the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2025) **124,** e70591 doi: 10.1111/tpj.70591

RESOURCE

Constructing epigenetic regulatory landscapes of plant lncRNAs—an exploration utilizing the novel specialized platform PERlncDB

Yan Li[1,2,†], Wenjing Yang[1,†], Jiazhi Liu[2,3], Baolin Yao[4] and Changning Liu[1,*]

[1]*CAS Key Laboratory of Tropical Plant Resources and Sustainable Use, Yunnan Key Laboratory of Crop Wild Relatives Omics, State Key Laboratory of Plant Diversity and Specialty Crops, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Kunming 650223, China,*

[2]*University of Chinese Academy of Sciences, Beijing 100049, China,*

[3]*Germplasm Bank of Wild Species & Yunnan Key Laboratory of Crop Wild Relatives Omics, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan 650201, China, and*

[4]*School of Life Sciences, University of Science and Technology of China, Hefei 230026, China*

Received 20 January 2025; revised 30 September 2025; accepted 7 November 2025.

*For correspondence (e-mail liuchangning@xtbg.ac.cn).

†These authors contributed equally to this work.

SUMMARY

Long non-coding RNAs (lncRNAs), once overlooked as transcriptional byproducts, are now recognized for their crucial roles in plant growth, development, and stress responses, with increasing focus on their epigenetic regulation. However, studies investigating epigenomic signals to explore the functions of lncRNAs in plants remain relatively limited. This study collected a comprehensive dataset of over 160 000 high-quality lncRNAs from 19 representative plant species and integrated 6715 ChIP-seq, BS-seq, and RNA-seq datasets to analyze epigenomic patterns at lncRNA loci. Results showed elevated DNA methylation in lncRNA regions. The highest levels occurred in transposable element-associated lncRNAs. Additionally, activating histone modifications at lncRNA loci showed tissue specificity, with epigenetic preferences differed from those at protein-coding gene (PCG) loci. Differential site analysis in epigenetic mutants further highlighted the selective regulation of lncRNA loci by specific epigenetic factors. To facilitate research, we developed PERlncDB, a platform that provides species-specific lncRNA browsing, epigenetic annotation, cross-species conservation analysis, and visualization of epigenomic landscapes. Case studies on *MARS* and *LINC-AP2* emphasized the platform's utility. Conserved epigenetic mechanisms regulating lncRNAs across species, exemplified by a syntenic conserved MET1-regulated lncRNA pair in Arabidopsis and tomato, suggested the stability of regulatory mechanisms underlying lncRNA functions. This work provides critical insights and resources for understanding plant lncRNA epigenetic regulation.

Keywords: DNA methylation, epigenome, histone modification, LncRNAs, omics database, plant.

INTRODUCTION

Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nucleotides that lack protein-coding potential. Most lncRNAs are transcribed by RNA polymerase II, with a subset in plants being produced by the plant-specific RNA polymerases Pol IV and Pol V (Mattick et al., 2023). Based on their genomic locations relative to protein-coding genes (PCGs), three main types of plant lncRNAs are primarily studied: intergenic, intronic, and antisense lncRNAs. Plant lncRNAs are generally shorter, contain fewer exons (Yu et al., 2019), and often include highly variable transposable element-derived fragments (Sigman & Slotkin, 2016). However, lncRNAs exhibit significant nucleotide divergence at the interspecies level (Wang, Niu, et al., 2015; Zhou et al., 2022), but their genomic locations tend to be more conserved, underscoring their functional stability within biological systems (Mattick et al., 2023). In plants, lncRNAs often show tissue-specific

expression during particular developmental stages, with highly conserved lncRNAs being constitutively transcribed, while low-expression lncRNAs are typically tissue-specific (Deng et al., 2018). For example, 575 orthologous lncRNA pairs have been identified in *Arabidopsis*, whereas fewer were found in rice and its related species (Deng et al., 2018). In contrast to mRNAs, lncRNAs are preferentially enriched in certain tissues, demonstrating their functional importance (Cabili et al., 2011; Li et al., 2016; Liu et al., 2016). This highlights the distinct roles of lncRNAs in gene regulation, with their conserved genomic locations and structural features playing crucial roles in their regulatory functions across species.

Epigenetics refers to heritable variations caused by chromosomal changes without altering the DNA sequence. These variations involve marks like histone modifications, DNA methylation, chromatin accessibility, and non-coding RNAs. Among these, histone modifications and DNA methylation are the two most extensively studied and well-resourced types (Lu et al., 2018; Niederhuth et al., 2016; Zhao et al., 2020). Recent studies have shown that lncRNAs regulate chromatin structure, transcription, and RNA processing by interacting with DNA, RNA, and proteins (Statello et al., 2021). Based on their regulatory interactions with epigenetic modifiers, lncRNAs have been classified into two categories: those that collaborate with epigenetic modifiers and those that are controlled by them, thereby regulating target genes (Yang et al., 2023). lncRNAs have the ability to interact with histone modification enzymes, thereby influencing chromatin structure and gene expression. For example, *MAS* recruits WDR5a to *MAF4* genomic region, promoting H3K4me3 deposition and activating *MAF4* expression (Zhao et al., 2018). *LAIR* upregulates the expression of the *LRK1* by binding to the histone modification proteins OsMOF and OsWDR5, leading to the enrichment of H3K4me3 and H4K16ac in the activated *LRK1* genomic region and its non-coding regions (Wang et al., 2018). In response to phosphate starvation, lncRNA *At4* is directly regulated through H3K9/14 acetylation mediated by the histone acetyltransferase GCN5 (Wang et al., 2019). In addition, lncRNA expression is influenced by DNA methylation. *HOTAIR* can guide DNA methyltransferases (DNMTs) to specific gene loci, leading to increased gene methylation and subsequent gene repression (Wen et al., 2024). In rice, mutations in *LDMAR* affect DNA methylation at the promoter region of its locus, suppressing *LDMAR* expression and ultimately resulting in photoperiod-sensitive male sterility (Ding et al., 2012). Studying the interactions between lncRNAs and epigenetic modifications can provide valuable insights into the functional roles of lncRNAs.

With the rapid advancement of high-throughput sequencing technologies and the deepening of genomics research, a large number of lncRNAs have been discovered in the genomes of plants and animals, though only a small fraction has been experimentally validated. To better characterize and explore the mysteries of lncRNAs, the integration of data resources has pushed lncRNA-related research to a deeper level (Kornienko et al., 2023; Rai et al., 2019; Zhao et al., 2024). However, compared to research on animal lncRNAs (Gao, Li, et al., 2021; Gao, Shang, et al., 2021; Li et al., 2021; Mazurov et al., 2022; Zhao et al., 2016), the development of data platforms and resources for plant lncRNAs remains underdeveloped, especially in the utilization of epigenetic information. TAIR and RiceLncPedia provide valuable resources for Arabidopsis and rice lncRNAs, respectively (Berardini et al., 2015; Zhang et al., 2021), and others like LncRNAdb, PlantNATsDB, GreeNC, and CANTATAdb cover multiple species and offer sequence and functional annotation (Chen et al., 2012; Di Marsico et al., 2022; Quek et al., 2015; Szczesniak & Wanowska, 2024). PLncDB is one of the most comprehensive plant lncRNA databases offering expression data and some epigenomics data resources seem to be rather limited (Jin et al., 2021; Yang et al., 2023). Furthermore, platforms like CARMO, Plant Regulomics, and ChIP-Hub integrate multi-omics data to analyze gene functions but fail to address non-coding regions, particularly lncRNA loci (Fu et al., 2022; Ran et al., 2020; Wang, Qi, et al., 2015). Thus, there is a clear need to understand the characteristics of epigenomic information at lncRNA loci and to develop a specialized data platform dedicated to leveraging epigenetic modifications to discover the functions of plant lncRNAs. Compared to directly using experimental methods, utilizing data resources for systematic exploration not only saves time but also enables large-scale analysis, cross-species comparisons, and the identification of regulatory mechanisms that might be difficult to uncover through traditional experimental approaches alone.

This study focused on utilizing epigenetic modification information to explore potential functions of plant lncRNAs. We integrated high-quality lncRNA data from 19 species with large-scale datasets from ChIP-seq, BS-seq, and RNA-seq to identify specific epigenetic patterns on lncRNAs. Through genome-wide epigenomic signal analysis, we found that activating histone modification signals at lncRNA loci exhibit tissue specificity, while different types of epigenetic modifications at lncRNA loci within the same tissue display distinct preferences than protein-coding gene loci. Analysis of differential sites in various epigenetic modification mutants further indicates that lncRNA loci tend to be regulated by different epigenetic modification factors. Additionally, the DNA methylation levels in lncRNA regions are significantly higher than those in PCG regions. To further explore the epigenetic regulation of lncRNAs in plants, we developed a specialized platform PERlncDB, which browsing species-specific lncRNA, searching lncRNA epigenetic annotations, analyzing

cross-species conservation, visualizing epigenetic landscapes, and integrating other bioinformatics tools. Using two published lncRNAs, *MARS* and *LINC-AP2*, as examples, we presented the richness and accessibility of the existing data for lncRNA research. Notably, this study explored the functional mechanisms of a synteny-conserved lncRNA pair which is regulated by homologous MET1 in Arabidopsis and tomato, revealing that lncRNAs could be modulated through conserved epigenetic regulatory mechanisms across species. Our research provides valuable data resources and new insights for the in-depth exploration of plant lncRNA epigenetic regulation.

RESULTS

Global resources for high-throughput omics data on plant lncRNAs

To explore the epigenetic regulatory patterns at plant lncRNA loci, this study primarily collects plant regulomics data stored in public databases, including ChIP-seq, BS-seq, and RNA-seq (Figure 1a). Although most datasets were generated in model organisms such as *Arabidopsis thaliana* (Atha, *A. thaliana*), *Oryza sativa* (Osat, *O. sativa*), and *Zea mays* (Zmay, *Z. mays*), high-throughput regulomics experiments are also widely applied to non-model plants (Table S1). We manually curated all datasets by reviewing original publications and biological projects, covering genomic data from 19 species (25 Gb). Through literature review and global searches, a total of 6715 omics datasets were obtained, including 1826 BS-seq sample datasets (9 Tb), 4244 ChIP-seq sample datasets (19 Tb), and 646 RNA-seq sample datasets (2.7 Tb), with a total data volume of approximately 31 TB (Figure 1b). The ChIP-seq datasets included 1613 related to transcription factors (TFs), 1413 related to histone methylation, 590 related to histone acetylation, and the remaining 628 datasets related to histone variants and other modifications (Figure 1b). As well, RNA-seq data primarily focus on datasets collected under mutant or stress conditions.

LncRNA data for these species were downloaded from public databases, filtered, and screened using standardized criteria, resulting in over 160 000 high-quality lncRNA entries. When categorizing lncRNAs based on their location relative to protein-coding genes (PCGs) (Figure 1c), we found that a significant proportion of lncRNAs are located in intergenic regions, with over 75% in crops like soybean, maize, and wheat. Approximately 30% of the lncRNAs were of antisense type, while only 2–3% were intronic. Notably, in rice and Arabidopsis, more than half of the lncRNAs were antisense, possibly due to the higher gene density in their genomes.

Additionally, this study identified a total of 21,486,313 histone modification regions, 10,688,976 methylation sites, and 473,144 transcription factor binding sites associated with lncRNAs (Figure 1d). Among these sites, *A. thaliana* exhibited the highest number of annotated data types, followed by species such as rice, maize, and soybean. These multiple epigenetic marks distributed in the non-coding regions could provide valuable insights into lncRNA transcription and their functions.

Characteristics of histone modifications at lncRNA loci across plant species

We initially monitored histone modification marks at lncRNA loci and found that most lncRNAs could be annotated with histone modification marks. Among the collected samples such as *A. thaliana*, soybean, rice, and maize, approximately 90% of the lncRNAs exhibited a broad range of histone modification types, particularly focused on H3 histone modifications (Figure 2a, Figure S1). In contrast, fewer lncRNAs were annotated in apple and peach species, possibly due to the limited samples. To mediate the complex gene expression patterns in an organism, chromatin needs to undergo dynamic transitions between euchromatin (transcriptionally active) and heterochromatin (silenced state). During gene transcription regulation, methylation of the H3 histone at K4 and K36 sites is typically associated with transcriptional activation, while H3K9 and H3K27 modifications are thought to be related to transcriptional repression (Xu et al., 2008; Zhao et al., 2019). Additionally, histone acetylation modifications are more widespread than methylation modifications and are generally associated with gene transcriptional activation (Liu et al., 2014). Therefore, further analysis focused on the relationship between lncRNA transcription and histone modification levels. Using Arabidopsis leaves as an example, we found that the distribution of histone modification levels at PCG loci was similar to those described in previously published plant epigenomes (Lu et al., 2019) (Figure 2b), confirming the reliability of our data processing methods.

