequencing). High coverage ($>$ 60 million reads) avoids any biases due to uneven sequencing depth or heterogeneity of reads. Deep RNA-Seq involves repeated sequencing of the cDNA fragments and it has been used to identify novel lncRNAs in many plant species (Table 2).

3. Conclusion

The development of high throughput sequencing techniques and bioinformatics analysis has aided in the identification of lncRNAs. To date, a number of functionally important lncRNAs from different species have been reported; however, research in the case of plants is still in its infancy. The reported plant lncRNAs belong to a limited range of species, however their functional relevance stipulates further research. Key areas of future plant lncRNA research include their roles in cell type specific gene regulation and genome organization through chromatin remodeling. LncRNAs could also be exploited as target mimics in transient assays for functional identification of genes or miRNAs.

References

[1] S. Anders, P.T. Pyl, W. Huber, HTSeq—a Python framework to work with high-throughput sequencing data, Bioinformatics 31 (2015) 166–169.
[2] F. Ariel, T. Jegu, D. Latrasse, N. Romero-Barrios, A. Christ, M. Benhamed, M. Crespi, Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop, Mol. Cell 55 (2014) 383–396.
[3] F. Ariel, N. Romero-Barrios, T. Jégu, M. Benhamed, M. Crespi, Battles and hijacks: noncoding transcription in plants, Trends Plant Sci. 20 (2015) 362–371.
[4] R.T. Arrial, R.C. Togawa, M. Brigido Mde, Screening non-coding RNAs in transcriptomes from neglected species using PORTRAIT: case study of the pathogenic fungus paracoccidioides brasiliensis, BMC Bioinforma. 10 (2009) 239.
[5] F. Bardou, F. Ariel, C.G. Simpson, N. Romero-Barrios, P. Laporte, S. Balzergue, J.W. Brown, M. Crespi, Long noncoding RNA modulates alternative splicing regulators in *Arabidopsis*, Dev. Cell 30 (2014) 166–176.
[6] G. Bhatia, N. Goyal, S. Sharma, S. Upadhyay, K. Singh, Present scenario of long non-coding RNAs in plants, Non-Coding RNA 3 (2017) 16.
[7] G. Bohmdorfer, S. Sethuraman, M.J. Rowley, M. Krzyszton, M.H. Rothi, L. Bouzit, A.T. Wierzbicki, Long non-coding RNA produced by RNA polymerase V determines boundaries of heterochromatin, elife (2016) 5.
[8] E. Bonnet, Y. Van de Peer, P. Rouze, The small RNA world of plants, New Phytol. 171 (2006) 451–468.
[9] S. Cawley, S. Bekiranov, H.H. Ng, P. Kapranov, E.A. Sekinger, D. Kampa, A. Piccolboni, V. Sementchenko, J. Cheng, A.J. Williams, R. Wheeler, B. Wong,

J. Drenkow, M. Yamanaka, S. Patel, S. Brubaker, H. Tammana, G. Helt, K. Struhl, T.R. Gingeras, Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs, Cell 116 (2004) 499–509.

[10] D. Chen, C. Yuan, J. Zhang, Z. Zhang, L. Bai, Y. Meng, L.L. Chen, M. Chen, PlantNATsDB: a comprehensive database of plant natural antisense transcripts, Nucleic Acids Res. 40 (2012) D1187–1193.

[11] J. Chen, M. Quan, D. Zhang, Genome-wide identification of novel long non-coding RNAs in Populus tomentosa tension wood, opposite wood and normal wood xylem by RNA-seq, Planta 241 (2015) 125–143.

[12] M. Chen, C. Wang, H. Bao, H. Chen, Y. Wang, Genome-wide identification and characterization of novel lncRNAs in populus under nitrogen deficiency, Mol. Gen. Genomics. 291 (2016) 1663–1680.

[13] M.D. Crespi, E. Jurkevitch, M. Poiret, Y. d'Aubenton-Carafa, G. Petrovics, E. Kondorosi, A. Kondorosi, enod40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth, EMBO J. 13 (1994) 5099–5112.

[14] T. Csorba, J.I. Questa, Q. Sun, C. Dean, Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 16160–16165.

[15] J. Cui, Y. Luan, N. Jiang, H. Bao, J. Meng, Comparative transcriptome analysis between resistant and susceptible tomato allows the identification of lncRNA16397 conferring resistance to Phytophthora infestans by co-expressing glutaredoxin, Plant J. 89 (2017) 577–589.

