BMC Genomics

RESEARCH

Open Access

Genome-wide identification of long non-coding RNA for *Botrytis cinerea* during infection to tomato (*Solanum lycopersicum*) leaves

Haojie Shi[1†], Guijuan Ding[1†], Yun Wang[1], Jiaqi Wang[1], Xiaoli Wang[3], Dan Wang[2*] and Ping Lu[1*]

Abstract

Long non-coding RNA (lncRNA) plays important roles in animals and plants. In filamentous fungi, however, their biological function in infection stage has been poorly studied. Here, we investigated the landscape and regulation of lncRNA in the filamentous plant pathogenic fungus *Botrytis cinerea* by strand-specific RNA-seq of multiple infection stages. In total, 1837 lncRNAs have been identified in *B. cinerea*. A large number of lncRNAs were found to be antisense to mRNAs, forming 743 sense-antisense pairs, of which 55 antisense lncRNAs and their respective sense transcripts were induced in parallel as the infection stage. Although small RNAs were produced from these overlapping loci, antisense lncRNAs appeared not to be involved in gene silencing pathways. In addition, we found the alternative splicing events occurred in lncRNA. These results highlight the developmental stage-specific nature and functional potential of lncRNA expression in the infection stage and provide fundamental resources for studying infection stage-induced lncRNAs.

Keywords *Botrytis cinerea*, Transcriptome, Long noncoding RNAs, *Solanum lycopersicum*, Alternative splicing

Introduction

Eukaryotic genomes encode a large number of non-coding transcripts, which function in key biological processes [1]. Long non-coding RNAs (lncRNAs) are defined as non-coding RNAs longer than 200 nucleotides, which are transcribed by RNA polymerase II and share common features with mRNAs except for protein-coding capacity [2]. While lncRNAs may exhibit lower levels of conservation and expression compared to mRNAs, they demonstrate a greater degree of tissue specificity [3–5]. LncRNA acts as a functional biomolecule that interacts with other components in the cell, including DNA, RNA and proteins [6–8]. The regulatory element responsible for gene regulation was found within the region of the lncRNA, and the level of expression of the lncRNA was observed to have a positive correlation with the activity

†Haojie Shi and Guijuan Ding contributed equally to this work.

*Correspondence:
Dan Wang
wangdan_star@163.com
Ping Lu
luping@zafu.edu.cn
[1]The Key Lab for Biology of Crop Pathogens and Insect Pests and Their Ecological Regulation of Zhejiang Province, College of Advanced Agricultural Sciences, Zhejiang A&F University, Hangzhou 311300, China
[2]State Key Laboratory of Subtropical Silviculture, School of Forestry and Biotechnology, Zhejiang A & F University, Hangzhou 311300, China
[3]Jiangsu Provincial Key Construction Laboratory of Probiotics Preparation, College of Life Science and Food Engineering, Huaiyin Institute of Technology, Huai'an 223003, China

of the regulatory element [9]. LncRNA can affect gene activity by influencing the process of transcription [10, 11]. Notably, recent studies have revealed that lncRNA can also function by encoding micropeptides consisting of no more than 100 amino acids [12–14].

The prevalence of lncRNAs has been extensively studied in animals and plants. Numerous studies have reported an association between long non-coding RNAs (lncRNAs) in mammals and cancer, including prostate cancer, breast cancer, lung cancer [15]. In plants, there has been a rise in comprehensive research exploring the roles of lncRNAs, including cold stress tolerance [16], drought/salt stress tolerance [17], and immune responses [18]. In recent study, lncRNAs also have been reported to be involved in plant responses to infection by *B. cinerea* and *Verticillium dahliae* [19].

Research on the biological function of lncRNAs in fungi is limited, with a majority of studies centered around yeast, including meiotic divisions [20], repress meiosis [21], and homologous chromosome pairing [22]. In addition, lncRNAs play an important regulatory role in the virulence and growth of plant pathogenic fungi. Antisense lncRNA *GzmetE-AS* was transcribed from the opposite strand of *GzmetE*, regulating its sense gene through the RNAi pathway in *Fusarium graminearum* [22]. The *lncRsp1* could directly regulate the expression of *Fgsp1*, thereby indirectly regulating the expression of several deoxynivalenol (DON) biosynthesis genes, including TRI4, TRI5, TRI6 and TRI13 in *F. graminearum* [23]. The latest research shows that several deoxynivalenol (DON) biosynthesis genes (including TRI5, TRI6 and TRI11) were also regulated by antisense lncRNA in *F. graminearum* [24].

B. cinerea causes disease in more than 1400 plant species, including crucial commercial crops such as grapes, strawberries, and tomatoes [25, 26]. The complete genome sequence of *B. cinerea* (strain B05.10) has significantly facilitated the detection and characterization of genes associated with virulence, metabolism, signal transduction, and resistance to fungicides [27–30]. Despite the prevalence of non-coding regions (52.4%) in the genome (data not shown), no functional lncRNAs have been identified in *B. cinerea*.

In this study, we aimed to identify lncRNAs present in *B. cinerea* during the inoculation of tomato leaves that may be involved in the infection process. This study report the genome-wide identification of lncRNAs during successive stages of infection, including conidia germination (6 h), pre-penetration (12 h), biotrophic stage (24 h), and necrotrophic stage (48 h) [31]. Combining strand-specific RNA-Seq data from vegetative and infection stages, we identified infection-specifically expressed lncRNAs. Thousands of lncRNAs that exhibit dynamic expression patterns were found, expanding the known roles of non-coding RNAs in infection.

Materials and methods

Plant, fungal pathogen and inoculations

The tomato (*Solanum lycopersicum* cv. Moneymaker) seedlings were planted in green house at 25 °C with 16 h light/8 h dark for four weeks and used for inoculation. Wild type strain B05.10 [32] of *B. cinerea* was used in this study for infection experiments. Conidia were harvested from the cultures, washed with sterile water, and the inoculum suspension with the final concentration adjusted to 5×10^6 conidia/mL was used for inoculating tomato leaves. Leaves of the seedlings were immersed in the conidial suspension for 3 min, and the leaves samples were collected at 6, 12, 24, and 48 h, respectively. Conidia not used for the inoculation were cultured in the PDB medium at 22 °C with 180 rpm as the control treatment. Two replicates of the experiments were performed during the infection stage (each replicate included 6 seedlings). Three replicates of the experiments were performed during vegetative stage. All samples were frozen in liquid nitrogen and stored at -80 °C.

RNA library construction and strand-specific sequencing

Total RNA were extracted with the RNA prep Pure Plant Kit (Tiangen, Beijing, China). Strand-specific cDNA synthesis with NEBNext® Ultra™ DNA Library Prep Kit (San Diego, CA, USA) and sequencing was performed on the Illumina HiSeq® 2500 System, with a 2×150 bp paired-end read mode. The RNA-seq data generated in this project has been deposited in the NCBI Sequence Read Archive (SRA) under accession numbers listed in Table S1.

Transcriptome assembly

Raw reads were processed to remove low-quality reads and trim adapter sequences using NGS QC Toolkit v2.3.3 [33]. The clean reads were mapped to the *B. cinerea* reference genome (ASM83294v1, Ensembl fungi v53) using HISAT2 v2.04 [34]. The transcriptome was assembled with de novo annotation using StringTie v2.1.3 [35]. Transcripts per kilobase million (TPM) were used as the expression value. If the expression value of a transcript was <1 TPM across all samples, the transcript was defined to be predicted but not detected. Detected transcripts were used for subsequent analysis.

LncRNA identification

Transcripts with sequences shorter than 200 nucleotides were filtered out. The assembled transcripts were compared with coding genes and categorized using gffcompare [36]. Intergenic transcripts (class codes "u" and "p") were regarded, sense transcripts (class codes "m",

"n", "o" and "j"), antisense transcripts (class codes "x") and intronic transcript (class codes "i"). Transcripts containing any known Pfam domain and non-coding RNAs were removed using the Rfam databases release and Pfam database. The coding potential of transcripts was assessed using CPAT v.1.2.2 [37].