To systematically capture histone modification signals on a genome-wide scale, over 20 histone modifications and histone variants were analyzed using ChIP-seq samples in Arabidopsis. Most histone mark peaks in various tissues were enriched at PCG loci (Figure S2a,b). However, 8–75% functionally distinct regions associated with gene repression were located with lncRNA loci. As shown in Figure S2c, compared to activating marks, lncRNA loci exhibit significantly higher repressive histone marks enrichment, like H3K9me1, H3K9me2, H3K27me1, and H3K27me3. While only about 20% of the histone modification peaks were enriched on lncRNAs within each sample, a cumulative analysis of all samples revealed that almost all lncRNAs were annotated, indicating that lncRNAs may exhibit specificity under certain modifications and tissue conditions. Since the genomic positions of antisense lncRNAs and intronic lncRNAs overlap with PCGs, these

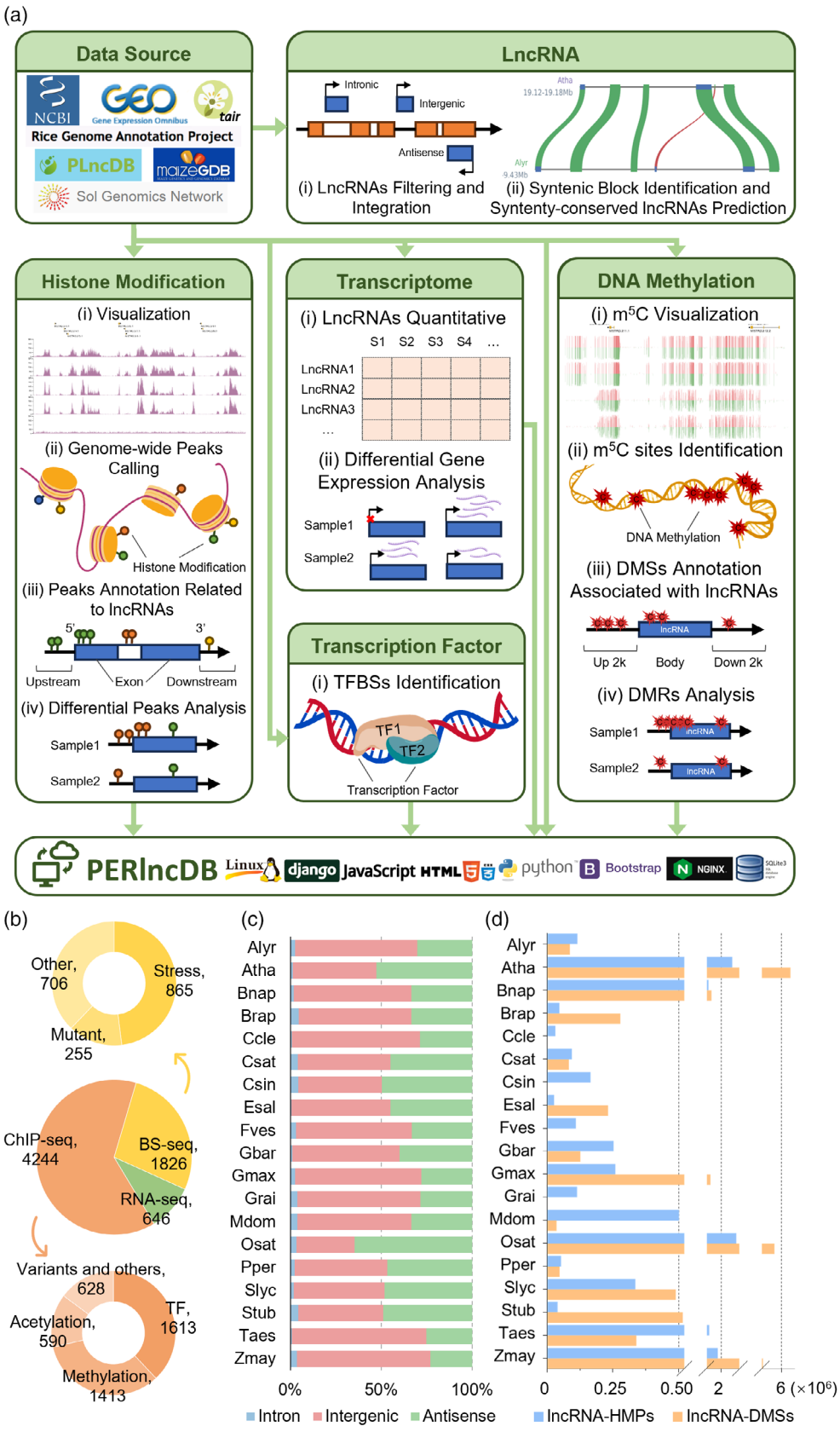


Figure 1. Workflow and key statistics of plant lncRNAs associated with epigenomic features.
(a) Plant lncRNAs and epigenomic datasets' analysis pipeline. Some components of histone modification (ii) chromatin model, transcription factor (i) TFBSs model and DNA methylation (iii) DNA model were derived from GDP (https://BioGDP.com).
(b) Statistics of high-throughput omics data sources (center), including classification of BS-seq samples (top) and ChIP-seq samples (bottom).
(c) Statistics of three types of lncRNAs based on their genomic locations in plants.
(d) Statistics of histone modification peaks and DNA methylation sites associated with lncRNAs. lncRNA-HMPs, histone modification peaks associated with lncRNAs; lncRNA-DMSs, DNA methylation sites associated with lncRNAs.
Alyr, *Arabidopsis lyrata*; Atha, *Arabidopsis thaliana*; Bnap, *Brassica napus*; Brap, *Brassica rapa*; Ccle, *Citrus clementina*; Csat, *Cucumis sativus*; Csin, *Citrus sinensis*; Esal, *Eutrema salsugineum*; Fves, *Fragaria vesca*; Gbar, *Gossypium barbadense*; Gmax, *Glycine max*; Grai, *Gossypium raimondii*; Mdom, *Malus domestica*; Osat, *Oryza sativa*; Pper, *Prunus persica*; Slyc, *Solanum lycopersicum*; Stub, *Solanum tuberosum*; Taes, *Triticum aestivum*; Zmay, *Zea mays*.

two types of lncRNAs may co-express with PCGs due to shared histone marking regulations. To minimize signal interference from PCG regions, we categorized the extracted lincRNAs into: (i) PCG-proximal lincRNAs (PCG-lincRNAs) and (ii) distal intergenic lincRNAs (DS-lincRNAs), based on their distance from adjacent genes (see Methods section for details). We identified 13,846 DS-lincRNAs across representative crop species, such as *A. thaliana*, rice, cucumber, soybean, cotton, tomato, and maize, which accounted for 50.0–70.8% of the total lincRNAs. Then, we compared the distribution of various histone marks in lncRNA and PCG regions across different tissues. As shown in Figure 2c, the percentage of DS-lincRNAs enriched with a single modification type (44.3–65%) was significantly higher than that of PCGs (21.6–35.9%), indicating that lncRNAs exhibit greater specificity in histone mark occupancy. Furthermore, when exploring lncRNAs enriched with the same modification type across different tissues and growth stages, we observed that lncRNAs marked by repressive histone modifications (H3K4me1, H3K27me3, and H3K9me2) tended to display stronger tissue specificity compared to those occupied by active marks (H3K4me3 and H3K36me3) (Figure 2d). This finding suggests that repressive histone modifications may play an important role in the tissue-specific expression of lncRNAs.

Although the above analysis demonstrates that various histone modification sites are widely distributed across lncRNA loci in all wild-type samples, we aimed to explore the factors influencing these modifications at lncRNA loci. We investigated 135 mutant- and 84 stress-responsive samples from *A. thaliana*, which had the most extensive sample collection, and quantified the number of lncRNA loci with significantly altered modification levels in these samples. We found that almost all lncRNAs and PCGs were annotated in at least one stress or mutant sample (Figure 2e). However, when analyzing the frequency of lncRNAs and PCGs annotated to differential histone modification regions and correlating this with sample size, we found that lncRNAs were more likely to appear in the altered signal regions under fewer mutant or stress conditions compared to PCGs (Figure 2f,g). This suggests that lncRNAs may be more selectively regulated by the histone modifiers than PCGs.

DNA methylation patterns and the role of transposable elements in lncRNA

In addition to histone modifications, DNA methylation is another common epigenetic modification signal. In plants, DNA methylation primarily occurs at three types of sites: the symmetric CG and CHG sites and the asymmetric CHH sites (where H represents C, T, or A) (Law & Jacobsen, 2010). Figure S3 summarizes the average DNA methylation levels in leaf tissues across 15 plant species. In most species, lncRNAs exhibited higher methylation levels in the gene body region compared to the flanking regions (Figure S3). The whole content revealed that the methylation patterns of lncRNAs are similar to those of PCGs (Zhao et al., 2021), with higher levels of methylation at CG and CHG sites, while methylation at CHH sites remains relatively low. Additionally, CG methylation was 3′-biased in PCGs, whereas lncRNAs maintained relatively uniform CG methylation. Although lncRNAs share core methylation mechanisms with PCGs, they possess distinct epigenetic regulatory features. Transposable elements (TEs) can move within an organism's genome through transposition or retro-transposition, and DNA methylation is considered a heritable epigenetic modification that silences TEs (Law & Jacobsen, 2010; Lei et al., 2015; Zhang et al., 2018). It also plays a role in maintaining genome stability and regulating gene expression. TEs can induce chromatin modifications near genes, thereby influencing their expression under specific conditions (Hirsch & Springer, 2017). Given the high frequency of genomic overlap between lncRNAs and TEs (Ariel & Manavella, 2021), we categorized lncRNAs into TE-related and non-TE lncRNAs and examined their DNA methylation levels in three species (Arabidopsis, tomato, and maize) that differ in genome size and TE content. Our statistical analysis revealed that 25–33% of lncRNAs are TE-related. Notably, lncRNAs exhibited significantly higher methylation levels compared to PCGs across all three species, as shown in Figure 3a, with TE-related lncRNAs showing the highest methylation levels in all methylation contexts. This may be associated with transcriptional silencing of these elements.

Based on the expression levels of lincRNAs and PCGs in leaves, we categorized them into three groups: high expression, low expression, and non-expression. By examining the

distribution of DNA methylation levels across the whole genome, it was found that non-expressed lncRNAs exhibit higher methylation levels in the gene body region compared to expressed lncRNAs (Figure 3b). Due to the low proportion of expressed lncRNAs (39.1%) in Arabidopsis, the distribution of methylation levels was not smooth.

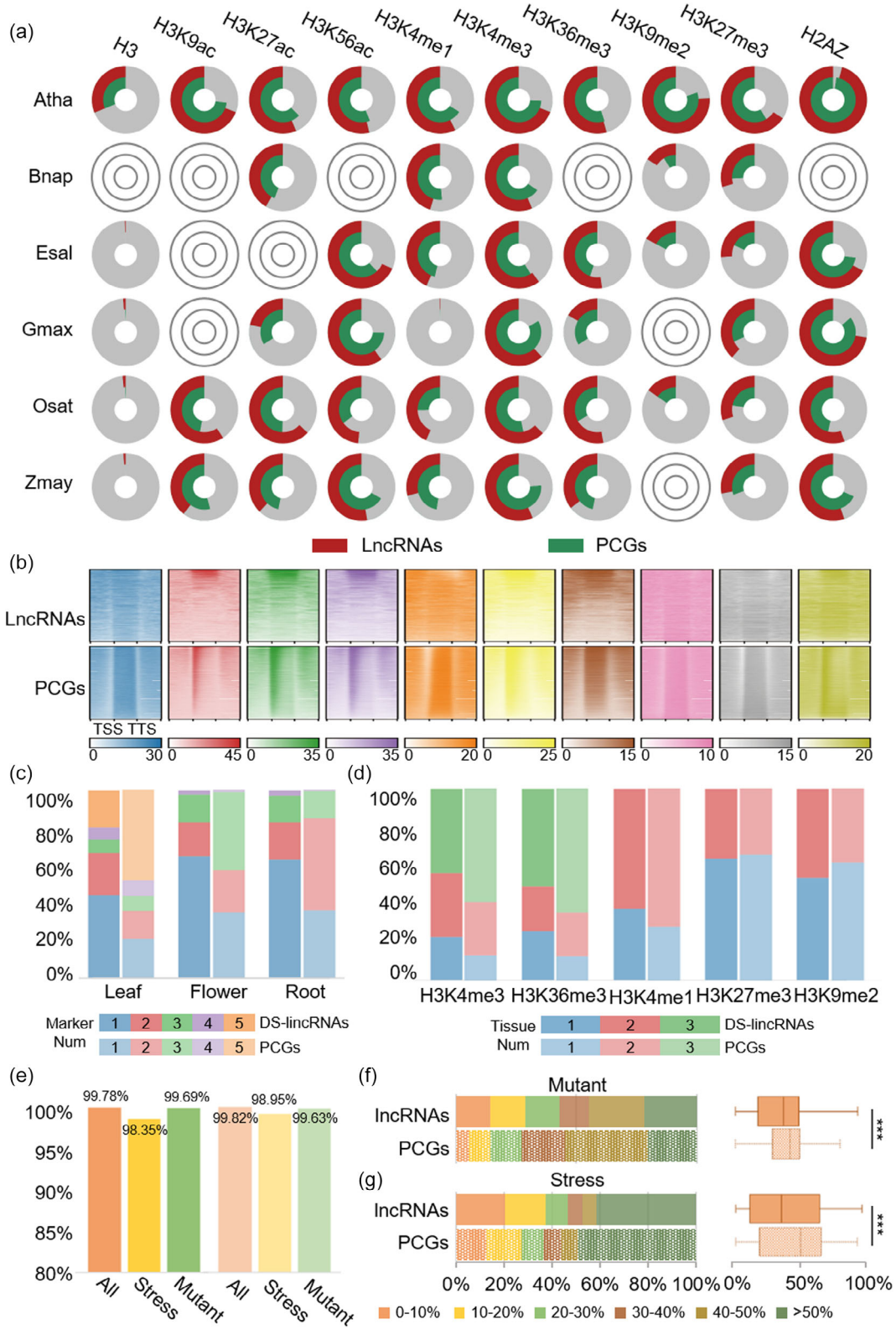


Figure 2. Statistics of histone modifications on lncRNAs and protein-coding genes (PCGs).
(a) Proportions of lncRNAs and PCGs occupied by 10 common histone modification marks across six plant species.
(b) Distribution of 10 common histone modification peaks on lncRNAs and PCGs' region in Arabidopsis leaf samples.
(c) Percentage of distal intergenic lincRNAs (DS-lincRNAs) and PCGs annotated by different histone modification markers in Arabidopsis leaf, flower, and root samples.
(d) Percentage of DS-lincRNAs and PCGs in different tissues occupied by H3K4me3, H3K36me3, H3K4me1, H3K27me3, and H3K9me2, respectively.
(e) Percentage of lncRNAs and PCGs with significantly changed histone modification enrichment in mutant- and stress-related samples.
(f, g) Frequency distribution characteristics of lncRNAs and PCGs with significantly changed histone modification in mutant (f) and stress-related samples (g). Stacked bar plots (left) represent the percentage of two gene types in distinct sample coverage ranges. Boxplots (right) display the comparison of frequencies between gene types based on the Wilcoxon test (***$P < 0.001$).