[16] A.P. de Koning, W. Gu, T.A. Castoe, M.A. Batzer, D.D. Pollock, Repetitive elements may comprise over two-thirds of the human genome, PLoS Genet. 7 (2011) e1002384.

[17] J. Ding, Q. Lu, Y. Ouyang, H. Mao, P. Zhang, J. Yao, C. Xu, X. Li, J. Xiao, Q. Zhang, A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 2654–2659.

[18] I. Efroni, K.D. Birnbaum, The potential of single-cell profiling in plants, Genome Biol. 17 (2016) 65.

[19] I. Efroni, P.L. Ip, T. Nawy, A. Mello, K.D. Birnbaum, Quantification of cell identity from single-cell gene expression profiles, Genome Biol. 16 (2015) 9.

[20] H. Fedak, M. Palusinska, K. Krzyczmonik, L. Brzezniak, R. Yatusevich, Z. Pietras, S. Kaczanowski, S. Swiezewski, Control of seed dormancy in *Arabidopsis* by a cis-acting noncoding antisense transcript, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) E7846–E7855.

[21] N.M. Florez-Zapata, M.H. Reyes-Valdes, O. Martinez, Long non-coding RNAs are major contributors to transcriptome changes in sunflower meiocytes with different recombination rates, BMC Genomics 17 (2016) 490.

[22] J.M. Franco-Zorrilla, A. Valli, M. Todesco, I. Mateos, M.I. Puga, I. Rubio-Somoza, A. Leyva, D. Weigel, J.A. Garcia, J. Paz-Ares, Target mimicry provides a new mechanism for regulation of microRNA activity, Nat. Genet. 39 (2007) 1033–1037.

[23] Y. Han, X. Chu, H. Yu, Y.-K. Ma, X.-J. Wang, W. Qian, Y. Jiao, Single-cell transcriptome analysis reveals widespread monoallelic gene expression in individual rice mesophyll cells, Sci. Bull. 62 (2017) 1304–1314.

[24] F. He, Q. Liu, L. Zheng, Y. Cui, Z. Shen, L. Zheng, RNA-Seq analysis of rice roots reveals the involvement of post-transcriptional regulation in response to cadmium stress, Front. Plant Sci. 6 (2015) 1136.

[25] M.A. Held, B. Penning, A.S. Brandt, S.A. Kessans, W. Yong, S.R. Scofield, N.C. Carpita, Small-interfering RNAs from natural antisense transcripts derived from a cellulose synthase gene modulate cell wall biosynthesis in barley, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 20534–20539.

[26] R. Henriques, H. Wang, J. Liu, M. Boix, L.F. Huang, N.H. Chua, The antiphasic regulatory module comprising CDF5 and its antisense RNA FLORE links the circadian clock to photoperiodic flowering, New Phytol. 216 (2017) 854–867.

[27] J.B. Heo, S. Sung, Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA, Science 331 (2011) 76–79.

[28] J.B. Heo, Y.S. Lee, S. Sung, Epigenetic regulation by long noncoding RNAs in plants, Chromosom. Res. 21 (2013) 685–693.

[29] R. Ietswaart, S. Rosa, Z. Wu, C. Dean, M. Howard, Cell-size-dependent transcription of FLC and its antisense long non-coding RNA COOLAIR explain cell-to-cell expression variation, Cell Syst 4 (2017) 622–635 (e629).

[30] M. Jabnoune, D. Secco, C. Lecampion, C. Robaglia, Q. Shu, Y. Poirier, A rice cis-natural antisense RNA acts as a translational enhancer for its cognate mRNA and contributes to phosphate homeostasis and plant fitness, Plant Cell 25 (2013) 4166–4182.

[31] J. Jin, J. Liu, H. Wang, L. Wong, N.H. Chua, PLncDB: plant long non-coding RNA database, Bioinformatics 29 (2013) 1068–1071.

[32] T. Kanno, B. Huettel, M.F. Mette, W. Aufsatz, E. Jaligot, L. Daxinger, D.P. Kreil, M. Matzke, A.J. Matzke, Atypical RNA polymerase subunits required for RNA-directed DNA methylation, Nat. Genet. 37 (2005) 761–765.