Differential expression analysis of lncRNAs during infection stage

Significantly differentially expressed lncRNAs were identified from four comparisons, including Inf6 h/Hyp6 h, Inf12 h/Hyp12 h, Inf24 h/Hyp24 h and Inf48 h/Hyp48 h. Transcripts per kilobase million (TPM) were used to determine expression values. Salmon v1.9 [38] was used to calculate the TPM of lncRNAs in each sample. The fold-change in transcript expression value was calculated using DESeq2 [39]. Transcripts were identified as differentially expressed between treatment and control with parameters of |Log2 FC| >1 and adj *P*-value < 0.05.

RNA extraction and reverse transcription-quantitative PCR (RT-qPCR)

Total RNA was extracted from *B. cinerea* using RNA Purification Kit (TianGen, Beijing, China) and stored at -80 °C, and the first strand cDNA was synthesized using the TranScript One-Step gDNA Removal and cDNA Synthesis SuperMix Kit (TransGen, Beijing, China). RT-qPCR was performed using a qPCR SYBR premix Ex TaqII kit (TaKaRa, Tokyo, Japan). Relative transcript levels of different genes among various treatments were evaluated using $2^{-\Delta\Delta CT}$ method [40]. The mRNA expression levels were normalized using GAPDH. Three biological replicates were performed for each sample, with three technical replicates. The specific primers were listed in Table S3.

sRNA-seq analysis

sRNA-Seq data for vegetative and infection stages were downloaded from NCBI SRA (PRJNA496584) [41]. Reads quality control was performed with the FastQCv.0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapter trimming, read size was performed with the Cutadapt v2.6 using the settings -m 60 -q 30,30 [42]. sRNA-Seq data were aligned to the *B. cinerea* reference genome (ASM83294v1, Ensembl fungi v53) using Bowtie2 v2.22 [43]. After read mapping, reads with 17–27 nt were extracted, using the 'reformat.sh' script in BBMap tools (https://sourceforge.net/projects/bbmap). To conve rt the location of sRNA into a bed file format, samtools depth was used with the alignment data of each sample. A customized Python script was then used to process the depth data and define sRNA loci by binding together adjacent nucleotides with a depth value more than 10 reads. Using the feature counts program v2.0.1 [44] and R script to calculate sRNA read counts and TPM for different transcript types.

Identification of alternative splicing

Alternative splicing landscape were extracted using SUPPA2 [45]. Percent Spliced-In (PSI) index using SUPPA2 for each AS event.

Results

Genome-wide identification of lncRNAs in *Botrytis cinerea*

To identify lncRNAs in *B. cinerea*, RNA-Seq datasets from *B. cinerea* strain B05.10 were cataloged for both control and treatment samples that were non-inoculated and inoculated on tomato leaves, respectively (Fig. 1A; Fig. S1). In total, 386.1 million reads were mapped to the *B. cinerea* genome with 29,424 predicted transcripts from 18,063 gene loci (Table S1). Established pipelines [46] were used to detect lncRNAs (Fig. 1B). To distinguish lncRNAs, following sequential stringent filters of the transcripts were employed. Firstly, 26,425 transcripts were detected longer than 200 nucleotides and with TPM ≥ 1 at least one developmental or infection stage. Novel transcripts (14291) were identified using gffcompare [36]. Transcripts with low coding potential were further scanned against the Pfam and Rfam databases to filter out transcripts encoding protein domains and/or harboring any known structural RNA motifs (E value < 0.001). Finally, lncRNAs were distinguished by coding potentials of < 0.54. The resulting lncRNA (1837) candidates were classified into four categories based on their positions (Fig. 1C): 51.4% (945) from intergenic regions, 8.1% (148) from the sense strand, 40.4% (743) from the antisense strand, and 0.01% (1) from intronic regions (Fig. 1D).

Characteristics of lncRNAs in *Botrytis cinerea*

Properties including exon number, length, GC ratio and ORF length were investigated by mRNA comparisons. Exon number distribution revealed that lncRNAs generally contained one or two exons (Fig. 2A). LncRNAs (median 865 nt) had shorter transcript lengths than did mRNAs (median 1977 nt) (Fig. 2B). GC content results revealed that lncRNAs exhibited significantly lower than those of protein-coding genes (Fig. 2C). lncRNAs typically have no protein-coding potential, and the results revealed that the length of ORFs in lncRNAs was significantly shorter than protein-coding genes (Fig. 2D). Collectively, these results demonstrated that lncRNAs possessed fewer exons, had shorter ORF lengths, transcript lengths and lower GC content.

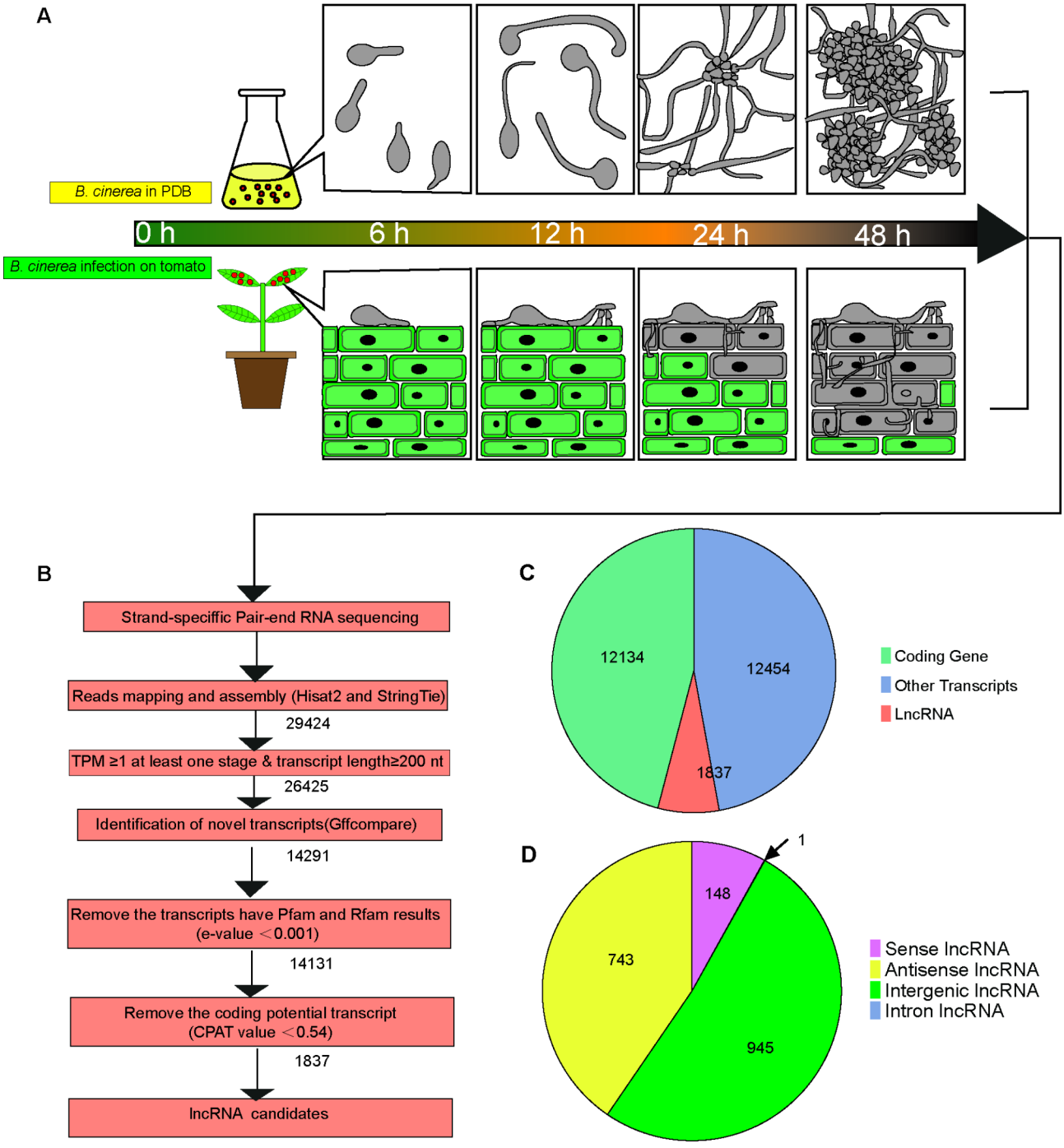

Fig. 1 Schematic pipeline for identification of lncRNAs in *Botrytis cinerea*. **A** Key time nodes of the infectious process of *B. cinerea*. **B** Bioinformatic pipeline for lncRNA identification. **C** Number of predicted transcripts. **D** Number of lncRNAs classification