Additionally, we analyzed classic DNA methylation mutant datasets to investigate their impact on lncRNAs and PCGs. Statistical analysis revealed that the *MET1* deletion mutant (*met1*), which is responsible for maintaining DNA methylation (Zhao et al., 2022), exhibited more significant changes in both lncRNAs and PCGs compared to mutants lacking other methylation-maintaining genes (Figure 3c). This finding further supports the idea that CG methylation at lncRNA loci, like at PCG loci, is generally higher than other types of methylation in plants. Moreover, RNA-dependent RNA polymerase 2 (RDR2) is an essential component of the RNA-directed DNA methylation (RdDM) silencing pathway (Du, 2024). In *rdr2* mutants, the proportion of methylation-associated lncRNAs showed the most significant differences in rice and Arabidopsis, indicating species-specific variations for lncRNA loci (Figure 3c).

Comprehensive functional features of PERlncDB

To facilitate the presentation of our analysis data and provide a user-friendly visual query interface for researchers, we developed the PERlncDB database website (http://perlncdb.liu-lab.com/). This platform offers detailed information and visualizations for lncRNA-related TEs, TFs, histone modifications, and DNA methylation across 19 representative plant species (Figure 4a). Additionally, it provides insights into the differential epigenetic landscapes of lncRNAs under various mutant or stress conditions. The platform comprises key modules, including "Browse," "Search," "Cross-species Analysis," "Dynamics," "Visualization," and "Tools." On the homepage, a navigation bar links to six modules, with quick search options for Histone Modification, DNA Methylation, and TFs search function. It also features a global lncRNA ID search function and links to detailed pages for each species.

The "Browse" module presents comprehensive genomic information for each species, including reference genome versions, genome sizes, the number of lncRNAs, and data sources (Figure 4b). It also provides a browsing interface for species-specific lncRNAs, displaying basic details along with annotations for TEs, TFs, histone modifications, and DNA methylation. By clicking on individual lncRNA IDs, users can access detailed pages that include sections such as Basic Information, Transcripts, TEs, TFs, Histone Modifications, and DNA Methylation. Utilizing the detailed genomic and epigenomic information for each lncRNA, the "Search" module offers several search options to retrieve specific lncRNA data, thereby enhancing the platform's accessibility and functionality. The module includes four search options: (i) direct search by lncRNA ID, (ii) search by TFs, (iii) search by DNA methylation-associated genomic regions, and (iv) search by histone modification-associated genomic regions. Each search returns detailed information about the corresponding lncRNA (Figure 4c).

The "Cross-species Analysis" module enables the identification and exploration of synteny-conserved lncRNAs, lncRNA synteny scan, and correlation analysis of epigenetic features (Figure 4d). Previous studies have shown that the positional conservation of lncRNAs on the genome is significantly higher than their sequence conservation (Wang, Niu, et al., 2015; Zhou et al., 2022). Therefore, we delineated syntenic blocks across the genome and predicted a total of 80,308 synteny-conserved lncRNAs (Figure S4). Among the 19 species, the proportion of syntenically conserved lncRNAs ranged from 20.0% to 67.4%, with the highest proportion observed in *Gossypium raimondii* and relatively fewer conserved lncRNAs in *Cucumis sativus* (Figure S4). Although the proportion of syntenically conserved lncRNAs varies across species from different families, it tends to be more consistent or closely related within the same family. This suggests that lncRNAs exhibit a high degree of positional conservation, especially at the family level. We have integrated an additional layer of analysis by implementing sequence conservation evaluation for each candidate lncRNA pair. This analysis is performed within the "Synteny-conserved LncRNA Prediction" module, enabling users to assess candidates based on their sequence characteristics. In addition to querying syntenically conserved lncRNA pairs across species, this module allows users to automatically calculate the correlation of epigenomic signal levels between lncRNA pairs across different samples, enabling the exploration of epigenetic regulatory mechanisms of lncRNAs across species.

Besides, "Dynamics" module provides differential analysis data for epigenetic modification signals under mutant or stress conditions, with detailed information provided in Figure 4e. Data can be retrieved by selecting the

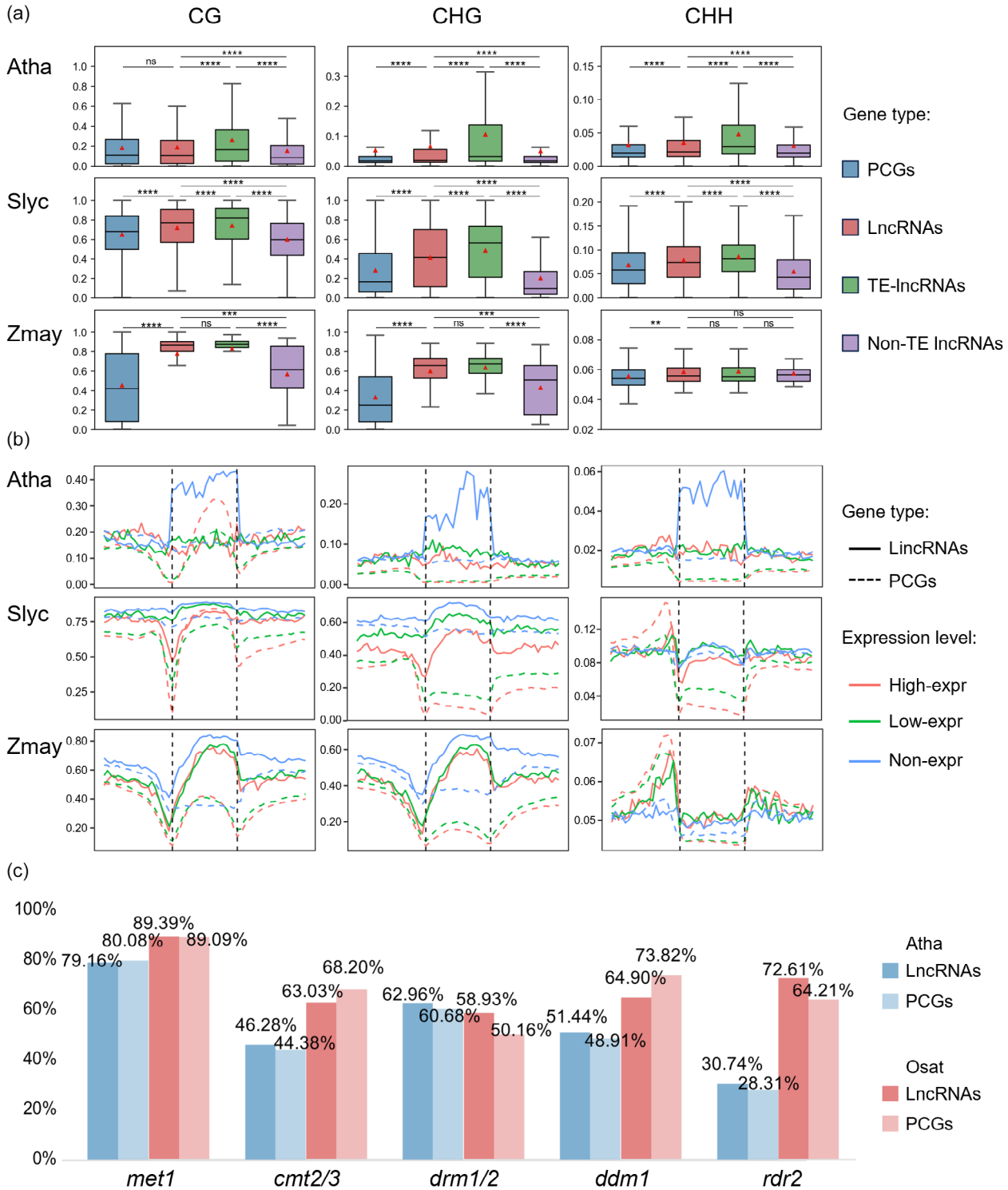


Figure 3. Statistics of DNA methylation level on lncRNAs and PCGs.
(a) DNA methylation levels of four gene types in Arabidopsis, tomato, and maize. The gene types include PCGs (blue), lncRNAs (red), transposable element-associated lncRNA (TE-lncRNAs) (green), and non-TE lncRNAs (purple).
(b) DNA methylation levels of lncRNAs (solid) and PCGs (dashed) grouped by gene expression levels in Arabidopsis, tomato, and maize leaves.
(c) Proportions of lncRNAs (dark) and PCGs (light) with DMRs in five classic DNA methylation-related mutants from Arabidopsis and rice.

species of interest, modification type, and sample details. Moreover, this module offers access to differential epigenetic regions, analysis results, and sample information, facilitating a comprehensive understanding of epigenetic changes.

Additionally, to enhance the visualization and presentation of data resources on the website, the database integrates multiple bioinformatics tools to optimize the data analysis experience. In the "Visualization" module, the platform allows users to explore epigenetic landscapes by selecting the desired species, modification types, and sample tracks (Figure 4f). Moreover, the "Tools" module primarily offers the BLAST alignment and "Auto Annotation" tool. BLAST provides sequence alignment information across multiple species, while "Auto Annotation" delivers detailed information on all differentially epigenetic regions analyzed in this study. Species, samples, tissues, specific lncRNA regions, or other genomic regions of interest can be submitted to examine the modification levels in the targeted areas. The platform also includes a "Download" section where related data and resources can be accessed.

To comprehensively demonstrate the core advantages of PERlncDB, we conducted a systematic comparative analysis with six mainstream plant lncRNA platforms. Through a comprehensive evaluation of 16 metrics across four key dimensions: species coverage, lncRNA annotation scale, epigenetic features, and functional modules. PERlncDB currently provides the most comprehensive epigenetic characterization with 163,609 lncRNA entries spanning 19 plant species (Table 1). It also offers specialized functions: (i) differential epigenetic mark analysis; (ii) transposable element-associated lncRNA (TE-lncRNA) annotation; (iii) synteny-conserved lncRNA analysis; and (iv) auto-annotation tools, features that establish its unparalleled status among existing plant lncRNA resources.

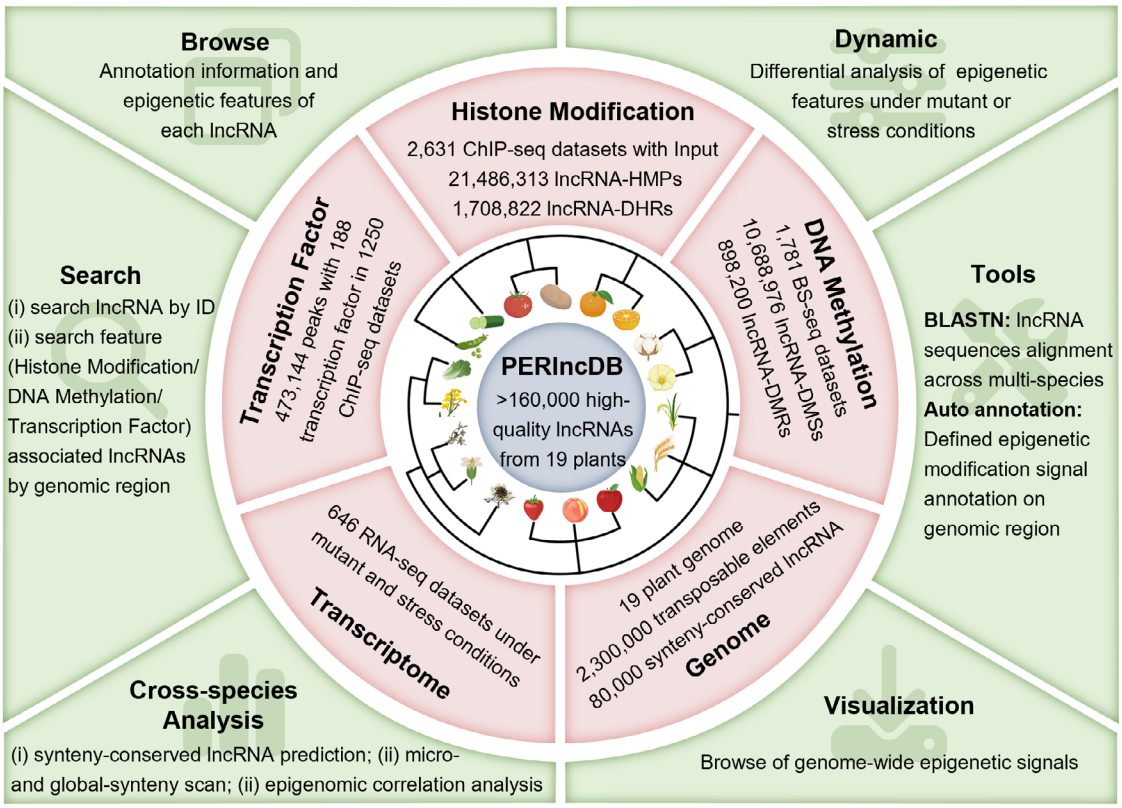
In summary, PERlncDB provides a powerful and intuitive interface for exploring and analyzing lncRNAs, significantly enhancing the research experience and supporting advanced studies in plant epigenomics.