[33] G. Karakulah, T. Unver, Barley Long Non-Coding RNAs and Their Tissue-Specific Co-expression Pattern with Coding-Transcripts, BioRxiv (2017), http://dx.doi.org/10.1101/229559.

[34] G. Karakulah, K. Yucebilgili Kurtoglu, T. Unver, PeTMbase: a database of plant endogenous target mimics (eTMs), PLoS One 11 (2016) e0167698.

[35] K. Kashi, L. Henderson, A. Bonetti, P. Carninci, Discovery and functional analysis of lncRNAs: methodologies to investigate an uncharacterized transcriptome, Biochim. Biophys. Acta 1859 (2016) 3–15.

[36] S.C. Kerr, F. Gaiti, C.A. Beveridge, M. Tanurdzic, De novo transcriptome assembly reveals high transcriptional complexity in *Pisum sativum* axillary buds and shows rapid changes in expression of diurnally regulated genes, BMC Genomics 18 (2017) 221.

[37] A.M. Khalil, M. Guttman, M. Huarte, M. Garber, A. Raj, D. Rivea Morales, K. Thomas, A. Presser, B.E. Bernstein, A. van Oudenaarden, A. Regev, E.S. Lander, J.L. Rinn, Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 11667–11672.

[38] E.D. Kim, S. Sung, Long noncoding RNA: unveiling hidden layer of gene regulatory networks, Trends Plant Sci. 17 (2012) 16–21.

[39] E.D. Kim, Y. Xiong, Y. Pyo, D.H. Kim, B.H. Kang, S. Sung, Spatio-temporal analysis of coding and long noncoding transcripts during maize endosperm development, Sci. Rep. 7 (2017) 3838.

[40] L. Kong, Y. Zhang, Z.Q. Ye, X.Q. Liu, S.Q. Zhao, L. Wei, G. Gao, CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine, Nucleic Acids Res. 35 (2007) W345–349.

[41] T. Koster, K. Meyer, Plant Ribonomics: proteins in search of RNA partners, Trends Plant Sci. 23 (2018) 352–365.

[42] H. Kouchi, K.-i. Takane, R.B. So, J.K. Ladha, P.M. Reddy, Rice ENOD40: isolation and expression analysis in rice and transgenic soybean root nodules, Plant J. 18 (1999) 121–129.

[43] J.T. Kung, D. Colognori, J.T. Lee, Long noncoding RNAs: past, present, and future, Genetics 193 (2013) 651–669.

[44] S. Kwenda, P.R. Birch, L.N. Moleleki, Genome-wide identification of potato long intergenic noncoding RNAs responsive to pectobacterium carotovorum subspecies brasiliense infection, BMC Genomics 17 (2016) 614.

[45] C. Lee, N. Kikyo, Strategies to identify long noncoding RNAs involved in gene regulation, Cell Biosci. 2 (2012) 37.

[46] J.Z. Levin, M. Yassour, X. Adiconis, C. Nusbaum, D.A. Thompson, N. Friedman, A. Gnirke, A. Regev, Comprehensive comparative analysis of strand-specific RNA sequencing methods, Nat. Methods 7 (2010) 709–715.

[47] A. Li, J. Zhang, Z. Zhou, PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme, BMC Bioinforma. 15 (2014) 311.

[48] L. Li, S.R. Eichten, R. Shimizu, K. Petsch, C.T. Yeh, W. Wu, A.M. Chettoor, S.A. Givan, R.A. Cole, J.E. Fowler, M.M. Evans, M.J. Scanlon, J. Yu, P.S. Schnable, M.C. Timmermans, N.M. Springer, G.J. Muehlbauer, Genome-wide discovery and characterization of maize long non-coding RNAs, Genome Biol. 15 (2014) R40.

[49] R. Li, H. Zhu, Y. Luo, Understanding the functions of long non-coding RNAs through their higher-order structures, Int. J. Mol. Sci. 17 (2016).

[50] S. Li, M. Yamada, X. Han, U. Ohler, P.N. Benfey, High-resolution expression map of the *Arabidopsis* root reveals alternative splicing and lincRNA regulation, Dev. Cell 39 (2016) 508–522.

[51] S. Li, L.E. Vandivier, B. Tu, L. Gao, S.Y. Won, S. Li, B. Zheng, B.D. Gregory, X. Chen, Detection of pol IV/RDR2-dependent transcripts at the genomic scale in *Arabidopsis* reveals features and regulation of siRNA biogenesis, Genome Res. 25 (2015) 235–245.