Antisense lncRNA was the main type of differentially expressed lncRNA in *Botrytis cinerea* during infection stage

The expression dynamics of lncRNAs were assessed by generating heatmaps based on TPM from detected 14,236 mRNAs and 1837 lncRNAs (Fig. 3A and B). The result showed that there were stage-specific expression patterns in both mRNAs and lncRNAs. TPM values indicated that expression levels of lncRNAs were much lower than expression levels of mRNAs (Fig. 3C). mRNAs had similar expression levels across the vegetative and infection stages. Surprisingly, expression levels of lncRNAs increased gradually during the vegetative stage but maintained high expression levels across all infection stages, which indicated lncRNAs played a crucial role in

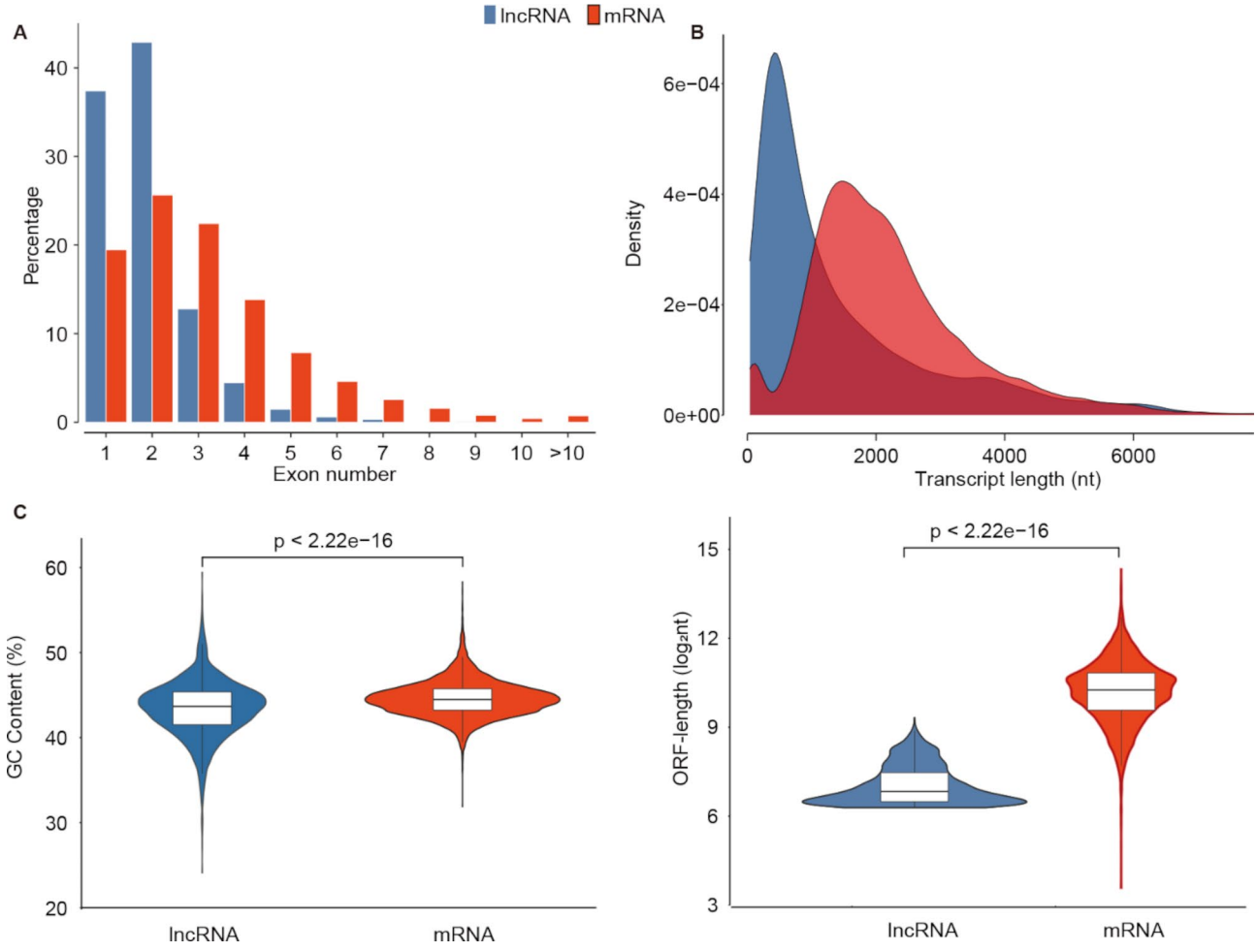

Fig. 2 Characteristics of lncRNAs in *Botrytis cinerea*. **A** The number of exons of lncRNAs compared with protein-coding genes. **B** Distribution of transcript length of lncRNAs compared with protein-coding genes. **C** The GC content of lncRNAs and protein-coding genes. **D** Distribution of ORF length of lncRNAs compared with protein-coding genes. *P* values are from two-tailed Wilcoxon rank sum test

disease development. The expression levels of intergenic lncRNAs, sense lncRNAs and antisense lncRNAs were similar, suggesting that all types of lncRNAs participated in the infection stages (Fig. 3D).

To further discover lncRNAs potentially involved in pathogenicity, we compared the expression level of lncRNAs between the vegetative stage at 6, 12, 24 and 48 h and infections stages at 6, 12, 24 and 48 h, respectively. Differentially expressed lncRNAs were identified from four comparisons, including Inf6 h/Hyp6 h, Inf12 h/Hyp12 h, Inf24 h/Hyp24 h, Inf48 h/Hyp48 h. There were 33, 125, 98, 274 differentially expressed lncRNAs in these four comparisons, respectively (Fig. 3E and F), and the detailed differentially expressed lncRNAs were listed in Table S2. The number of differentially expressed lncRNAs increased during the infection progress. However, down-regulated lncRNAs were the main types and multiplied with the course of infection (Fig. 3E), of which antisense lncRNAs (149, 72.7%) were the main types of differentially expressed lncRNAs (Fig. 3G). Differentially expressed lncRNAs were divided into two categories including high expression in vegetative and infection stages (Fig. 3G), suggesting that lncRNAs had two modes of positive and negative regulation. Taken together, down-regulated antisense lncRNAs were the main types of lncRNA and may have important functions in *B. cinerea* during infection.

Antisense lncRNAs and target sense transcripts

Generally, antisense lncRNAs can participate in a wide range of controlling sense gene expression [47–49]. Interestingly, gene expression correlation in many differentially expressed antisense lncRNAs and sense mRNA pairs (55 out of 149 pairs with Pearson's correlation $|r|>0.7$; Fisher's exact test, $P<0.05$) (Fig. 4A), all (55 pairs) of which were positively correlated (Fig. 4B). We investigated whether the differentially expressed antisense lncRNAs are antisense to genes involved in a specific biological process. The results demonstrated that genes with positively correlated were significantly enriched for