Epigenetic regulation of two reported lncRNAs in PERlncDB

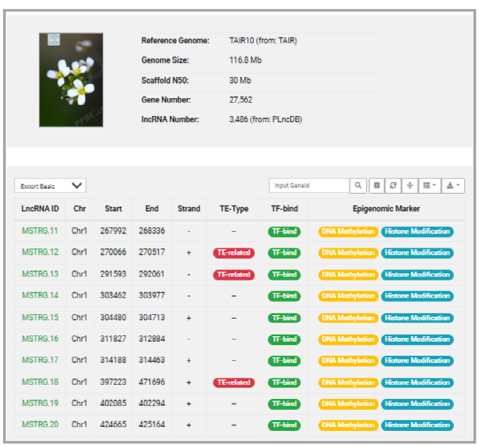
To illustrate the wealth of data available and its accessibility for lncRNA research, we used two previously published lncRNAs as examples to explore their detailed information in our database. In the previous studies, *MARS* was shown to regulate the local epigenetic activation of its surrounding region in response to abscisic acid (ABA) by modulating the dose-dependent binding of LHP1, which subsequently activated H3K27me3 deposition and chromatin condensation (Roule et al., 2022). In the Visualization module, through clicking the "H3K27me3_seedling"-related samples and setting the positional range of *MARS*, we observed enriched LHP1 binding signals at the *MARS* locus and its adjacent regions, along with the distribution of H3K27me3 in both wild-type and *lhp1* mutant samples (Figure 5a), which are consistent with the previous report (Roule et al., 2022; Veluchamy et al., 2016). Additionally, several ABA-related TFs from five gene families collected in PERlncDB showed clear binding signals at the *MARS* locus (Figure S5), replicating the ABA-responsive binding motifs found at the *MARS* locus (Roule et al., 2022; Song et al., 2016). Further conserved synteny-lncRNA prediction for *MARS* in the Brassicaceae family (Figure 5b) implied that the positionally conserved lncRNAs in *Arabidopsis lyrate* (*A. lyrate*) may function similar to *MARS*. The analysis of the correlation between epigenetic modification signals at *MARS* and its syntenic conserved lncRNA in *A. lyrate* revealed a significant correlation in H3K27me3 modification level at both upstream and downstream regions in wild-type samples (Figure 5c). Additionally, we sought to determine whether *MARS* transcription is influenced by altered epigenetic modifications. By mining the comprehensive data resources in PERlncDB, we observed that the CG methylation levels at the *MARS* locus were significantly reduced in *ddm1* mutants compared to wild-type plants, accompanied by the upregulation of *MARS* transcription (Figure 5d, Figure S6a). To further investigate the potential link between DNA methylation regulation and ABA stress response, we integrated and analyzed two independent transcriptome datasets under ABA treatment (PRJNA274888 and PRJNA389285). Differential expression analysis revealed that *DDM1* expression was significantly downregulated upon ABA treatment (Figure 5e). Based on these findings, we hypothesize that ABA stress signals may suppress *DDM1* expression, leading to reduced DNA methylation at the *MARS* locus and consequently activating its expression. This discovery uncovers the complexity of epigenetic regulation in ABA signaling and suggests a potential cooperative mechanism between DNA methylation and H3K27me3 in modulating MARS-mediated ABA responses.

The other lncRNA, *LINC-AP2*, can suppress the expression of its downstream target gene AP2 under TCV virus infection, leading to abnormalities in floral organs (Gao et al., 2016). Due to its similar phenotype to certain Arabidopsis mutants of RNA-dependent RNA polymerase 6 (*rdr6*), gene silencing suppressor 3 (*sgs3*), and Dicer-like 4 (*dcl4*), it was hypothesized that *LINC-AP2* might be associated with DNA methylation (Gao et al., 2016). To explore this further, we examined the expression patterns of *LINC-AP2* in these three mutants and observed significant alterations in CHH methylation levels (Figure 5f; Figure S6), suggesting that *RDR6*, *SGS3*, and *DCL4* are likely regulated the methylation modification at *LINC-AP2* locus. Furthermore, our analysis of the PERlncDB revealed a significant enrichment of the transcription factor ARF3/ETT at the *LINC-AP2* locus (Figure 5g). Given that *ett* loss-of-function mutants have been previously reported to exhibit severe

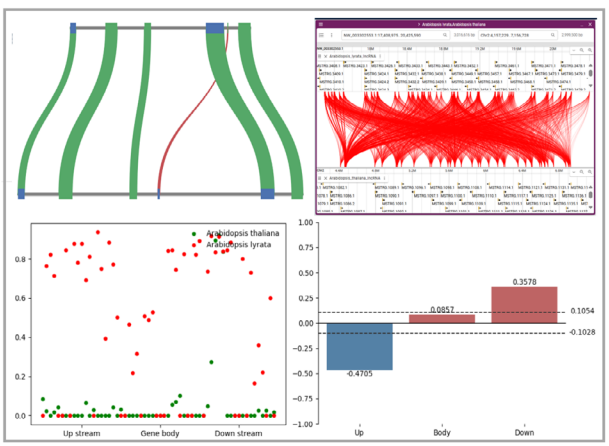
(a)



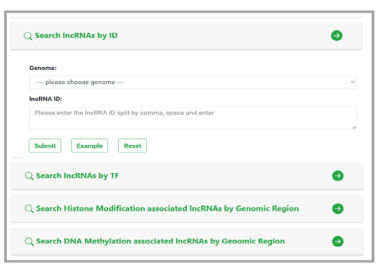
(b)



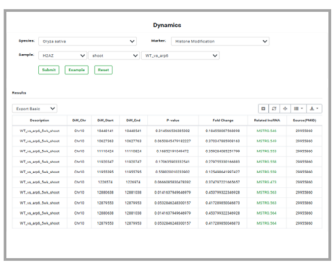
(d)



(c)



(e)



(f)

Figure 4. Overview and functional demonstrations of PERlncDB modules and features.
(a) Details of functional modules and data resources in PERlncDB. lncRNA-DHRs, differential histone modification regions associated with lncRNAs; lncRNA-DMRs, differential DNA methylation regions with lncRNAs. The plant icons were derived from GDP (https://BioGDP.com).
(b) "Browse" module—an example of Arabidopsis lncRNAs browse, including genomic location, epigenomic future and detailed information with link.
(c) Demonstration of four items of "Search" functions, including "Search lncRNAs by ID," "Search Transcription Factor and associated lncRNAs by Genomic Region," "Search Histone Modification and associated lncRNAs by Genomic Region," and "Search DNA Methylation and associated lncRNAs by Genomic Region."
(d) "Cross-Species Analysis" module—an example using lncRNA pairs from Arabidopsis and *Arabidopsis lyrate*, featuring synteny scan and visualization of epigenetic modification level correlation.
(e) "Dynamics" module demonstration—an example using Arabidopsis lncRNAs detailed changes in histone modification enrichment.
(f) "Visualization" module—an example using Arabidopsis with omics data tracks, including gene annotation, histone modification enrichment, DNA methylation sites, TF binding sites, and gene expression level.

Table 1 Comparison of PERlncDB with other published plant lncRNA databases

Data item	PERlncDB	PLncDB V2.0	GREENC v2	CANTATAdb 2.0	PNRD	EVLncRNAs 3.0	NONCODEV6
Plant species	19 species	80 species	94 species	39 species	150 species	64 species	23 species
ChIP-seq libraries	4244	456	NA	NA	34	NA	NA
BS-seq libraries	1826	NA	NA	NA	NA	NA	NA
RNA-seq libraries	645	13 836	NA	328	NA	NA	38 (plant)
LncRNA entries	163 609	1 246 372	495 000	239 631	5573	578	94 697 (plant)
Experimentally identified lncRNAs	√	√	NA	NA	Partial	√	NA
Genomic annotation	√	√	√	√	√	√	√
Expression	√	√	NA	√	NA	√	√
Histone modification peaks annotation	√	√	NA	NA	NA	NA	NA
DNA methylation annotation	√	√	NA	NA	NA	NA	NA
Transcription factor binding sites annotation	√	NA	NA	NA	NA	NA	NA
TE-lncRNAs	√	NA	√	NA	NA	NA	NA
Differential epigenetic marks analysis	√	NA	NA	NA	NA	NA	NA
Auto-annotation tools	√	NA	NA	NA	NA	NA	NA
Synteny-conserved lncRNA analysis	√	NA	NA	NA	NA	NA	NA
Accessible	√	√	√	√	√	√	√
Reference (PMID)	This study	33 079 992	34 723 326	30 945 201	25 398 903	37 953 349	33 196 801

The statistics in NONCODEv6 are mainly focus on plant species. Plant species: Total number of plant species covered. Library scale: Number of integrated datasets (ChIP-seq for histone modifications/TF binding; BS-seq for DNA methylation; RNA-seq for lncRNA expression). LncRNA entries: Total number of cataloged lncRNAs. Experimentally identified: LncRNAs functionally validated by experiments. Genomic annotation: Basic genomic information for lncRNAs. Expression profile: Quantitative expression data. Epigenetic annotation: Annotations for histone modification peaks, DNA methylation levels (CG/CHG/CHH), and transcription factor binding sites at lncRNA loci. TE-lncRNAs: LncRNAs derived from transposable elements. Differential epigenetic marks analysis: Pre-computed results for comparing epigenetic marks across conditions. Auto-annotation tools: Online tools for automated epigenetic annotation of user-submitted lncRNA data. Synteny-conserved lncRNAs: Identification of conserved lncRNAs via genomic synteny. Accessible: Website functionality and data availability during the assessment. Reference (PMID): PubMed ID of the primary publication. Symbols: "√" indicates that the feature is provided; "NA" indicates not available or no evidence found.

floral developmental defects (Simonini et al., 2017), we further performed transcriptomic analysis and found that *LINC-AP2* expression was significantly upregulated in *ett* mutants compared to the wild type (Figure 5g, Figure S6b), indicating that ETT acts as a transcriptional repressor of *LINC-AP2*. Based on these findings, we propose that during floral development, *LINC-AP2* is likely regulated by distinct epigenetic mechanisms: DNA methylation-mediated chromatin state changes and ETT-dependent transcriptional repression. Thus, by leveraging the large-scale datasets in PERlncDB, researchers can investigate the dynamics of lncRNA epigenomic signals to infer their potential functions.