[52] M.F. Lin, I. Jungreis, M. Kellis, PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions, Bioinformatics 27 (2011) i275–282.

[53] C. Liu, U.S. Muchhal, K.G. Raghothama, Differential expression of TPS11, a phosphate starvation-induced gene in tomato, Plant Mol. Biol. 33 (1997) 867–874.

[54] J. Liu, J. Gough, B. Rost, Distinguishing protein-coding from non-coding RNAs through support vector machines, PLoS Genet. 2 (2006) e29.

[55] J. Liu, C. Jung, J. Xu, H. Wang, S. Deng, L. Bernad, C. Arenas-Huertero, N.H. Chua, Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in *Arabidopsis*, Plant Cell 24 (2012) 4333–4345.

[56] X. Liu, L. Hao, D. Li, L. Zhu, S. Hu, Long non-coding RNAs and their biological roles in plants, Genomics Proteomics Bioinformatics 13 (2015) 137–147, http://dx.doi.org/10.1016/j.pbi.2015.08.003.

[57] X. Liu, D. Li, D. Zhang, D. Yin, Y. Zhao, C. Ji, X. Zhao, X. Li, Q. He, R. Chen, S. Hu, L. Zhu, A novel antisense long noncoding RNA, TWISTED LEAF, maintains leaf blade flattening by regulating its associated sense R2R3-MYB gene in rice, New Phytol. 218 (2018) 774–788.

[58] T. Lu, C. Zhu, G. Lu, Y. Guo, Y. Zhou, Z. Zhang, Y. Zhao, W. Li, Y. Lu, W. Tang, Q. Feng, B. Han, Strand-specific RNA-seq reveals widespread occurrence of novel cis-natural antisense transcripts in rice, BMC Genomics 13 (2012) 721.

[59] Z. Lu, X. Xia, B. Jiang, K. Ma, L. Zhu, L. Wang, B. Jin, Identification and characterization of novel lncRNAs in *Arabidopsis thaliana*, Biochem. Biophys. Res. Commun. 488 (2017) 348–354.

[60] L. Ma, V.B. Bajic, Z. Zhang, On the classification of long non-coding RNAs, RNA Biol. 10 (2013) 925–933.

[61] J. Mach, The long-noncoding RNA ELENA1 functions in plant immunity, Plant Cell 29 (2017) 916.

[62] A. Matsui, J. Ishida, T. Morosawa, M. Okamoto, J.M. Kim, Y. Kurihara, M. Kawashima, M. Tanaka, To, T.K, K. Nakaminami, E. Kaminuma, T.A. Endo, Y. Mochizuki, S. Kawaguchi, N. Kobayashi, K. Shinozaki, T. Toyoda, M. Seki, *Arabidopsis* tiling array analysis to identify the stress-responsive genes, Methods Mol. Biol. 639 (2010) 141–155.

[63] J.S. Mattick, Non-coding RNAs: the architects of eukaryotic complexity, EMBO Rep. 2 (2001) 986–991.

[64] J.D. Mills, Y. Kawahara, M. Janitz, Strand-specific RNA-Seq provides greater resolution of transcriptome profiling, Curr. Genom. 14 (2013) 173–181.

[65] T.C. Mockler, S. Chan, A. Sundaresan, H. Chen, S.E. Jacobsen, J.R. Ecker, Applications of DNA tiling arrays for whole-genome analysis, Genomics 85 (2005) 1–15.

[66] A. Moieni, M. Salehi, N. Safaie, Novel medium for enhancing callus growth in hazel (Corylus avellana L.), J. Plant Physiol. Pathol. 05 (2017).