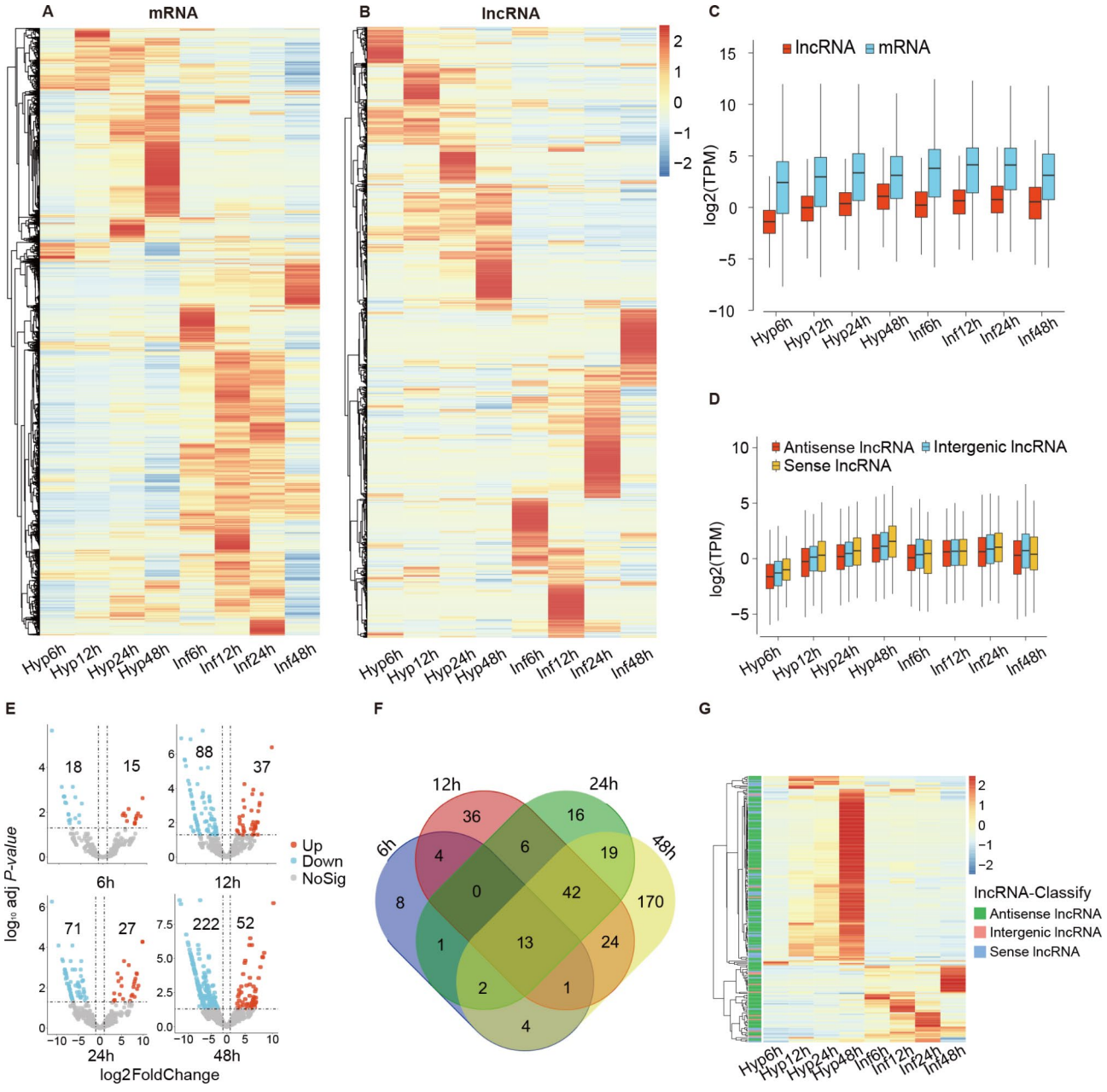

Fig. 3 lncRNA expression level and pattern in *Botrytis cinerea*. **A** and **B** Heatmap of the expression of mRNAs and lncRNAs. High (orange to red) and low (yellow to blue) expression levels are depicted as Z-scores for each gene. **C** Boxplot of mRNA and lncRNA expression patterns across developmental and infection stages. **D** Boxplot of different types of lncRNA expression patterns across developmental and infection stages. **E** Volcano plot of the significantly up-regulated and down-regulated lncRNAs in the Inf6 h/Hyp6 h, Inf12 h/Hyp12 h, Inf24 h/Hyp24 h and Inf48 h/Hyp48 h (|log2 Fold Change| ≥ 1, adjusted *P*-value < 0.05). **F** Venn diagram showing non-overlap and overlap of different expressed lncRNAs. **G** Heatmap of the different expressed lncRNAs. High (orange to red) and low (yellow to blue) expression levels are depicted as Z-scores for each gene. Three different types (intergenic lncRNAs, sense lncRNAs, and antisense lncRNAs) of lncRNAs are indicated with different colors as marked

regulation of cellular homeostasis, inorganic ion homeostasis, and regulation of mitotic spindle organization (Fig. 4C). The positively correlated sense mRNAs were enriched for the inorganic ion homeostasis, siderophore metabolic process, iron import into cell, hydroxymate-containing siderophore biosynthetic process, ferricrocin metabolic process, ferricrocin biosynthetic process, ferrichrome metabolic process, ferrichrome biosynthetic process, siderophore biosynthetic process (Fig. 4C). So far, several studies have shown that iron ion play import roles in plant pathogenic fungi during infection [50, 51]. These results suggested that antisense lncRNA may participate in regulating the iron ions of target genes in *B. cinerea* during infection.

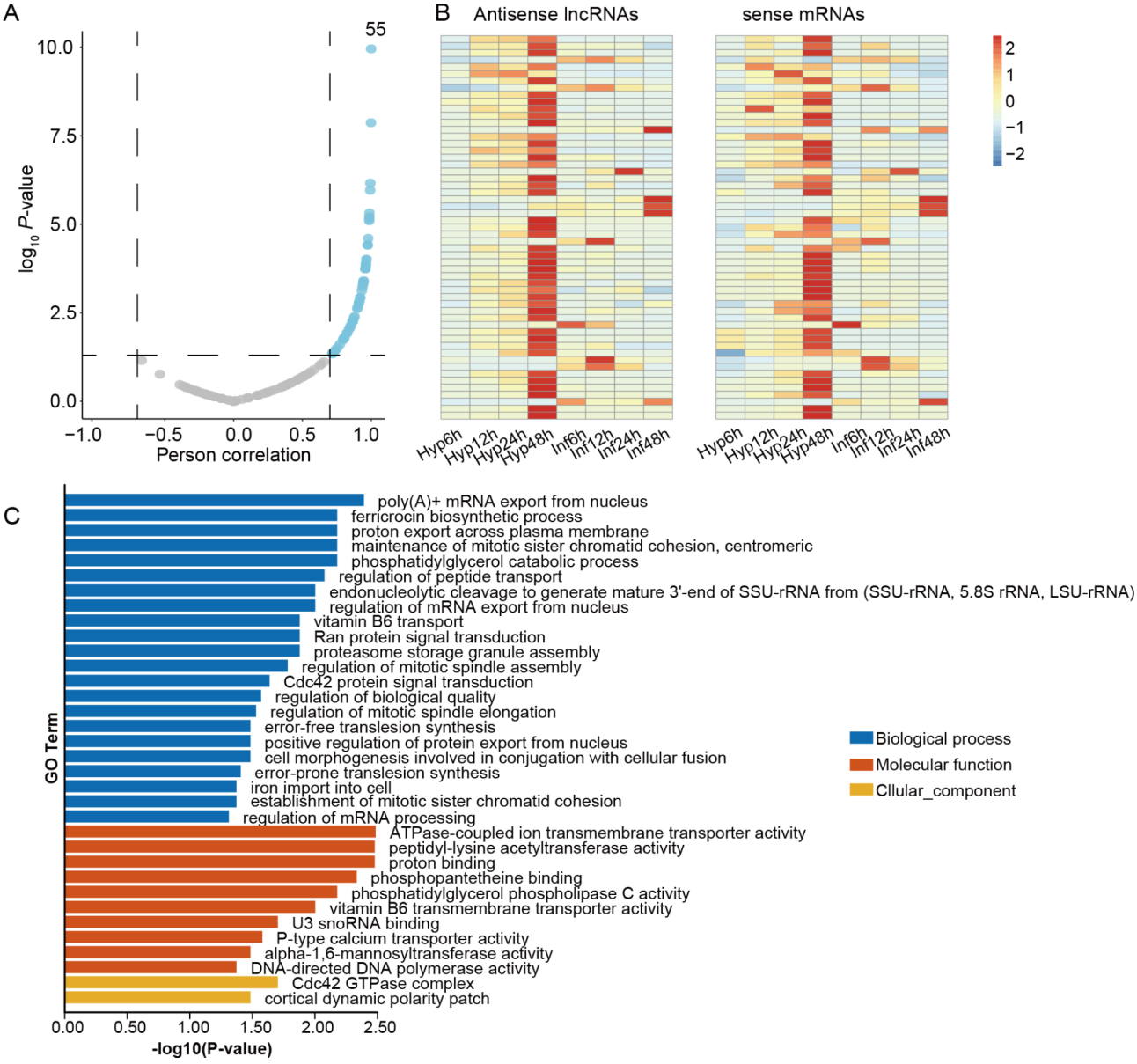

Fig. 4 Co-expression of antisense lncRNA and their sense mRNA. **A** Pearson's correlation between antisense lncRNA and sense mRNA ($|r| \geq 0.7$; Fisher's exact test, *P*-value < 0.05). **B** Parallel induction of sense mRNA and antisense lncRNA pairs during infection stages. High (orange to red) and low (yellow to blue) expression levels are depicted as Z-scores for each gene. **C** Enriched gene ontology (GO) terms to specific in positively correlated sense mRNAs (*P*-value < 0.05)