Identification of lncRNA pairs with conserved epigenetic regulatory mechanisms across species

The poor sequence conservation and diverse functional roles of lncRNAs present significant challenges in studying

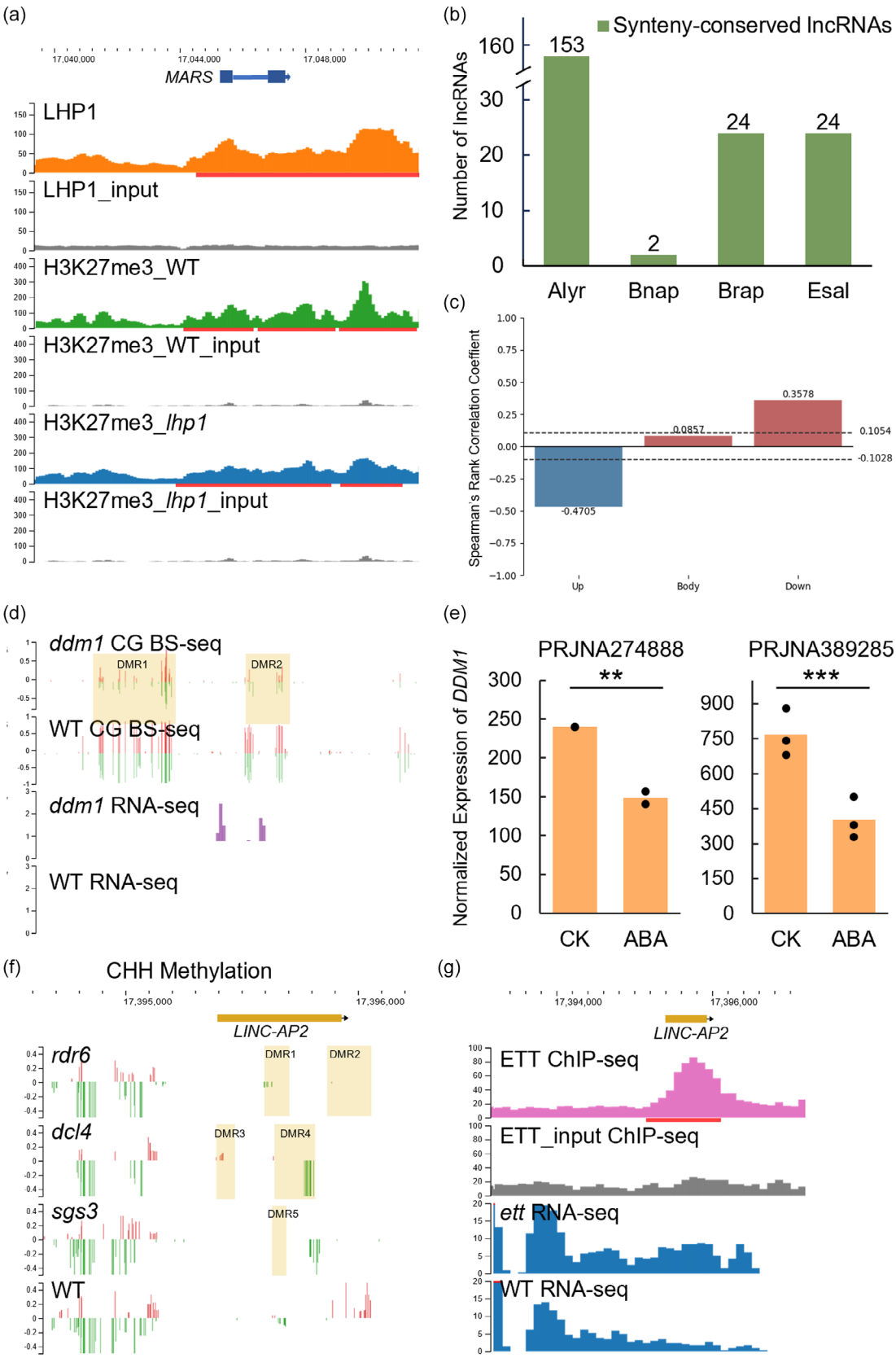


Figure 5. Epigenetic regulation information for two reported lncRNAs in PERlncDB.
(a) Occupancy of LHP1 and the H3K27me3 modification levels at the *MARS* locus. The red line under each track represents the statistically significant peaks.
(b) Statistical analysis of the lncRNAs showing syntenic conservation of *MARS* across four Brassicaceae species.
(c) The strongest correlation of epigenetic modification signal between *MARS* and its synteny-conserved lncRNA in *Arabidopsis lyrata*. Red indicates the degree of positive correlation, while blue represents negative correlation. The dashed line marks the correlation threshold corresponding to a significance level of 0.05. Correlation values exceeding the absolute value of this threshold suggest a more reliable level of correlation in the epigenetic modifications of the two lncRNAs in the given sample.
(d) CG DNA methylation and expression pattern of *MARS* in Arabidopsis seedlings across *ddm1* mutant and wild-type sample. Compared with wild type, the DNA methylation levels in DMRs reduce by 16.5% for DMR1 and 12.7% for DMR2 in *ddm1* mutant.
(e) Normalized expression levels of *DDM1* under ABA treatment versus control conditions in two public projects.
(f) CHH DNA methylation levels at the *LINC-AP2* locus and its adjacent regions in three different mutants, including *rdr6*, *dcl4*, and *sgs3*. Compared with wild type, the DNA methylation levels in DMRs reduce by 3.7% for DMR1 and 17.4% for DMR2 in *rdr2* mutant, 7.9% for DMR3 and 7.4% for DMR4 in *dcl4* mutant, and 8.5% for DMR5 in *sgs3* mutant. Red lines represent methylation sites on the sense strand, and green lines on the antisense strand.
(g) ETT binding characteristics at the *LINC-AP2* locus and expression analysis of *LINC-AP2* in *ett* mutant and wild-type sample.

their functions, especially across species. A recent study revealed that, despite the lack of sequence homology between human *UPAT* and Arabidopsis *APOLO*, *UPAT* can exert similar regulatory functions as *APOLO* in plants, highlighting the conservation of lncRNAs across species and the similarity of their epigenetic mechanisms between plants and animals (Fonouni-Farde et al., 2022). In this study, we sought to explore how to identify lncRNA pairs with conserved epigenetic regulatory mechanisms using PERlncDB. Excluding sequence factors, we focused on synteny-conserved lncRNA pairs from different species and investigated whether they could be regulated by homology proteins. Ultimately, our research centered on lncRNAs from Arabidopsis and tomato.

In both tomato and Arabidopsis, studies on MET1 highlight its functional conservation, as it affects gene expression by increasing DNA methylation levels (He et al., 2022; Mathieu et al., 2007; Yang et al., 2019). By retrieving differentially methylated regions (DMRs) of *met1*, 6127 and 55,471 lncRNAs were annotated in *A. thaliana* and *Solanum lycopersicum* (*S. lycopersicum*), respectively. In the Arabidopsis *met1* sample, 423 lncRNAs showed increased methylation levels, while 950 lncRNAs exhibited decreased methylation levels (Figure 6a). In the tomato *met1* sample, 5322 lncRNAs displayed increased methylation levels, whereas 4793 showed decreased levels (Figure 6b). Quantitative and differential expression analyses of lncRNAs revealed that 32 and 94 lncRNAs were highly expressed in the *met1* mutants of *A. thaliana* and *S. lycopersicum*, respectively (Figure 6a,b). Consequently, we identified 31 and 73 lncRNAs in *A. thaliana* and *S. lycopersicum*, respectively, as potential targets regulated by MET1.

To further investigate the positional conservation of lncRNAs between the two species, we identified that 11 lncRNAs in *A. thaliana* and 7 in *S. lycopersicum* were positionally conserved using the "LncRNA Microsynteny Scan" module. Among them, *MSTRG.2210.1* from Arabidopsis and *MSTRG.2095.1* from tomato caught our attention. These two lncRNAs exhibited syntenic conservation (Figure 6c), yet their sequences were highly divergent (Figure 6d). Sequence feature and secondary structure predictions revealed that their local structures are similar (Figure 6e, Figure S7). The DMRs were identified at both the *MSTRG.2210.1* and *MSTRG.2095.1* loci, showing a significant decrease in DNA methylation levels in the *met1* mutant compared to the wild type, with a reduction of 63.7% (DMR1) and 42.3% (DMR2), respectively (Figure 6f). This consistent dynamics of DNA methylation levels in *met1* mutant of both *A. thaliana* and *S. lycopersicum* represented that this lncRNA pair was influenced by MET1. Through further synteny analysis, we identified 16 lncRNAs across three families (Rosaceae, Juglandaceae, Malvaceae) that exhibit syntenic relationships with both Arabidopsis *MSTRG.2210.1* and tomato *MSTRG.2095.1* (Figure 6g), which implies these two lncRNAs retain a stable position relationship. A previous study has demonstrated that PCGs exhibiting body methylation changes evolve more slowly and often possess crucial biological functions (Takuno & Gaut, 2012). For lncRNAs, beyond sequence conservation, the remarkable synteny conservation associated with epigenetic regulator may provide an excellent perspective for investigating functional conservation across species.

FUTURE PERSPECTIVES

Considering the strong tissue specificity of lncRNAs and the variability in quantity and quality of RNA-seq datasets used for their identification, lncRNAs identified in the same species may exhibit significant discrepancies across different studies (Chen & Zhu, 2022; Palos et al., 2022; Waseem et al., 2020). It is essential to consider how our data platform can provide extended epigenetic information for lncRNAs that have not yet been cataloged, as the current databases (PLncDB, CANTATAdb, and GreeNC) typically offer fixed epigenomic information for a predefined set of lncRNAs without accounting for this limitation. To address this issue, we have introduced an "Auto Annotation" tool that allows researchers to explore modification signal characteristics of lncRNAs that are not collected in our PERlncDB platform, helping to bridge data gaps. Moreover,

compared to previous study, lncRNA loci exhibit greater histone modification specificity and tissue specificity (Figure 2c,d) (Chow et al., 2022), suggesting that the abundant epigenomic data resources collected in PERlncDB are highly valuable for precisely capturing the epigenetic modification states of lncRNAs under specific

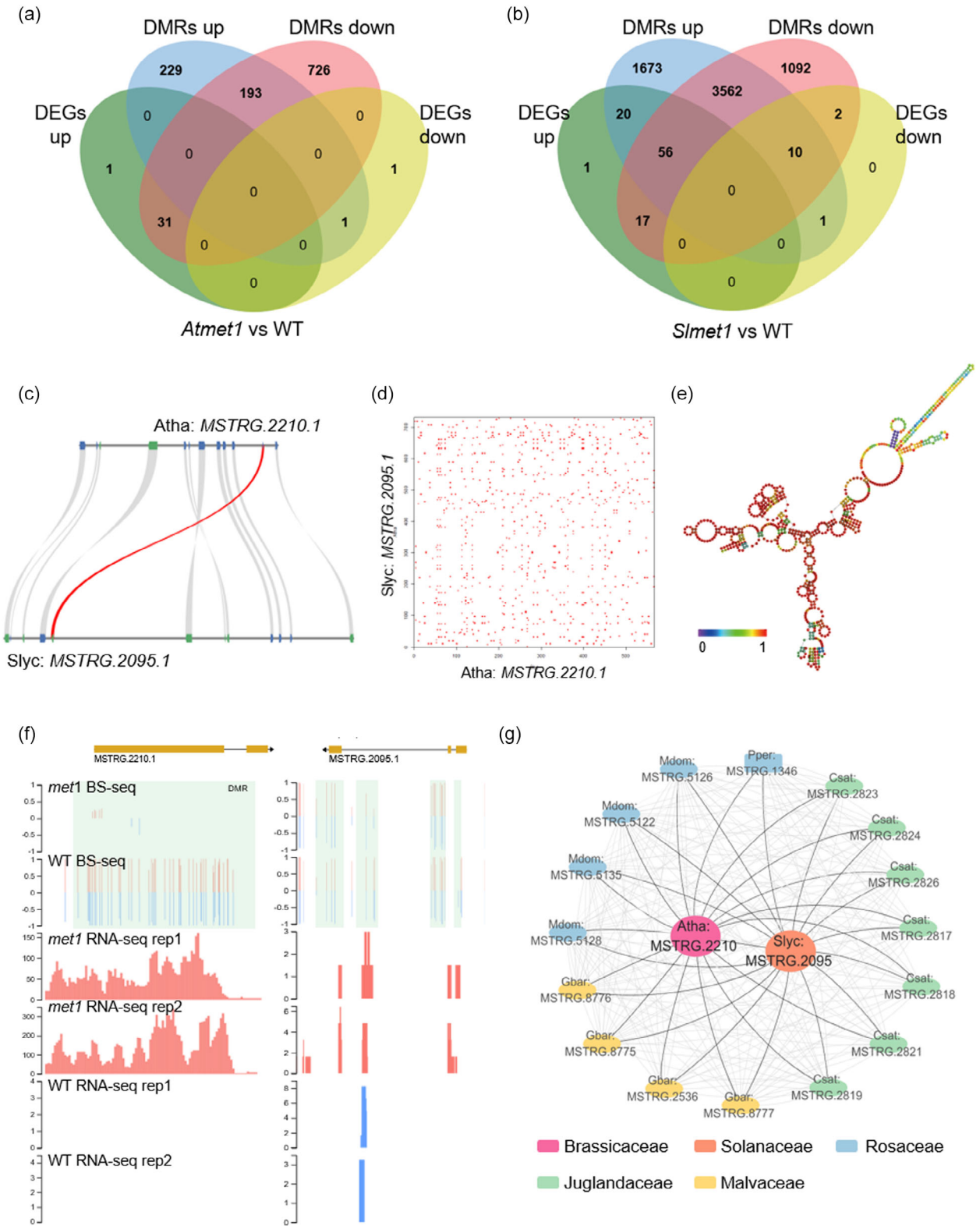


Figure 6. A key study on predicting conserved epigenetic regulatory mechanisms of lncRNAs.
(a, b) Statistical analysis of lncRNA expression regulated by *met1*. Venn diagrams showing lncRNAs with significantly changed altered expression and DNA methylation levels in Arabidopsis (a) and tomato (b).
(c) Local synteny block of Atha *MSTRG.2210.1* and Slyc *MSTRG.2095.1*.
(d) Sequence similarity alignment between Atha *MSTRG.2210.1* and Slyc *MSTRG.2095.1*.
(e) Consensus secondary structure prediction for Atha *MSTRG.2210.1* and Slyc *MSTRG.2095.1*. The color bar indicates the reliability of base pairing, with red representing highly conserved secondary structure.
(f) DNA methylation and expression levels at the loci of Atha *MSTRG.2210.1* and Slyc *MSTRG.2095.1* in the *met1* mutant and the wild type.
(g) Conserved lncRNA network based on synteny analysis in five families. Nodes represent lncRNAs, and edges indicate syntenically conserved lncRNA pairs, with thick edges indicating conservation with both Atha *MSTRG.2210.1* and Slyc *MSTRG.2095.1*.

conditions. Such insights facilitate a better understanding of lncRNA transcription and provide deeper insights into their functions. Furthermore, tools like "Auto Annotation," which can automatically retrieve functional annotations for lncRNAs, remain largely underdeveloped. This indicates that future data platforms for functional annotation of lncRNAs should place greater emphasis on integrating such tools, rather than solely focusing on updating the information of cataloged lncRNAs.