[67] R.A. Mosher, F. Schwach, D. Studholme, D.C. Baulcombe, PolIVb influences RNA-

directed DNA methylation independently of its role in siRNA biogenesis, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 3145–3150.
[68] N. Nejat, N. Mantri, Emerging roles of long non-coding RNAs in plant response to biotic and abiotic stresses, Crit. Rev. Biotechnol. (2017) 1–13.
[69] A. Paytuvi Gallart, A. Hermoso Pulido, I. Anzar Martinez de Lagran, W. Sanseverino, R. Aiese Cigliano, GREENC: a Wiki-based database of plant lncRNAs, Nucleic Acids Res. 44 (2016) D1161–1166.
[70] K.V. Prasanth, D.L. Spector, Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum, Genes Dev. 21 (2007) 11–42.
[71] X. Qi, S. Xie, Y. Liu, F. Yi, J. Yu, Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing, Plant Mol. Biol. 83 (2013) 459–473.
[72] T. Qin, H. Zhao, P. Cui, N. Albesher, L. Xiong, A nucleus-localized long non-coding RNA enhances drought and salt stress tolerance, Plant Physiol. 175 (2017) 1321–1336.
[73] X.C. Quek, D.W. Thomson, J.L. Maag, N. Bartonicek, B. Signal, M.B. Clark, B.S. Gloss, M.E. Dinger, lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs, Nucleic Acids Res. 43 (2015) D168–173.
[74] H. Rehrauer, C. Aquino, W. Gruissem, S.R. Henz, P. Hilson, S. Laubinger, N. Naouar, A. Patrignani, S. Rombauts, H. Shu, Y. Van de Peer, M. Vuylsteke, D. Weigel, G. Zeller, L. Hennig, AGRONOMICS1: a new resource for *Arabidopsis* transcriptome profiling, Plant Physiol. 152 (2010) 487–499.
[75] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics 26 (2010) 139–140.
[76] F. Roudier, I. Ahmed, C. Berard, A. Sarazin, T. Mary-Huard, S. Cortijo, D. Bouyer, E. Caillieux, E. Duvernois-Berthet, L. Al-Shikhley, L. Giraut, B. Despres, S. Drevensek, F. Barneche, S. Derozier, V. Brunaud, S. Aubourg, A. Schnittger, C. Bowler, M.L. Martin-Magniette, S. Robin, M. Caboche, V. Colot, Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*, EMBO J. 30 (2011) 1928–1938.
[77] M.J. Rowley, M.H. Rothi, G. Bohmdorfer, J. Kucinski, A.T. Wierzbicki, Long-range control of gene expression via RNA-directed DNA methylation, PLoS Genet. 13 (2017) e1006749.
[78] D. Scarano, R. Rao, G. Corrado, In silico identification and annotation of non-coding RNAs by RNA-seq and De novo assembly of the transcriptome of tomato fruits, PLoS One 12 (2017) e0171504.
[79] J.S. Seo, H.-X. Sun, B.S. Park, C.-H. Huang, S.-D. Yeh, C. Jung, N.-H. Chua, ELF18-INDUCED LONG-NONCODING RNA associates with mediator to enhance expression of innate immune response genes in *Arabidopsis*, Plant Cell 29 (2017) 1024–1038.
[80] P. Shuai, D. Liang, S. Tang, Z. Zhang, C.Y. Ye, Y. Su, X. Xia, W. Yin, Genome-wide identification and functional prediction of novel and drought-responsive lincRNAs in Populus trichocarpa, J. Exp. Bot. 65 (2014) 4975–4983.
[81] Shumayla, S. Sharma, M. Taneja, S. Tyagi, K. Singh, S.K. Upadhyay, Survey of high throughput RNA-Seq data reveals potential roles for lncRNAs during development and stress response in bread wheat, Front. Plant Sci. 8 (2017) 1019.
[82] X. Song, G. Liu, Z. Huang, W. Duan, H. Tan, Y. Li, X. Hou, Temperature expression patterns of genes and their coexpression with LncRNAs revealed by RNA-Seq in non-heading Chinese cabbage, BMC Genomics 17 (2016) 297.
[83] C. Sousa, C. Johansson, C. Charon, H. Manyani, C. Sautter, A. Kondorosi, M. Crespi, Translational and structural requirements of the early nodulin gene enod40, a short-open reading frame-containing RNA, for elicitation of a cell-specific growth response in the alfalfa root cortex, Mol. Cell. Biol. 21 (2001) 354–366.