The transcriptional levels of lncRNAs in *Botrytis cinere* a are related to host interaction

Generally, antisense lncRNAs tune sense mRNA expression by trans-regulation. According to the overlapping types of sense mRNA, antisense lncRNAs were classified into three categories, as described in Fig. 5A [52]: "head to head", where antisense lncRNAs and sense mRNAs overlap on the 5' ends; "tail to tail", where antisense lncRNAs and sense mRNAs overlap on the 3' ends; "overlap", where one of the mRNAs overlaps the other. In this study, we found that the "Overlap" comprised the majority (71.8%) of differentially expressed antisense lncRNAs, "Tail to tail" was the second most common (17.5%) differentially expressed antisense lncRNAs, followed by "head to head" (10.7%) (Fig. 5B). To explore the overlapping position distribution of differentially expressed antisense lncRNAs and sense mRNA, we counted the overlapping distribution in sense mRNA. Compared to the other positions, the distribution of overlapping region was skewed toward to the 3' end of transcript (Fig. 5C). To further confirm the expression pattern of the lncRNAs during vegetative and infection stages, we randomly selected three examples that showed each type

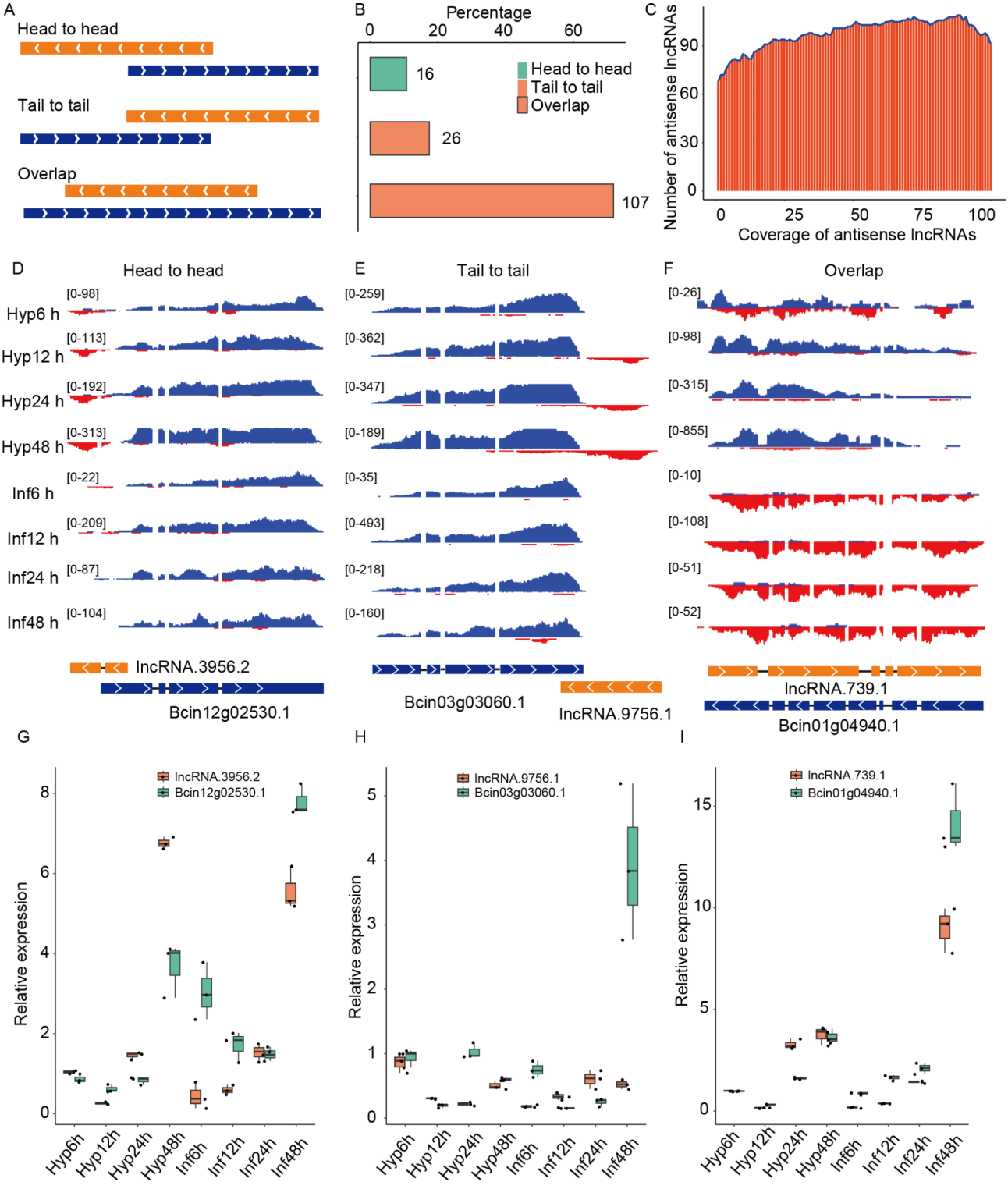

Fig. 5 Coverage of differentially expressed antisense lncRNAs. **A** Classification of antisense lncRNA. **B** Number of different coverage types of differentially expressed antisense lncRNA. **C** Identification of coverage in sense mRNA. **D-F** RNA sequencing read coverage of three antisense lncRNA and their sense transcript. Antisense lncRNAs are shown in orange and their sense transcripts in blue. **G-I** Expression of lncRNA and target genes detected by RT-qPCR

of antisense lncRNAs to verify their expression by RT-qPCR at 6, 12, 24 and 48 h across vegetative and infections stages, respectively (Fig. 5G-I). For all the examples, we observed expression trends that were consistent with RNA-seq data (Fig. 5D-F). Primers for RT-qPCR were designed specifically to distinguish lncRNAs and their sense mRNAs (Table S3). Collectively, these results suggested that lncRNAs are present in *B. cinerea* and their expression patterns correlate with phases of interactions with tomato leaves.

Antisense lncRNA are not the major source for endogenous sRNA production in *Botrytis cinerea*

The most common mechanism involving antisense transcripts is the RNA interference (RNAi) pathway, which incorporates small RNAs (sRNAs) generated from the double-stranded RNA regions. To explore the effect of sRNA production in the antisense lncRNAs loci, previously published sRNA-seq data for vegetative and infection stages were analyzed [41]. Compared to the vegetative stage, a large number of sRNAs were identified during the infection stage. The majority of sRNA lengths ranged from 20 to 24 nt with a peak at 21 nt (Fig. 6A). Surprisingly, sRNA reads mapped to antisense lncRNA were not increased compared to mRNAs (Fig. 6B). To search for any lncRNAs associated with sRNA enriched loci that could be the indication of transcriptional gene silencing events, we performed statistics on the overlap of 27,918 unique sRNA from 109,518 loci identified in previous studies with lncRNAs and mRNAs [41]. The analysis results manifested that most (65.9%) of the sRNA reads were mapped in the non-coding region (Fig. 6C), while only a part of sRNA reads (41) were mapped to antisense lncRNA region. Taken together, these results suggested that the antisense lncRNA may not be the major source for endogenous sRNA production.

Alternative splicing events in lncRNAs

Similar with mRNA, lncRNA also can generate different transcript isoforms by alternative splicing [53]. However, the biological functions of lncRNA transcript isoforms have not been discovered and studied in fungi. Surprisingly, we found that multiple antisense lncRNAs may simultaneously regulate the same target gene, suggesting that alternative splicing also occurred in lncRNAs. To test whether alternative splicing is involved in lncRNA regulation, we further analyzed the alternative splicing (AS) events in lncRNAs. AS events were classified into four basic types: intron retention (IR), alternative 5'-donor (A5), alternative 3'-acceptor (A3), and exon skipping (ES) [54] (Fig. 7A). In this research, a total of 233 AS events were identified in lncRNAs (Fig. 7B). Among these AS events, A3 comprised the majority (37.8%) of AS events, IR was the second most common (33.5%) type of AS events, and followed by A5 (24.5%), the number of ES events was the least (4.3%).