With the rapid development of high-throughput sequencing technologies, the exploration of lncRNA regulatory functions and the improvement of data platforms remain expansive research frontiers. Epigenetic modifications, such as DNA methylation, histone modifications, RNA modifications, open chromatin regions, and 3D genome structures, represent diverse and critical resources for understanding gene regulation (He et al., 2021; Liang et al., 2020; Zhang & Zhu, 2025). Currently, our research mainly focuses on the two most extensively studied epigenetic features—DNA methylation and histone modifications. In plants, several lncRNAs with well-defined functions, such as *APOLO* and *LAIR*, have been shown to interact with multiple epigenetic modifications in their regulatory roles (Ariel et al., 2020; Wang et al., 2018). Although PERlncDB has successfully integrated and facilitated the retrieval of lncRNA-related epigenetic data, the comprehensive crosstalk analysis of multiple modifications at lncRNA loci still requires improvement. Therefore, in the future, we plan to expand our research scope to comprehensively integrate and utilize various epigenetic data resources, enabling a more comprehensive understanding of the regulatory mechanisms that govern lncRNAs.

In this study, we investigated the conserved mechanisms of a synteny-conserved lncRNA pair in Arabidopsis and tomato using the data resources of PERlncDB (Figure 6). Future efforts will focus on expanding the platform by incorporating more bioinformatics tools and diverse data resources to better explore the epigenetic features of conserved lncRNAs across species. We also plan to include epigenetic variant sites and lncRNAs associated with specific traits at the population level, thereby enriching the data support available to researchers. Furthermore, single-cell epigenomics has the potential to provide new insights into decoding the gene regulatory landscape (Luo et al., 2020; Zhang et al., 2024).

In conclusion, exploring lncRNA functions and optimizing data platforms is a long-term and challenging endeavor. We believe that the continuous integration of diverse data resources will provide valuable support for advancing our understanding of lncRNA epigenetic regulation in future research.

MATERIALS AND METHODS

Data acquisition and classification

The genome sequences and gene annotations were downloaded from public databases (Table S2). Chromatin immunoprecipitation sequencing (ChIP-seq) and bisulfite sequencing (BS-seq) samples and projects were retrieved from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) and/or Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo), ensuring that each dataset was supported by reliable, published literature. By manually reviewing the literature, we labeled each dataset with sample information, tissue type, study factors (e.g., TFs or histone modification marks), biological replicates, and other relevant details. To facilitate the study of regulatory characteristics of interest (e.g., TFs, histone modifications, and DNA methylation) under specific tissues or conditions, we processed each experiment independently.

The collection of epigenome data (including ChIP-seq and BS-seq datasets) was primarily conducted through two approaches: manual collection of raw data by reviewing the literature and keyword-based searches of public databases such as NCBI and GEO. The raw sequencing files (in FASTQ format) for each sample were downloaded from the European Nucleotide Archive (ENA, https://www.ebi.ac.uk/ena). To ensure data reliability, ChIP-seq datasets were selected from experiments containing "Input" samples.

Integration and classification of plant lncRNAs

Plant lncRNA data and annotation information were mainly obtained from public databases, including PLncDB (Jin et al., 2021), EVLncRNAs (Zhou et al., 2019), RNAcentral (Consortium, 2021), and RefSeq (https://www.ncbi.nlm.nih.gov/refseq/). These lncRNAs from different sources were mapped to the same genome version using BLASTN v2.16.0 (with parameters "-evalue 1e-5 -outfmt 6 -max_target_seqs 1") and then merged with StringTie v2.2.1 (Shumate et al., 2022). Next, we employed the GffCompare v0.11.2 (Pertea & Pertea, 2020) tool to identify lncRNA types according to the protein-coding gene positions. Finally, lncRNA sequences were extracted using Gffread v0.12.7 (Pertea & Pertea, 2020), and transcripts with low protein-coding potential were

retained based on CPC2 predictions (Kang et al., 2017). In summary, we selected high-quality lncRNA transcripts that conform to the following criteria: (i) class code labels of i, x, or u; (ii) classified as non-coding; (iii) transcript length ≥200 bp. Based on genomic location, lncRNAs were categorized into three types: (i) intergenic lncRNAs, transcribed from DNA sequences between two protein-coding genes; (ii) intronic lncRNAs, originating from introns within protein-coding genes; and (iii) antisense lncRNAs, transcribed in the opposite direction to protein-coding genes.

To minimize interference at coding gene sites, intergenic lncRNAs were further classified based on their distances from adjacent genes into PCG-lincRNAs and DS-lincRNAs. For lncRNAs with a median distance of >5 kb from neighboring genes, those beyond the median were classified as DS-lincRNAs, while those within 5 kb were classified as PCG-lincRNAs. If the median distance was <5 kb, transcripts were categorized based on whether they were farther or closer than the median distance.

TE-related lncRNA identification

Species-specific transposable element information was downloaded from public platforms such as APTEdb (Pedro et al., 2021) and RepetDB (Amselem et al., 2019). TEs were aligned to the reference genome of each species using BLAST, and the best alignment sequences were extracted. We used the BEDtools v2.30.0 (Quinlan & Hall, 2010) "intersect" method to identify lncRNAs with overlapping regions of more than 10 bp with TEs as TE-related lncRNAs, while lncRNAs with no overlapping regions or overlapping regions of less than 10 bp were identified as non-TE related lncRNAs.

Histone modification data analysis

Enrichment regions of histone modifications are identified using the MACS2 v2.2.7.1 tool (Zhang et al., 2008). For broad peak histone modifications (H3K36me3, H3K20me1, H3K4me1, H3K79me2, H3K79me3, H3K27me3, H3K9me3, and H3K9me1), the parameter "--broad" is set. For narrow peaks, the parameter is set to "--call-summits," and a lenient *P*-value threshold ($P < 1e\text{-}2$) is set to correctly calculate the irreproducible discovery rate. Since IDR v2.0.4.2 requires input peak data spanning the entire spectrum of high confidence (signal) and low confidence (noise) to fit a bivariate model that separates signal from noise, we follow the recommendations for assessing consistency and reproducibility between replicates (Li et al., 2011). Peak calling is performed for biological replicates in three groups: all replicates, merged data for each replicate (combined replicates), pseudo-replicates (each sample splits into two subsamples), and pseudo-replicates of merged samples (combined replicate samples split into two subsamples), using the same merged control as input. For samples without replicates, peaks with "-log10 *q*value > 2" and "fold enrichment > 2" are extracted as the final results. Overlap, merging, and summarization of peaks were analyzed using BEDTools. Genome-wide localization and annotation were performed using ChIPseeker v1.30.1 (Yu et al., 2015) and BEDTools, with detailed genetic regions categorized as follows: upstream, 5′UTR, CDS, intron, 3′UTR, downstream, and intergenic regions for protein-coding regions and upstream 1 k, intron, exon, downstream, and intergenic regions for non-coding regions.

ChIP-seq data were analyzed for statistical significance using DiffBind v3.4.11 with the DESeq2 method, identifying differential peaks by applying thresholds for fold change and *P*-value. Significant differential peaks were further analyzed, including genome-wide annotation and localization of differential regions, lncRNA annotation, and identification of TF binding sites (TFBS).

DNA m5C data analysis

Raw sequencing data used Trim Galore v0.4.1 to remove potential adapter with the parameters "-q 20 --phred33 --stringency 3 --length 20 -e 0.1" and FastQC (Brown et al., 2017) to assess the quality. The clean reads were mapped to the reference genome of each germplasm using the BSseeker2 v2.1.8 with the following parameters "--aligner=bowtie2 --XSteve --bt2-p 4 --bt2--end-to-end" (Guo et al., 2013). Redundant reads and PCR duplicates were removed using Picard v2.25.5 and SAMtools v1.9 (Li et al., 2009). CGmapTools v0.1.2 (Guo et al., 2018) was used to convert BAM files to CGmap files and methylation sites were identified with BSseeker2.

Differentially methylated regions (DMRs) refer to genomic regions exhibiting different methylation patterns between samples. This study identifies genomic regions showing significant methylation level changes under different samples or conditions using the CGmapTools "dmr" strategy, including DMR detection, localization, and calculation of average methylation levels. LncRNA annotation of DMRs is performed using bedtools and ChIPseeker, and visualization is accomplished using deepTools v3.5.1 (Ramirez et al., 2014).

Transcriptome data analysis

RNA-seq data primarily focused on mutants related to DNA methylation and histone modifications as well as stress-related samples. Raw data were downloaded from NCBI, and raw reads were filtered using Trim Galore, with sequencing quality assessed by FastQC. Mapping to the genome and gene quantification followed published pipelines (Kim et al., 2015). Differential gene expression between tissues was analyzed using DESeq2 in the R environment, with "$|\log 2(\text{fold change})| \geq 1$" and "*P*adj < 0.05."

Synteny-conserved lncRNA prediction

JCVI v1.4.20 (Tang et al., 2024) was used to construct whole-genome syntenic blocks based on protein-coding gene sequence homology between species. Then, we employ two core criteria to screen for synteny-conserved lncRNAs based on each block. LncRNAs were classified as synteny-conserved when meeting dual criteria: (i) at least five (moderate criterion) adjacent homologous protein-coding genes must be syntenic between the two species and (ii) transcriptional orientation concordance with any neighboring homologous coding gene. To accommodate diverse research requirements, we implemented a tiered classification system incorporating based on the number of adjacent homologous coding genes: (i) a lenient threshold (≥3 genes) optimized for maximal sensitivity in lncRNA detection, particularly useful for exploratory analyses; (ii) a moderate threshold (≥5 genes) serving as our primary classification standard that balances sensitivity and specificity; and (iii) a strict threshold (≥10 genes and transcriptional orientation are consistent with the nearest homologous coding gene upstream or downstream) designed for high-confidence conservation assessment when studying deeply conserved functional elements. Additionally, a microsynteny scan for each synteny-conserved lncRNA was performed using the "jcvi.graphics.synteny" method within the "LncRNA Synteny Scan" module, enabling comprehensive visualization and evaluation of genomic structural conservation patterns. The conserved lncRNA networks across multiple species were visualized by Cytoscape (v3.9.0).

Sequence-conserved lncRNA prediction

The BLASTN tool was performed for sequence alignment analysis: (i) sequence similarity, evaluated using reciprocal BLASTN (*e*-value = 1×10^{-5}) to assess direct sequence identity between lncRNAs of the two species, and (ii) flanking sequence conservation, assessed by using the 200 nucleotides at the 5′ or 3′ end of each lincRNA as queries in a best BLASTN search against the reference genome. Specifically, flanking sequences were defined as 200 nucleotides upstream of the transcription start site (promoter region) or 200 nucleotides downstream of the transcription termination site. It provided local genomic context evidence for synteny. Bedtools intersect was used to verify whether BLASTN hits corresponded to the appropriate syntenic locations.

RNA second structure analysis

A. thaliana MSTRG.2210.1 and *S. lycopersicum MSTRG.2095.1* were predicted using RNAfold with default parameters (Gruber et al., 2008), along with consensus secondary structure predictions generated by the RNAalifold software (Bernhart et al., 2008).

Correlation analysis of epigenetic modification features

Regions of synteny-conserved lncRNAs were divided into upstream 2 kb, gene body, and downstream 2 kb regions, with each part further divided into 20 subregions. The modification level for each subregion was calculated using CGmaptools or deepTools. Finally, the correlation of epigenetic modification levels between two specified samples was computed using a Python script with the Spearman method.

Website implementation

The construction of the PERlncDB (http://perlncdb.liu-lab.com) platform adopts the architecture pattern of front-end separation, where the frontend handles visualization and display and the backend is responsible for data processing and transmission. PERlncDB relies on the efficient and fast web development framework Django, with integrated table appearance and regulation data stored and managed in the lightweight SQLite3 data management system. Additionally, to enhance web visualization and display of data resources, the database integrates SequenceServer and JBrowse2 bioinformatics tools. The project is deployed on an Nginx web server based on Linux (https://www.linux.org/), supported by the computing platform Alibaba Cloud (https://cn.aliyun.com/).

AUTHOR CONTRIBUTIONS

CL, WY, and YL conceived and conceptualized the project. YL and WY designed the database structure and wrote the original draft. YL collected omics datasets, analyzed the datasets, and built the database. CL supervised the project and reviewed and edited the manuscript. WY helped with data analysis. JL helped with database construction. BY helped with the secondary structure prediction of lncRNA. All authors agreed to the submitted version of the manuscript.

ACKNOWLEDGMENTS

We also thank the Information Center of Xishuangbanna Tropical Botanical Garden, CAS for providing the high-performance computing cluster and technical support. This work was supported by the National Natural Science Foundation of China (No. 31970609); Startup Fund from Xishuangbanna Tropical Botanical Garden; and "Top Talents Program in Science and Technology" from Yunnan Province.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Proportions of lncRNAs and PCGs occupied by 10 common histone modification marks in plant species.

Figure S2. Stacked bar charts showing histone modification enrichment across different genomic regions.

Figure S3. Three DNA methylation levels of lncRNAs (solid) and PCGs (dashed) in leaf samples from 15 plants.