[84] H. Stroud, M.V. Greenberg, S. Feng, Y.V. Bernatavichute, S.E. Jacobsen, Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome, Cell 152 (2013) 352–364.
[85] K. Sun, X. Chen, P. Jiang, X. Song, H. Wang, H. Sun, iSeeRNA: identification of long intergenic non-coding RNA transcripts from transcriptome sequencing data, BMC Genomics 14 (Suppl. 2) (2013) S7.
[86] Q. Sun, T. Csorba, K. Skourti-Stathaki, N.J. Proudfoot, C. Dean, R-loop stabilization represses antisense transcription at the *Arabidopsis* FLC locus, Science 340 (2013) 619–621.
[87] D. Swarbreck, C. Wilks, P. Lamesch, T.Z. Berardini, M. Garcia-Hernandez, H. Foerster, D. Li, T. Meyer, R. Muller, L. Ploetz, A. Radenbaugh, S. Singh, V. Swing, C. Tissier, P. Zhang, E. Huala, The *Arabidopsis* information resource (TAIR): gene structure and function annotation, Nucleic Acids Res. 36 (2008) D1009–1014.
[88] S. Swiezewski, F. Liu, A. Magusin, C. Dean, Cold-induced silencing by long antisense transcripts of an *Arabidopsis* polycomb target, Nature 462 (2009) 799–802.
[89] M.W. Szczesniak, W. Rosikiewicz, I. Makalowska, CANTATAdb: a collection of plant long non-coding RNAs, Plant Cell Physiol. 57 (2016) e8.
[90] J. Tian, Y. Song, Q. Du, X. Yang, D. Ci, J. Chen, J. Xie, B. Li, D. Zhang, Population genomic analysis of gibberellin-responsive long non-coding RNAs in Populus, J. Exp. Bot. 67 (2016) 2467–2482.
[91] M. Todesco, I. Rubio-Somoza, J. Paz-Ares, D. Weigel, A collection of target mimics for comprehensive analysis of microRNA function in *Arabidopsis thaliana*, PLoS Genet. 6 (2010) e1001031.
[92] C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks, Nat. Protoc. 7 (2012) 562–578.
[93] Y. Tu, T.S. Shimizu, H.C. Berg, Modeling the chemotactic response of *Escherichia coli* to time-varying stimuli, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 14855–14860.
[94] C.Y. Wang, S.R. Liu, X.Y. Zhang, Y.J. Ma, C.G. Hu, J.Z. Zhang, Genome-wide screening and characterization of long non-coding RNAs involved in flowering development of trifoliate orange (*Poncirus trifoliata* L. Raf.), Sci. Rep. 7 (2017) 43226.
[95] D. Wang, Z. Qu, L. Yang, Q. Zhang, Z.H. Liu, T. Do, D.L. Adelson, Z.Y. Wang, I. Searle, J.K. Zhu, Transposable elements (TEs) contribute to stress-related long intergenic noncoding RNAs in plants, Plant J. 90 (2017) 133–146.
[96] H. Wang, P.J. Chung, J. Liu, I.C. Jang, M.J. Kean, J. Xu, N.H. Chua, Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in *Arabidopsis*, Genome Res. 24 (2014) 444–453.
[97] H. Wang, Q.W. Niu, H.W. Wu, J. Liu, J. Ye, N. Yu, N.H. Chua, Analysis of non-coding transcriptome in rice and maize uncovers roles of conserved lncRNAs associated with agriculture traits, Plant J. 84 (2015) 404–416.
[98] J. Wang, W. YU, Y. Yang, X. Li, T. Chen, T. Liu, N. Ma, X. Yang, R. Liu, B. Zhang, Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection, Sci. Rep. 5 (2015) 16946.
[99] K.C. Wang, Y.W. Yang, B. Liu, A. Sanyal, R. Corces-Zimmerman, Y. Chen, B.R. Lajoie, A. Protacio, R.A. Flynn, R.A. Gupta, J. Wysocka, M. Lei, J. Dekker, J.A. Helms, H.Y. Chang, A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression, Nature 472 (2011) 120–124.
[100] L. Wang, S. Zhao, C. Gu, Y. Zhou, H. Zhou, J. Ma, J. Cheng, Y. Han, Deep RNA-Seq uncovers the peach transcriptome landscape, Plant Mol. Biol. 83 (2013) 365–377.
[101] M. Wang, D. Yuan, L. Tu, W. Gao, Y. He, H. Hu, P. Wang, N. Liu, K. Lindsey, X. Zhang, Long noncoding RNAs and their proposed functions in fibre development of cotton (*Gossypium spp.*), New Phytol. 207 (2015) 1181–1197.