To examine the variation of splicing events throughout development, Percent Spliced-In (PSI) index were calculated using SUPPA2 for each AS event across different samples [45]. PSI index was calculated as the fraction of the inclusion reads to the total reads (both inclusion and exclusion reads) to measure the inclusion level of a given splicing event. Hierarchical clustering revealed that the PSI values were variable in different stages (Fig. 7C), suggesting that they are regulated in a stage-specific manner. To validate the accuracy of the AS events detected, we randomly selected one example that showed two type of AS events (including IR and A3) for reverse transcription (RT)-PCR and Sanger sequencing (Fig. 7D). Primers (Table S3) suitable for distinguishing different splice isoforms were designed and used for RT-PCR with RNA from vegetative and infection stages, respectively. The results of agarose gel electrophoresis showed that the relative brightness of the three bands from vegetative and infection stages were different for lncRNA.3649,

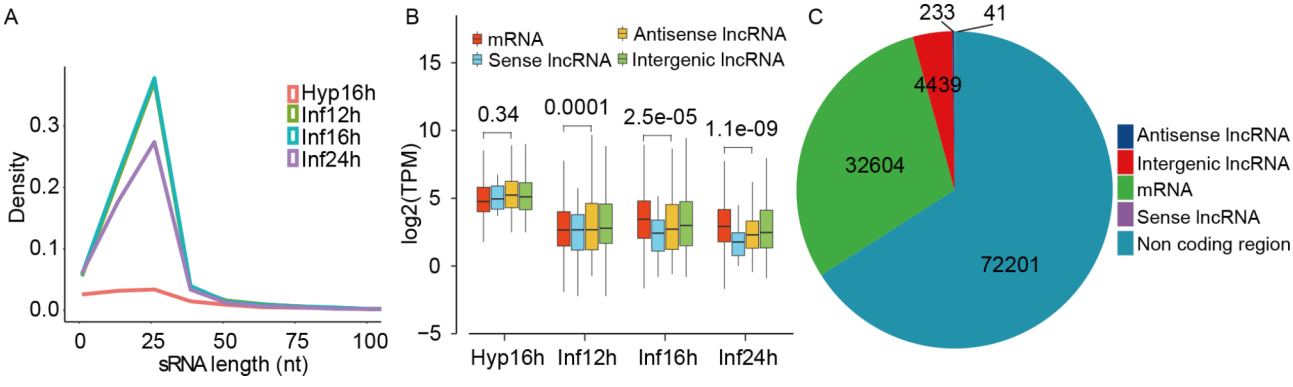

Fig. 6 Antisense lncRNA associated with small RNA-enrich loci. **A** Read length distribution of unique sRNA sequences. **B** sRNA reads mapped to mRNAs without antisense transcript, sense lncRNAs, antisense lncRNAs and intergenic lncRNA are represented as TPM.*P* values are from two-tailed Wilcoxon rank sum test. **C** Fractions of sRNA reads mapped to mRNA, lncRNA and non-coding region

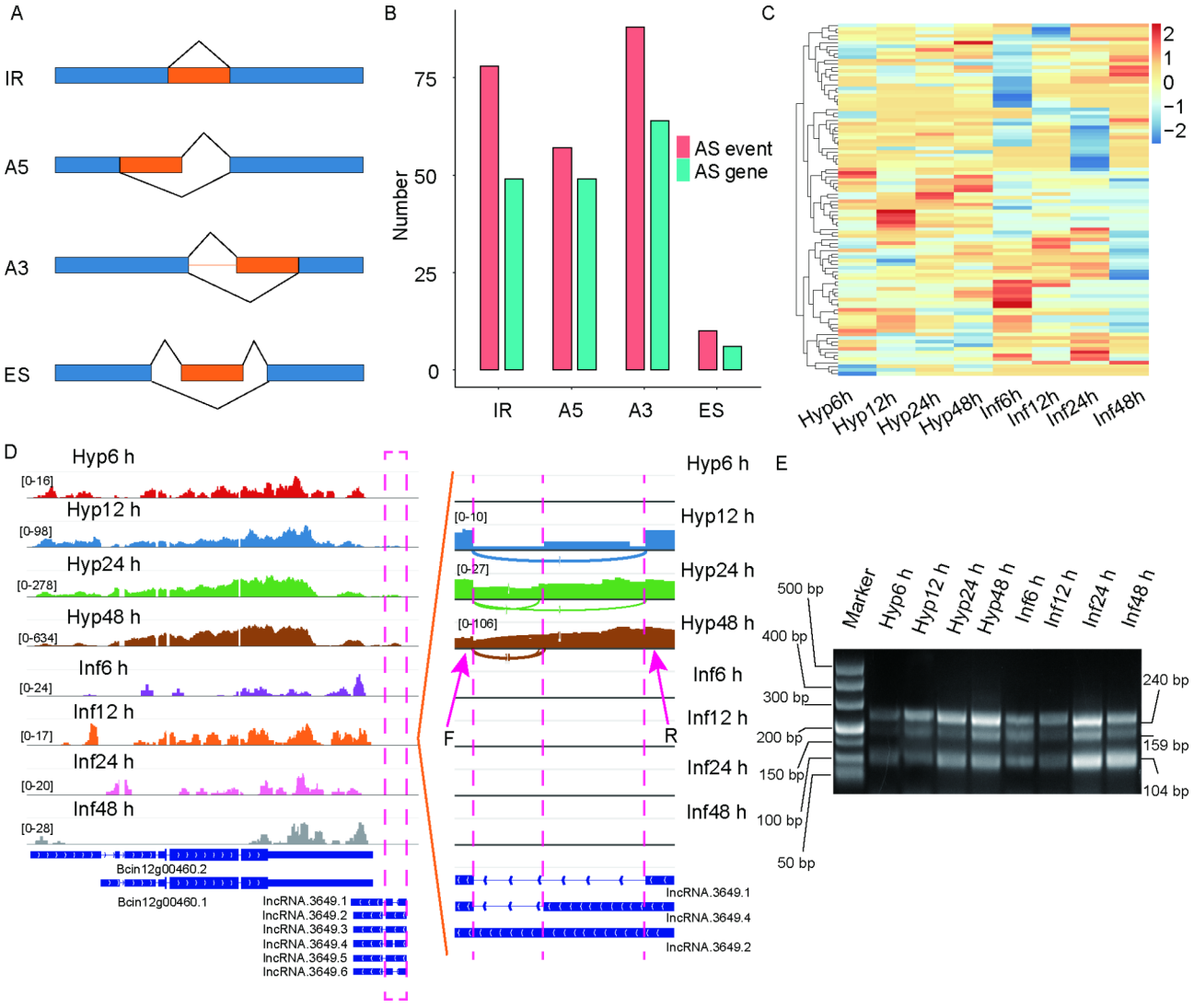

Fig. 7 Identification of alternative splicing (AS) in lncRNA. **A** Schematic drawing of four basic alternative splicing events. intron retention (IR), alternative 5′-donor (A5), alternative 3′-acceptor (A3) and exon skipping (ES). **B** Number of alternative splicing events and gene. **C** Heatmap of Percent Spliced-In (PSI) values across different samples. High (orange to red) and low (yellow to blue) PSI values are depicted as Z-scores for each AS event. **D** RNA sequencing (RNA-Seq) read coverage of the lncRNA.3649 is shown in left panel. Alternative splicing region of lncRNA.3649 is shown in right panel. PCR primers (F, forward and R, reverse) are designed to flank the splicing events. **E** The bands of agarose gel electrophoresis show DNA makers and PCR results in eight stages

indicating that expression of splice isoforms may exhibit a stage-preferential pattern (Fig. 7E, Fig. S2).

Discussion

In living organisms, complex post-transcript regulatory mechanisms are required to ensure regulation of tissue/stage specific gene expression. Much of the non-protein coding portion of the genome has historically been regarded as junk DNA. In particular, lncRNA have received considerable attention in recent years due to their widespread effects in human [55]. Numerous studies have shown that lncRNAs were involved in the regulation of prostate cancer, breast cancer, lung cancer in human [15]. LncRNAs also can participate in plant growth such as seed development, and improve plant resistance to abiotic stresses including heat stress, cold stress, salt tolerance and oxidative stress [56–58]. However, only fewer function studies involved in lncRNAs have been analyzed in fungi [20–22, 24]. In previous researches, some lncRNAs have been identified in filamentous pathogenic fungi, but the biological function of lncRNAs has not been explored in depth [46, 59, 60].