Figure S4. Statistical analysis of the number of synteny-conserved lncRNAs.

Figure S5. Occupancy of ABA-related transcription factors at the *MARS* locus and surrounding regions.

Figure S6. The RNA-seq datasets and visualizations presented in *MARS* and *LINC-AP2* case study.

Figure S7. The secondary structures of Atha *MSTRG.2210.1* and Slyc *MSTRG.2095.1* predicted by RNAfold.

Table S1. The detailed information of datasets.

Table S2. The reference genome in plant.

REFERENCES

Amselem, J., Cornut, G., Choisne, N., Alaux, M., Alfama-Depauw, F., Jamilloux, V. *et al.* (2019) RepetDB: a unified resource for transposable element references. *Mobile DNA*, **10**, 6. Available from: https://doi.org/10.1186/s13100-019-0150-y

Ariel, F., Lucero, L., Christ, A., Mammarella, M.F., Jegu, T., Veluchamy, A. *et al.* (2020) R-loop mediated transaction of the APOLO long noncoding RNA. *Molecular Cell*, **77**, 1055–1065. Available from: https://doi.org/10.1016/j.molcel.2019.12.015

Ariel, F.D. & Manavella, P.A. (2021) When junk DNA turns functional: transposon-derived non-coding RNAs in plants. *Journal of Experimental Botany*, **72**, 4132–4143. Available from: https://doi.org/10.1093/jxb/erab073

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. *et al.* (2015) The Arabidopsis information resource: making and mining the "gold standard" annotated reference plant genome. *Genesis*, **53**, 474–485. Available from: https://doi.org/10.1002/dvg.22877

Bernhart, S.H., Hofacker, I.L., Will, S., Gruber, A.R. & Stadler, P.F. (2008) RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics*, **9**, 474. Available from: https://doi.org/10.1186/1471-2105-9-474

Brown, J., Pirrung, M. & McCue, L.A. (2017) FQC dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics*, **33**, 3137–3139. Available from: https://doi.org/10.1093/bioinformatics/btx373

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A. *et al.* (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & Development*, **25**, 1915–1927. Available from: https://doi.org/10.1101/gad.17446611

Chen, D., **Yuan, C.**, **Zhang, J.**, **Zhang, Z.**, **Bai, L.**, **Meng, Y.** ***et al.*** (2012) PlantNATsDB: a comprehensive database of plant natural antisense transcripts. *Nucleic Acids Research*, **40**, D1187–D1193. Available from: https://doi.org/10.1093/nar/gkr823

Chen, L. & **Zhu, Q.H.** (2022) The evolutionary landscape and expression pattern of plant lincRNAs. *RNA Biology*, **19**, 1190–1207. Available from: https://doi.org/10.1080/15476286.2022.2144609

Chow, C.N., **Tseng, K.C.**, **Hou, P.F.**, **Wu, N.Y.**, **Lee, T.Y.** & **Chang, W.C.** (2022) Mysteries of gene regulation: promoters are not the sole triggers of gene expression. *Computational and Structural Biotechnology Journal*, **20**, 4910–4920. Available from: https://doi.org/10.1016/j.csbj.2022.08.058

RNAcentral Consortium. (2021) RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Research*, **49**, D212–D220. Available from: https://doi.org/10.1093/nar/gkaa921

Deng, P., **Liu, S.**, **Nie, X.**, **Weining, S.** & **Wu, L.** (2018) Conservation analysis of long non-coding RNAs in plants. *Science China. Life Sciences*, **61**, 190–198. Available from: https://doi.org/10.1007/s11427-017-9174-9

Di Marsico, M., **Paytuvi Gallart, A.**, **Sanseverino, W.** & **Aiese Cigliano, R.** (2022) GreeNC 2.0: a comprehensive database of plant long non-coding RNAs. *Nucleic Acids Research*, **50**, D1442–D1447. Available from: https://doi.org/10.1093/nar/gkab1014

Ding, J., **Lu, Q.**, **Ouyang, Y.**, **Mao, H.**, **Zhang, P.**, **Yao, J.** ***et al.*** (2012) A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 2654–2659. Available from: https://doi.org/10.1073/pnas.1121374109

Du, X. (2024) The cellular RNA-dependent RNA polymerases in plants. *The New Phytologist*, **244**, 2150–2155. Available from: https://doi.org/10.1111/nph.20046

Fonouni-Farde, C., **Christ, A.**, **Blein, T.**, **Legascue, M.F.**, **Ferrero, L.**, **Moison, M.** ***et al.*** (2022) The Arabidopsis APOLO and human UPAT sequence-unrelated long noncoding RNAs can modulate DNA and histone methylation machineries in plants. *Genome Biology*, **23**, 181. Available from: https://doi.org/10.1186/s13059-022-02750-7

Fu, L.Y., **Zhu, T.**, **Zhou, X.**, **Yu, R.**, **He, Z.**, **Zhang, P.** ***et al.*** (2022) ChIP-hub provides an integrative platform for exploring plant regulome. *Nature Communications*, **13**, 3413. Available from: https://doi.org/10.1038/s41467-022-30770-1

Gao, R., **Liu, P.**, **Irwanto, N.**, **Loh, R.** & **Wong, S.M.** (2016) Upregulation of LINC-AP2 is negatively correlated with AP2 gene expression with turnip crinkle virus infection in *Arabidopsis thaliana*. *Plant Cell Reports*, **35**, 2257–2267. Available from: https://doi.org/10.1007/s00299-016-2032-9

Gao, Y., **Li, X.**, **Shang, S.**, **Guo, S.**, **Wang, P.**, **Sun, D.** ***et al.*** (2021) LincSNP 3.0: an updated database for linking functional variants to human long non-coding RNAs, circular RNAs and their regulatory elements. *Nucleic Acids Research*, **49**, D1244–D1250. Available from: https://doi.org/10.1093/nar/gkaa1037

Gao, Y., **Shang, S.**, **Guo, S.**, **Li, X.**, **Zhou, H.**, **Liu, H.** ***et al.*** (2021) Lnc2Cancer 3.0: an updated resource for experimentally supported lncRNA/circRNA cancer associations and web tools based on RNA-seq and scRNA-seq data. *Nucleic Acids Research*, **49**, D1251–D1258. Available from: https://doi.org/10.1093/nar/gkaa1006

Gruber, A.R., **Lorenz, R.**, **Bernhart, S.H.**, **Neuböck, R.** & **Hofacker, I.L.** (2008) The Vienna RNA websuite. *Nucleic Acids Research*, **36**, W70–W74. Available from: https://doi.org/10.1093/nar/gkn188

Guo, W., **Fiziev, P.**, **Yan, W.**, **Cokus, S.**, **Sun, X.**, **Zhang, M.Q.** ***et al.*** (2013) BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics*, **14**, 774. Available from: https://doi.org/10.1186/1471-2164-14-774

Guo, W., **Zhu, P.**, **Pellegrini, M.**, **Zhang, M.Q.**, **Wang, X.** & **Ni, Z.** (2018) CGmapTools improves the precision of heterozygous SNV calls and supports allele-specific methylation detection and visualization in bisulfite-sequencing data. *Bioinformatics*, **34**, 381–387. Available from: https://doi.org/10.1093/bioinformatics/btx595

He, K., **Cao, X.** & **Deng, X.** (2021) Histone methylation in epigenetic regulation and temperature responses. *Current Opinion in Plant Biology*, **61**, 102001. Available from: https://doi.org/10.1016/j.pbi.2021.102001

He, L., **Huang, H.**, **Bradai, M.**, **Zhao, C.**, **You, Y.**, **Ma, J.** ***et al.*** (2022) DNA methylation-free Arabidopsis reveals crucial roles of DNA methylation in regulating gene expression and development. *Nature Communications*, **13**, 1335. Available from: https://doi.org/10.1038/s41467-022-28940-2

Hirsch, C.D. & **Springer, N.M.** (2017) Transposable element influences on gene expression in plants. *Biochimica et Biophysica Acta, Gene Regulatory Mechanisms*, **1860**, 157–165. Available from: https://doi.org/10.1016/j.bbagrm.2016.05.010

Jin, J., **Lu, P.**, **Xu, Y.**, **Li, Z.**, **Yu, S.**, **Liu, J.** ***et al.*** (2021) PLncDB V2.0: a comprehensive encyclopedia of plant long noncoding RNAs. *Nucleic Acids Research*, **49**, D1489–D1495. Available from: https://doi.org/10.1093/nar/gkaa910

Kang, Y.J., **Yang, D.C.**, **Kong, L.**, **Hou, M.**, **Meng, Y.Q.**, **Wei, L.** ***et al.*** (2017) CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Research*, **45**, W12–W16. Available from: https://doi.org/10.1093/nar/gkx428

Kim, D., **Langmead, B.** & **Salzberg, S.L.** (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, **12**, 357–360. Available from: https://doi.org/10.1038/nmeth.3317

Kornienko, A.E., **Nizhynska, V.**, **Molla Morales, A.**, **Pisupati, R.** & **Nordborg, M.** (2023) Population-level annotation of lncRNAs in Arabidopsis reveals extensive expression variation associated with transposable element-like silencing. *Plant Cell*, **36**, 85–111. Available from: https://doi.org/10.1093/plcell/koad233

Law, J.A. & **Jacobsen, S.E.** (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews. Genetics*, **11**, 204–220. Available from: https://doi.org/10.1038/nrg2719

Lei, M., **Zhang, H.**, **Julian, R.**, **Tang, K.**, **Xie, S.** & **Zhu, J.K.** (2015) Regulatory link between DNA methylation and active demethylation in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 3553–3557. Available from: https://doi.org/10.1073/pnas.1502279112

Li, H., **Handsaker, B.**, **Wysoker, A.**, **Fennell, T.**, **Ruan, J.**, **Homer, N.** ***et al.*** (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079. Available from: https://doi.org/10.1093/bioinformatics/btp352

Li, Q., **Brown, J.B.**, **Huang, H.** & **Bickel, P.J.** (2011) Measuring reproducibility of high-throughput experiments. *The Annals of Applied Statistics*, **5**, 1752–1779. Available from: https://doi.org/10.1214/11-AOAS466

Li, S., **Yamada, M.**, **Han, X.**, **Ohler, U.** & **Benfey, P.N.** (2016) High-resolution expression map of the Arabidopsis root reveals alternative splicing and lincRNA regulation. *Developmental Cell*, **39**, 508–522. Available from: https://doi.org/10.1016/j.devcel.2016.10.012

Li, Y., **Li, X.**, **Yang, Y.**, **Li, M.**, **Qian, F.**, **Tang, Z.** ***et al.*** (2021) TRInc: a comprehensive database for human transcriptional regulatory information of lncRNAs. *Briefings in Bioinformatics*, **22**, 1929–1939. Available from: https://doi.org/10.1093/bib/bbaa011

Liang, Z., **Riaz, A.**, **Chachar, S.**, **Ding, Y.**, **Du, H.** & **Gu, X.** (2020) Epigenetic modifications of mRNA and DNA in plants. *Molecular Plant*, **13**, 14–30. Available from: https://doi.org/10.1016/j.molp.2019.12.007

Liu, S.J., **Nowakowski, T.J.**, **Pollen, A.A.**, **Lui, J.H.**, **Horlbeck, M.A.**, **Attenello, F.J.** ***et al.*** (2016) Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biology*, **17**, 67. Available from: https://doi.org/10.1186/s13059-016-0932-1

Liu, X., **Yang, S.**, **Zhao, M.**, **Luo, M.**, **Yu, C.W.**, **Chen, C.Y.** ***et al.*** (2014) Transcriptional repression by histone deacetylases in plants. *Molecular Plant*, **7**, 764–772. Available from: https://doi.org/10.1093/mp/ssu033

Lu, P., **Yu, S.**, **Zhu, N.**, **Chen, Y.R.**, **Zhou, B.**, **Pan, Y.** ***et al.*** (2018) Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening. *Nature Plants*, **4**, 784–791. Available from: https://doi.org/10.1038/s41477-018-0249-z

Lu, Z., **Marand, A.P.**, **Ricci, W.A.**, **Ethridge, C.L.**, **Zhang, X.** & **Schmitz, R.J.** (2019) The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nature Plants*, **5**, 1250–1259. Available from: https://doi.org/10.1038/s41477-019-0548-z

Luo, C., **Fernie, A.R.** & **Yan, J.** (2020) Single-cell genomics and epigenomics: technologies and applications in plants. *Trends in Plant Science*, **25**, 1030–1040. Available from: https://doi.org/10.1016/j.tplants.2020.04.016

Mathieu, O., **Reinders, J.**, **Caikovski, M.**, **Smathajitt, C.** & **Paszkowski, J.** (2007) Transgenerational stability of the Arabidopsis epigenome is coordinated by CG methylation. *Cell*, **130**, 851–862. Available from: https://doi.org/10.1016/j.cell.2007.07.007

Mattick, J.S., **Amaral, P.P.**, **Carninci, P.**, **Carpenter, S.**, **Chang, H.Y.**, **Chen, L.L.** ***et al.*** (2023) Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nature Reviews. Molecular Cell Biology*, **24**, 430–447. Available from: https://doi.org/10.1038/s41580-022-00566-8

Mazurov, E., **Sizykh, A.** & **Medvedeva, Y.A.** (2022) HiMoRNA: a comprehensive database of human lncRNAs involved in genome-wide epigenetic regulation. *Noncoding RNA*, **8**, 18. Available from: https://doi.org/10.3390/ncrna8010018

Niederhuth, C.E., **Bewick, A.J.**, **Ji, L.**, **Alabady, M.S.**, **Kim, K.D.**, **Li, Q.** ***et al.*** (2016) Widespread natural variation of DNA methylation within angiosperms. *Genome Biology*, **17**, 194. Available from: https://doi.org/10.1186/s13059-016-1059-0

Palos, K., **Nelson Dittrich, A.C.**, **Yu, L.**, **Brock, J.R.**, **Railey, C.E.**, **Wu, H.L.** ***et al.*** (2022) Identification and functional annotation of long intergenic non-coding RNAs in Brassicaceae. *Plant Cell*, **34**, 3233–3260. Available from: https://doi.org/10.1093/plcell/koac166

Pedro, D.L.F., **Amorim, T.S.**, **Varani, A.**, **Guyot, R.**, **Domingues, D.S.** & **Paschoal, A.R.** (2021) An atlas of plant transposable elements. *F1000Res*, **10**, 1194. Available from: https://doi.org/10.12688/f1000research.74524.1

Pertea, G. & **Pertea, M.** (2020) GFF utilities: GffRead and GffCompare. *F1000Research*, **9**, 10.12688/f1000research.23297.2.