[102] N. Wang, P. Cao, W. Xia, L. Fang, H. Yu, Identification and characterization of long non-coding RNAs in response to early infection by *Melampsora larici*-Populina using genome-wide high-throughput RNA sequencing, Tree Genet. Genomes 13 (2017).
[103] Y. Wang, X. Fan, F. Lin, G. He, W. Terzaghi, D. Zhu, X.W. Deng, *Arabidopsis* noncoding RNA mediates control of photomorphogenesis by red light, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 10359–10364.
[104] Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics, Nat. Rev. Genet. 10 (2009) 57–63.
[105] Z. Wang, Y. Liu, L. Li, D. Li, Q. Zhang, Y. Guo, S. Wang, C. Zhong, H. Huang, Whole transcriptome sequencing of *Pseudomonas syringae* pv. actinidiae-infected kiwifruit plants reveals species-specific interaction between long non-coding RNA and coding genes, Sci. Rep. 7 (2017) 4910.
[106] J. Wasaki, R. Yonetani, T. Shinano, M. Kai, M. Osaki, Expression of the OsPI1 gene, cloned from rice roots using cDNA microarray, rapidly responds to phosphorus status, New Phytol. 158 (2003) 239–248.
[107] J.M. Wendte, C.S. Pikaard, The RNAs of RNA-directed DNA methylation, Biochim. Biophys. Acta 1860 (2017) 140–148.
[108] A.T. Wierzbicki, J.R. Haag, C.S. Pikaard, Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes, Cell 135 (2008) 635–648.
[109] J.E. Wilusz, H. Sunwoo, D.L. Spector, Long noncoding RNAs: functional surprises from the RNA world, Genes Dev. 23 (2009) 1494–1504.
[110] H.J. Wu, Z.M. Wang, M. Wang, X.J. Wang, Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants, Plant Physiol. 161 (2013) 1875–1884.
[111] M. Wunderlich, R. Gross-Hardt, F. Schoffl, Heat shock factor HSFB2a involved in gametophyte development of *Arabidopsis thaliana* and its expression is controlled by a heat-inducible long non-coding antisense RNA, Plant Mol. Biol. 85 (2014) 541–550.
[112] M. Xin, Y. Wang, Y. Yao, N. Song, Z. Hu, D. Qin, C. Xie, H. Peng, Z. Ni, Q. Sun, 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. 11 (2011) 61.
[113] H. Xuan, L. Zhang, X. Liu, G. Han, J. Li, X. Li, A. Liu, M. Liao, S. Zhang, PLNlncRbase: a resource for experimentally identified lncRNAs in plants, Gene 573 (2015) 328–332.
[114] M. Yamada, Functions of long intergenic non-coding (linc) RNAs in plants, J. Plant Res. 130 (2017) 67–73.
[115] J. Yan, Y. Gu, X. Jia, W. Kang, S. Pan, X. Tang, Effective small RNA destruction by the expression of a short tandem target mimic in *Arabidopsis*, Plant Cell 24 (2012) 415–427.
[116] H. Yang, M. Howard, C. Dean, Antagonistic roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at *Arabidopsis* FLC, Curr. Biol. 24 (2014) 1793–1797.
[117] W.-C. Yang, P. Katinakis, P. Hendriks, A. Smolders, F. de Vries, J. Spee, A. van Kammen, T. Bisseling, H. Franssen, Characterization of GmENOD40, a gene showing novel patterns of cell-specific expression during soybean nodule development, Plant J. 3 (1993) 573–585.
[118] C.Y. Ye, H. Xu, E. Shen, Y. Liu, Y. Wang, Y. Shen, J. Qiu, Q.H. Zhu, L. Fan, Genome-wide identification of non-coding RNAs interacted with microRNAs in soybean, Front. Plant Sci. 5 (2014) 743.
[119] S. Yotsukura, D. duVerle, T. Hancock, Y. Natsume-Kitatani, H. Mamitsuka, Computational recognition for long non-coding RNA (lncRNA): software and databases, Brief. Bioinform. 18 (2017) 9–27.
[120] X. Yu, J. Yang, X. Li, X. Liu, C. Sun, F. Wu, Y. He, Global analysis of cis-natural antisense transcripts and their heat-responsive nat-siRNAs in Brassica rapa, BMC Plant Biol. 13 (2013) 208.
[121] J. Yuan, Y. Zhang, J. Dong, Y. Sun, B.L. Lim, D. Liu, Z.J. Lu, Systematic characterization of novel lncRNAs responding to phosphate starvation in *Arabidopsis thaliana*, BMC Genomics 17 (2016) 655.