As far as know, there are not reports on the identification and relative function of lncRNAs in *B. cinerea*. In this research, we profiled strand-specific RNA-Seq data of vegetative and infection stages of *B. cinerea*, and a total

of 1837 lncRNAs were identified. Among these lncRNAs, a gradually increasing number of differentially expressed lncRNAs were discovered with a time course following infection of the host. It is worth noting that down-regulated lncRNAs accounted for the major proportion of differentially expressed lncRNAs in *B. cinerea* during the infection stage, which indicated these differentially expressed lncRNAs may participate in pathogenic process during the host infection of *B. cinerea*. Furthermore, antisense lncRNA has received an increasing attention due to its association with coding gene [61–63]. In *B. cinerea*, more than 70% of differentially expressed lncRNAs were antisense lncRNAs, which implied that antisense lncRNA may participate in infection stage through the regulation of the expression of its sense gene in *B. cinerea*.

In addition, we further discovered a strongly positive correlation ($r \geq 0.7$) between the differentially antisense lncRNA and sense mRNA. These antisense lncRNAs (37%) were induced in parallel with their target genes on the opposite DNA strand during infection. Antisense lncRNAs can participate in a wide range of cellular processes through target sense genes [64–66]. This correlation has also been found in other plant pathogenic fungi. In barley powdery mildew fungus *Blumeria hordei*, there is extensive positive and negative co-regulation of lncRNAs, transposable elements and coding genes during the asexual pathogenic life cycle of the fungis [67]. In our current research, the target genes that regulated by antisense lncRNAs were significantly enriched in GO terms of iron ion, including siderophore and ferrichrome processes. Iron is an indispensable element for all eukaryotes. In fungi, siderophores are functional for iron uptake, and ferrichrome contributes to iron storage [68, 69]. Many studies showed that the majority of fungi have the ability to produce siderophores, which play a crucial role in determining the virulence of pathogenic fungi, such as *Aspergillus fumigatus* and *Magnaporthe oryzae* [50, 51]. In *B. cinerea*, sense genes have a positive correlation with differentially expressed antisense lncRNAs that are involved in the siderophore and ferrichrome processes, which indicated the antisense lncRNAs may regulate iron absorption, transport and metabolism processes by targeting specific genes during the stages of vegetative and infection.

Antisense lncRNA has been reported to usually exert their biological functions through multiple mechanisms including transcriptional regulation, chromatin shape, epigenetics regulation, competition for endogenous RNA and miRNAs [70]. Heterochromatin formation-induced gene silencing requires the generation of sRNAs (sometimes from antisense lncRNA) through a co-transcriptional process [71]. However, in this study, the sRNA-seq data showed that reads mapped to region of antisense lncRNA were not enriched in *B. cinerea*, implying the antisense lncRNA was not preferentially target by RNAi machinery.

During the alternative splicing, some exons can be retained or excluded, resulting in different mature mRNAs generated from the same pre-mRNA. Targeted RNA sequencing revealed that lncRNA transcript consist of exons and introns and that lncRNAs can also occurred alternative splicing [72]. For example, lncRNA GAS5 (Growth-Arrest-Specific) has fifteen transcript isoforms in mice according to RefSeq [55]. Like protein-coding mRNAs, these different isoform of lncRNA GAS5 have distinct cellular localizations and involve in diverse pathologic functions in human diseases [73–76]. To further complicate matters, different lncRNA isoforms perform different functions when localized in same subcellular compartments. *lncRNA-PXN-AS1* affects the *PXN* expression at different levels through distinct lncRNA isoforms, thereby promoting the occurrence of liver cancer in human [77]. In *B. cinerea*, we also found that lncRNAs could perform alternative splicing, and a total of 233 alternative splicing events from 123 lncRNAs were identified. In fact, the biological significance of lncRNA has been overlooked in numerical studies, even in those involving pathogenic fungi. Thus, the studies on the functional verification of alternative splicing-related lncRNAs will be further investigated in future work.

Collectively, this study presents the first to report on the genome-wide characterization of lncRNAs during vegetative and infection stage in *B. cinerea*. The newly identified lncRNAs provide fundamental genomic resource to the *B. cinerea*. Our results also provide new insights into the functional understanding of lncRNA during infection stage of *B. cinerea*.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-11171-8.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We thank the Zhubiao lab in Zhejiang A&F University for providing tomato seeds.

Author contributions

SHJ and LP designed the experiments. LP, DGU, WY, and WJQ prepared samples and performed the experiments. LP and WD performed the analysis and pipelines. LP wrote the initial draft. WXL, WD and SHJ wrote, reviewed, and edited the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (32302305) and the Development Fund Project of Zhejiang A&F University (2022LFR092).

Data availability
Sequence data that support the findings of this study have been deposited in the NCBI Sequence Read Archive (SRA) with the primary accession code PRJNA1056687.

Declarations

Ethics approval and consent to participate
We ensure the collection of the tomato seed used in our study complied with local and national guidelines.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 14 May 2024 / Accepted: 19 December 2024
Published online: 06 January 2025