Quek, X.C., **Thomson, D.W.**, **Maag, J.L.**, **Bartonicek, N.**, **Signal, B.**, **Clark, M.B.** ***et al.*** (2015) lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Research*, **43**, D168–D173. Available from: https://doi.org/10.1093/nar/gku988

Quinlan, A.R. & **Hall, I.M.** (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842. Available from: https://doi.org/10.1093/bioinformatics/btq033

Rai, M.I., **Alam, M.**, **Lightfoot, D.A.**, **Gurha, P.** & **Afzal, A.J.** (2019) Classification and experimental identification of plant long non-coding RNAs. *Genomics*, **111**, 997–1005. Available from: https://doi.org/10.1016/j.ygeno.2018.04.014

Ramirez, F., **Dundar, F.**, **Diehl, S.**, **Gruning, B.A.** & **Manke, T.** (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Research*, **42**, W187–W191. Available from: https://doi.org/10.1093/nar/gku365

Ran, X., **Zhao, F.**, **Wang, Y.**, **Liu, J.**, **Zhuang, Y.**, **Ye, L.** ***et al.*** (2020) Plant Regulomics: a data-driven interface for retrieving upstream regulators from plant multi-omics data. *The Plant Journal*, **101**, 237–248. Available from: https://doi.org/10.1111/tpj.14526

Roule, T., **Christ, A.**, **Hussain, N.**, **Huang, Y.**, **Hartmann, C.**, **Benhamed, M.** ***et al.*** (2022) The lncRNA MARS modulates the epigenetic reprogramming of the marneral cluster in response to ABA. *Molecular Plant*, **15**, 840–856. Available from: https://doi.org/10.1016/j.molp.2022.02.007

Shumate, A., **Wong, B.**, **Pertea, G.** & **Pertea, M.** (2022) Improved transcriptome assembly using a hybrid of long and short reads with StringTie. *PLoS Computational Biology*, **18**, e1009730. Available from: https://doi.org/10.1371/journal.pcbi.1009730

Sigman, M.J. & **Slotkin, R.K.** (2016) The first rule of plant transposable element silencing: location, location, location. *Plant Cell*, **28**, 304–313. Available from: https://doi.org/10.1105/tpc.15.00869

Simonini, S., **Bencivenga, S.**, **Trick, M.** & **Østergaard, L.** (2017) Auxin-induced modulation of ETTIN activity orchestrates gene expression in Arabidopsis. *The Plant Cell*, **29**, 1864–1882. Available from: https://doi.org/10.1105/tpc.17.00389

Song, L., **Huang, S.C.**, **Wise, A.**, **Castanon, R.**, **Nery, J.R.**, **Chen, H.** ***et al.*** (2016) A transcription factor hierarchy defines an environmental stress response network. *Science*, **354**, aag1550. Available from: https://doi.org/10.1126/science.aag1550

Statello, L., **Guo, C.J.**, **Chen, L.L.** & **Huarte, M.** (2021) Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews. Molecular Cell Biology*, **22**, 96–118. Available from: https://doi.org/10.1038/s41580-020-00315-9

Szczesniak, M.W. & **Wanowska, E.** (2024) CANTATAdb 3.0: an updated repository of plant long non-coding RNAs. *Plant & Cell Physiology*, **65**, 1486–1493. Available from: https://doi.org/10.1093/pcp/pcae081

Takuno, S. & **Gaut, B.S.** (2012) Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. *Molecular Biology and Evolution*, **29**, 219–227. Available from: https://doi.org/10.1093/molbev/msr188

Tang, H., **Krishnakumar, V.**, **Zeng, X.**, **Xu, Z.**, **Taranto, A.**, **Lomas, J.S.** ***et al.*** (2024) JCVI: a versatile toolkit for comparative genomics analysis. *iMeta*, **3**, e211. Available from: https://doi.org/10.1002/imt2.211

Veluchamy, A., **Jegu, T.**, **Ariel, F.**, **Latrasse, D.**, **Mariappan, K.G.**, **Kim, S.K.** ***et al.*** (2016) LHP1 regulates H3K27me3 spreading and shapes the three-dimensional conformation of the Arabidopsis genome. *PLoS One*, **11**, e0158936. Available from: https://doi.org/10.1371/journal.pone.0158936

Wang, H., **Niu, Q.W.**, **Wu, H.W.**, **Liu, J.**, **Ye, J.**, **Yu, N.** ***et al.*** (2015) Analysis of non-coding transcriptome in rice and maize uncovers roles of conserved lncRNAs associated with agriculture traits. *The Plant Journal*, **84**, 404–416. Available from: https://doi.org/10.1111/tpj.13018

Wang, J., **Qi, M.**, **Liu, J.** & **Zhang, Y.** (2015) CARMO: a comprehensive annotation platform for functional exploration of rice multi-omics data. *The Plant Journal*, **83**, 359–374. Available from: https://doi.org/10.1111/tpj.12894

Wang, T., **Xing, J.**, **Liu, Z.**, **Zheng, M.**, **Yao, Y.**, **Hu, Z.** ***et al.*** (2019) Histone acetyltransferase GCN5-mediated regulation of long non-coding RNA At4 contributes to phosphate starvation response in Arabidopsis. *Journal of Experimental Botany*, **70**, 6337–6348. Available from: https://doi.org/10.1093/jxb/erz359

Wang, Y., **Luo, X.**, **Sun, F.**, **Hu, J.**, **Zha, X.**, **Su, W.** ***et al.*** (2018) Overexpressing lncRNA LAIR increases grain yield and regulates neighbouring gene cluster expression in rice. *Nature Communications*, **9**, 3516. Available from: https://doi.org/10.1038/s41467-018-05829-7

Waseem, M., **Liu, Y.** & **Xia, R.** (2020) Long non-coding RNAs, the dark matter: an emerging regulatory component in plants. *International Journal of Molecular Sciences*, **22**, 86. Available from: https://doi.org/10.3390/ijms22010086

Wen, J., **Wu, Y.** & **Luo, Q.** (2024) DNA methyltransferases-associated long non-coding RNA PRKCQ-AS1 regulate DNA methylation in myelodysplastic syndrome. *International Journal of Laboratory Hematology*, **46**, 878–884. Available from: https://doi.org/10.1111/ijlh.14297

Xu, L., **Zhao, Z.**, **Dong, A.**, **Soubigou-Taconnat, L.**, **Renou, J.P.**, **Steinmetz, A.** ***et al.*** (2008) Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in *Arabidopsis thaliana*. *Molecular and Cellular Biology*, **28**, 1348–1360. Available from: https://doi.org/10.1128/MCB.01607-07

Yang, W., **Bai, Q.**, **Li, Y.**, **Chen, J.** & **Liu, C.** (2023) Epigenetic modifications: allusive clues of lncRNA functions in plants. *Computational and Structural Biotechnology Journal*, **21**, 1989–1994. Available from: https://doi.org/10.1016/j.csbj.2023.03.008

Yang, Y., **Tang, K.**, **Datsenka, T.U.**, **Liu, W.**, **Lv, S.**, **Lang, Z.** ***et al.*** (2019) Critical function of DNA methyltransferase 1 in tomato development and regulation of the DNA methylome and transcriptome. *Journal of Integrative Plant Biology*, **61**, 1224–1242. Available from: https://doi.org/10.1111/jipb.12778

Yu, G., **Wang, L.G.** & **He, Q.Y.** (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**, 2382–2383. Available from: https://doi.org/10.1093/bioinformatics/btv145

Yu, Y., **Zhang, Y.**, **Chen, X.** & **Chen, Y.** (2019) Plant noncoding RNAs: hidden players in development and stress responses. *Annual Review of Cell and Developmental Biology*, **35**, 407–431. Available from: https://doi.org/10.1146/annurev-cellbio-100818-125218

Zhang, H., **Lang, Z.** & **Zhu, J.K.** (2018) Dynamics and function of DNA methylation in plants. *Nature Reviews. Molecular Cell Biology*, **19**, 489–506. Available from: https://doi.org/10.1038/s41580-018-0016-z

Zhang, H. & **Zhu, J.K.** (2025) Epigenetic gene regulation in plants and its potential applications in crop improvement. *Nature Reviews. Molecular Cell Biology*, **26**, 51–67. Available from: https://doi.org/10.1038/s41580-024-00769-1

Zhang, X., **Luo, Z.**, **Marand, A.P.**, **Yan, H.**, **Jang, H.**, **Bang, S.** ***et al.*** (2024) A spatially resolved multi-omic single-cell atlas of soybean development. *Cell*, **188**(2), 550–567. Available from: https://doi.org/10.1016/j.cell.2024.10.050

Zhang, Y., **Liu, T.**, **Meyer, C.A.**, **Eeckhoute, J.**, **Johnson, D.S.**, **Bernstein, B.E.** ***et al.*** (2008) Model-based analysis of ChIP-seq (MACS). *Genome Biology*, **9**, R137. Available from: https://doi.org/10.1186/gb-2008-9-9-r137

Zhang, Z., **Xu, Y.**, **Yang, F.**, **Xiao, B.** & **Li, G.** (2021) RiceLncPedia: a comprehensive database of rice long non-coding RNAs. *Plant Biotechnology Journal*, **19**, 1492–1494. Available from: https://doi.org/10.1111/pbi.13639

Zhao, L., **Xie, L.**, **Zhang, Q.**, **Ouyang, W.**, **Deng, L.**, **Guan, P.** ***et al.*** (2020) Integrative analysis of reference epigenomes in 20 rice varieties. *Nature Communications*, **11**, 2658. Available from: https://doi.org/10.1038/s41467-020-16457-5

Zhao, L., **Zhou, Q.**, **He, L.**, **Deng, L.**, **Lozano-Duran, R.**, **Li, G.** ***et al.*** (2022) DNA methylation underpins the epigenomic landscape regulating genome transcription in Arabidopsis. *Genome Biology*, **23**, 197. Available from: https://doi.org/10.1186/s13059-022-02768-x

Zhao, T., **Mei, H.**, **Cao, Z.**, **Wang, L.**, **Tao, X.**, **Feng, S.** ***et al.*** (2021) Absence of CG methylation alters the long noncoding transcriptome landscape in multiple species. *FEBS Letters*, **595**, 1734–1747. Available from: https://doi.org/10.1002/1873-3468.14100

Zhao, T., **Zhan, Z.** & **Jiang, D.** (2019) Histone modifications and their regulatory roles in plant development and environmental memory. *Journal of Genetics and Genomics*, **46**, 467–476. Available from: https://doi.org/10.1016/j.jgg.2019.09.005

Zhao, X., **Li, J.**, **Lian, B.**, **Gu, H.**, **Li, Y.** & **Qi, Y.** (2018) Global identification of Arabidopsis lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. *Nature Communications*, **9**, 5056. Available from: https://doi.org/10.1038/s41467-018-07500-7

Zhao, Y., **Li, H.**, **Fang, S.**, **Kang, Y.**, **Wu, W.**, **Hao, Y.** ***et al.*** (2016) NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Research*, **44**, D203–D208. Available from: https://doi.org/10.1093/nar/gkv1252

Zhao, Z., **Yang, Y.**, **Iqbal, A.**, **Wu, Q.** & **Zhou, L.** (2024) Biological insights and recent advances in plant long non-coding RNA. *International Journal of Molecular Sciences*, **25**, 11964. Available from: https://doi.org/10.3390/ijms252211964

Zhou, B., **Zhao, H.**, **Yu, J.**, **Guo, C.**, **Dou, X.**, **Song, F.** ***et al.*** (2019) Experimentally validated plant lncRNAs in EVLncRNAs database. *Methods in Molecular Biology*, **1933**, 431–437. Available from: https://doi.org/10.1007/978-1-4939-9045-0_27

Zhou, H., **Ren, F.**, **Wang, X.**, **Qiu, K.**, **Sheng, Y.**, **Xie, Q.** ***et al.*** (2022) Genome-wide identification and characterization of long noncoding RNAs during peach (*Prunus persica*) fruit development and ripening. *Scientific Reports*, **12**, 11044. Available from: https://doi.org/10.1038/s41598-022-15330-3