[122] P.D. Zamore, B. Haley, Ribo-gnome: the big world of small RNAs, Science 309 (2005) 1519–1524.
[123] G. Zhang, A. Duan, J. Zhang, C. He, Genome-wide analysis of long non-coding RNAs at the mature stage of sea buckthorn (*Hippophae rhamnoides* Linn) fruit, Gene 596 (2017) 130–136.
[124] H. Zhang, J. Zhang, J. Yan, F. Gou, Y. Mao, G. Tang, J.R. Botella, J.K. Zhu, Short tandem target mimic rice lines uncover functions of miRNAs in regulating important agronomic traits, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) 5277–5282.
[125] H. Zhang, W. Hu, J. Hao, S. Lv, C. Wang, W. Tong, Y. Wang, Y. Wang, X. Liu, W. Ji, Genome-wide identification and functional prediction of novel and fungi-responsive lincRNAs in *Triticum aestivum*, BMC Genomics 17 (2016) 238.
[126] W. Zhang, Z. Han, Q. Guo, Y. Liu, Y. Zheng, F. Wu, W. Jin, Identification of maize long non-coding RNAs responsive to drought stress, PLoS One 9 (2014) e98958.
[127] X. Zhang, Y.V. Bernatavichute, S. Cokus, M. Pellegrini, S.E. Jacobsen, Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*, Genome Biol. 10 (2009) R62.
[128] Y.C. Zhang, J.Y. Liao, Z.Y. Li, Y. Yu, J.P. Zhang, Q.F. Li, L.H. Qu, W.S. Shu, Y.Q. Chen, Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice, Genome Biol. 15 (2014) 512.
[129] Y. Zhao, H. Li, S. Fang, Y. Kang, W. Wu, Y. Hao, Z. Li, D. Bu, N. Sun, M.Q. Zhang, R. Chen, NONCODE 2016: an informative and valuable data source of long non-coding RNAs, Nucleic Acids Res. 44 (2016) D203–208.
[130] B. Zheng, Z. Wang, S. Li, B. Yu, J.Y. Liu, X. Chen, Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*, Genes Dev. 23 (2009) 2850–2860.
[131] Y. Zheng, B. Ding, Z. Fei, Y. Wang, Comprehensive transcriptome analyses reveal tomato plant responses to tobacco rattle virus-based gene silencing vectors, Sci. Rep. 7 (2017) 9771.
[132] B. Zhu, Y. Yang, R. Li, D. Fu, L. Wen, Y. Luo, H. Zhu, RNA sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening, J. Exp. Bot. 66 (2015) 4483–4495.
[133] Q.H. Zhu, M.B. Wang, Molecular functions of long non-coding RNAs in plants, Genes (Basel) 3 (2012) 176–190.
[134] Q.H. Zhu, S. Stephen, J. Taylor, C.A. Helliwell, M.B. Wang, Long noncoding RNAs responsive to fusarium oxysporum infection in *Arabidopsis thaliana*, New Phytol. 201 (2014) 574–584.
[135] Y. Zhu, L. Chen, C. Zhang, P. Hao, X. Jing, X. Li, Global transcriptome analysis reveals extensive gene remodeling, alternative splicing and differential transcription profiles in non-seed vascular plant *Selaginella moellendorffii*, BMC Genomics 18 (2017) 1042.
[136] C. Ziegenhain, B. Vieth, S. Parekh, B. Reinius, A. Guillaumet-Adkins, M. Smets, H. Leonhardt, H. Heyn, I. Hellmann, W. Enard, Comparative analysis of single-cell RNA sequencing methods, Mol. Cell 65 (2017) 631–643 (e634).
[137] C. Zou, Q. Wang, C. Lu, W. Yang, Y. Zhang, H. Cheng, X. Feng, M.A. Prosper, G. Song, Transcriptome analysis reveals long noncoding RNAs involved in fiber development in cotton (*Gossypium arboreum*), Sci. China Life Sci. 59 (2016) 164–171.
[138] E. Zubko, P. Meyer, A natural antisense transcript of the petunia hybrida sho gene suggests a role for an antisense mechanism in cytokinin regulation, Plant J. 52 (2007) 1131–1139.