References

1. Mattick JS, Amaral PP, Carninci P, Carpenter S, Chang HY, Chen LL, Chen R, Dean C, Dinger ME, Fitzgerald KA, et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat Rev Mol Cell Biol. 2023;24(6):430–47.
2. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermuller J, Hofacker IL, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science. 2007;316(5830):1484–8.
3. Mukherjee N, Calviello L, Hirsekorn A, de Pretis S, Pelizzola M, Ohler U. Integrative classification of human coding and noncoding genes through RNA metabolism profiles. Nat Struct Mol Biol. 2017;24(1):86–96.
4. Ulitsky I. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. Nat Rev Genet. 2016;17(10):601–14.
5. Zuckerman B, Ulitsky I. Predictive models of subcellular localization of long RNAs. RNA. 2019;25(5):557–72.
6. Ahn JH, Lee HS, Lee JS, Lee YS, Park JL, Kim SY, Hwang JA, Kunkeaw N, Jung SY, Kim TJ, et al. nc886 is induced by TGF-beta and suppresses the microRNA pathway in ovarian cancer. Nat Commun. 2018;9(1):1166.
7. Postepska-Igielska A, Giwojna A, Gasri-Plotnitsky L, Schmitt N, Dold A, Ginsberg D, Grummt I. LncRNA *Khps1* regulates expression of the proto-oncogene *SPHK1* via triplex-mediated changes in chromatin structure. Mol Cell. 2015;60(4):626–36.
8. Zealy RW, Fomin M, Davila S, Makowsky D, Thigpen H, McDowell CH, Cummings JC, Lee ES, Kwon SH, Min KW, et al. Long noncoding RNA complementarity and target transcripts abundance. Biochim Biophys Acta Gene Regul Mech. 2018;1861(3):224–34.
9. Schmidt K, Weidmann CA, Hilimire TA, Yee E, Hatfield BM, Schneekloth JS Jr., Weeks KM, Novina CD. Targeting the oncogenic long non-coding RNA *SLNCR1* by blocking its sequence-specific binding to the androgen receptor. Cell Rep. 2020;30(2):541–54.
10. Yin Y, Yan P, Lu J, Song G, Zhu Y, Li Z, Zhao Y, Shen B, Huang X, Zhu H, et al. Opposing roles for the lncRNA *haunt* and its genomic locus in regulating *HOXA* Gene activation during embryonic stem cell differentiation. Cell Stem Cell. 2015;16(5):504–16.
11. Ali T, Grote P. Beyond the RNA-dependent function of LncRNA genes. Elife. 2020;9:e60583.
12. Cai BL, Li ZH, Ma MT, Wang ZJ, Han PG, Abdalla BA, Nie QH, Zhang XQ. LncRNA-Six1 encodes a micropeptide to activate in Six1 in cis and is involved in cell proliferation and muscle growth. Front Physiol. 2017;8:203.
13. Guo B, Wu S, Zhu X, Zhang L, Deng J, Li F, Wang Y, Zhang S, Wu R, Lu J, et al. Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. Embo j. 2020;39(1):e102190.
14. Matsumoto A, Pasut A, Matsumoto M, Yamashita R, Fung J, Monteleone E, Saghatelian A, Nakayama KI, Clohessy JG, Pandolfi PP. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. Nature. 2017;541(7636):228–32.
15. Bhan A, Soleimani M, Mandal SS. Long noncoding RNA and cancer: a new paradigm. Cancer Res. 2017;77(15):3965–81.
16. Liu G, Liu F, Wang Y, Liu X. A novel long noncoding RNA *CIL1* enhances cold stress tolerance in *Arabidopsis*. Plant Sci. 2022;323:111370.
17. Qin T, Zhao H, Cui P, Albesher N, Xiong L. A nucleus-localized long non-coding RNA enhances drought and salt stress tolerance. Plant Physiol. 2017;175(3):1321–36.
18. Seo JS, Diloknawarit P, Park BS, Chua NH. ELF18-INDUCED LONG NONCODING RNA 1 evicts fibrillarin from mediator subunit to enhance *PATHOGENESIS-RELATED GENE 1 (PR1)* expression. New Phytol. 2019;221(4):2067–79.
19. Zhang L, Wang M, Li N, Wang H, Qiu P, Pei L, Xu Z, Wang T, Gao E, Liu J, et al. Long noncoding RNAs involve in resistance to *Verticillium Dahliae*, a fungal disease in cotton. Plant Biotechnol J. 2018;16(6):1172–85.
20. Hiriart E, Verdel A. Long noncoding RNA-based chromatin control of germ cell differentiation: a yeast perspective. Chromosome Res. 2013;21(6–7):653–63.
21. Yamashita A, Shichino Y, Tanaka H, Hiriart E, Touat-Todeschini L, Vavasseur A, Ding DQ, Hiraoka Y, Verdel A, Yamamoto M. Hexanucleotide motifs mediate recruitment of the RNA elimination machinery to silent meiotic genes. Open Biol. 2012;2(3):120014.
22. Ding DQ, Okamasa K, Yamane M, Tsutsumi C, Haraguchi T, Yamamoto M, Hiraoka Y. Meiosis-specific noncoding RNA mediates robust pairing of homologous chromosomes in meiosis. Science. 2012;336(6082):732–6.
23. Wang J, Zeng W, Cheng J, Xie J, Fu Y, Jiang D, Lin Y. *lncRsp1*, a long noncoding RNA, influences *Fgsp1* expression and sexual reproduction in *FusGraminearumnearum*. Mol Plant Pathol. 2022;23(2):265–77.
24. Huang P, Yu X, Liu H, Ding M, Wang Z, Xu JR, Jiang C. Regulation of TRI5 expression and deoxynivalenol biosynthesis by a long non-coding RNA in *Fusarium graminearu*m. Nat Commun. 2024;15(1):1216.
25. van Kan JA. Licensed to kill: the lifestyle of a necrotrophic plant pathogen. Trends Plant Sci. 2006;11(5):247–53.
26. Veloso J, van Kan JAL. Many shades of grey in *Botrytis*-host plant interactions. Trends Plant Sci. 2018;23(7):613–22.
27. Van Kan JA, Stassen JH, Mosbach A, Van Der Lee TA, Faino L, Farmer AD, Papasotiriou DG, Zhou S, Seidl MF, Cottam E, et al. A gapless genome sequence of the fungus *Botrytis cinerea*. Mol Plant Pathol. 2017;18(1):75–89.
28. Bi K, Scalschi L, Jaiswal N, Mengiste T, Fried R, Sanz AB, Arroyo J, Zhu W, Masrati G, Sharon A. The *Botrytis cinerea* Crh1 transglycosylase is a cytoplasmic effector triggering plant cell death and defense response. Nat Commun. 2021;12(1):2166.
29. Hou J, Feng HQ, Chang HW, Liu Y, Li GH, Yang S, Sun CH, Zhang MZ, Yuan Y, Sun J, et al. The H3K4 demethylase Jar1 orchestrates ROS production and expression of pathogenesis-related genes to facilitate *Botrytis cinerea* virulence. New Phytol. 2020;225(2):930–47.
30. Zhu W, Yu M, Xu R, Bi K, Yu S, Xiong C, Liu Z, Sharon A, Jiang D, Wu M, et al. *Botrytis cinerea* BcSSP2 protein is a late infection phase, cytotoxic effector. Environ Microbiol. 2022;24(8):3420–35.
31. Bi K, Liang Y, Mengiste T, Sharon A. Killing softly: a roadmap of pathogenicity. Trends Plant Sci. 2023;28(2):211–22.
32. van Kan JA, van't Klooster JW, Wagemakers CA, Dees DC, van der Vlugt-Bergmans CJ. Cutinase A of *Botrytis cinerea* is expressed, but not essential, during penetration of gerbera and tomato. Mol Plant Microbe Interact. 1997;10(1):30–8.
33. Patel RK, Jain M. NGS QC toolkit: a toolkit for quality control of next generation sequencing data. PLoS ONE. 2012;7(2):e30619.
34. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.
35. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016;11(9):1650–67.
36. Pertea G, Pertea M. GFF Utilities: GffRead and GffCompare. F1000Res. 2020;9.
37. Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res. 2013;41(6):e74.
38. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417–9.
39. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
40. Ballester M, Cordón R, Folch JM. DAG expression: high-throughput gene expression analysis of real-time PCR data using standard curves for relative quantification. PLoS ONE. 2013;8(11).

41. Qin S, Veloso J, Baak M, Boogmans B, Bosman T, Puccetti G, Shi-Kunne X, Smit S, Grant-Downton R, Leisen T, et al. Molecular characterization reveals no functional evidence for naturally occurring cross-kingdom RNA interference in the early stages of *Botrytis cinerea*-tomato interaction. Mol Plant Pathol. 2023;24(1):3–15.
42. Martin MJEJ. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17(1):10–2.
43. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.
44. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.
45. Trincado JL, Entizne JC, Hysenaj G, Singh B, Skalic M, Elliott DJ, Eyras E. SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. Genome Biol. 2018;19(1):40.
46. Choi G, Jeon J, Lee H, Zhou S, Lee YH. Genome-wide profiling of long non-coding RNA of the rice blast fungus *Magnaporthe oryzae* during infection. BMC Genomics. 2022;23(1):132.
47. Wunderlich M, Gross-Hardt R, Schoffl F. 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. 2014;85(6):541–50.
48. Wang J, Zeng W, Xie J, Fu Y, Jiang D, Lin Y, Chen W, Cheng J. A novel antisense long non-coding RNA participates in asexual and sexual reproduction by regulating the expression of *GzmetE* in *Fusarium graminearu*m. Environ Microbiol. 2021;23(9):4939–55.
49. Latos PA, Pauler FM, Koerner MV, Senergin HB, Hudson QJ, Stocsits RR, Allhoff W, Stricker SH, Klement RM, Warczok KE, et al. *Airn* transcriptional overlap, but not its lncRNA products, induces imprinted *Igf2r* silencing. Science. 2012;338(6113):1469–72.
50. Haas H, Eisendle M, Turgeon BG. Siderophores in fungal physiology and virulence. Annu Rev Phytopathol. 2008;46:149–87.
51. Schrettl M, Bignell E, Kragl C, Sabiha Y, Loss O, Eisendle M, Wallner A, Arst HN Jr., Haynes K, Haas H. Distinct roles for intra- and extracellular siderophores during *Aspergillus fumigatus* infection. PLoS Pathog. 2007;3(9):1195–207.
52. Makalowska I, Lin CF, Makalowski W. Overlapping genes in vertebrate genomes. Comput Biol Chem. 2005;29(1):1–12.
53. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, et al. The landscape of long noncoding RNAs in the human transcriptome. Nat Genet. 2015;47(3):199–208.
54. Kim E, Goren A, Ast G. Alternative splicing: current perspectives. BioEssays. 2008;30(1):38–47.
55. Bridges MC, Daulagala AC, Kourtidis A. LNCcation: lncRNA localization and function. J Cell Biol. 2021;220(2):e202009045.
56. Liu GC, Liu FX, Wang Y, Liu X. A novel long noncoding RNA *CIL1* enhances cold stress tolerance in *Arabidopsis*. Plant Sci. 2022;323:111370.
57. Qin T, Zhao HY, Cui P, Albesher N, Xiong LM. A nucleus-localized long non-coding RNA enhances drought and salt stress tolerance. Plant Physiol. 2017;175(3):1321–36.
58. Wunderlich M, Gross-Hardt R, Schöffl F. 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. 2014;85(6):541–50.
59. Kim W, Miguel-Rojas C, Wang J, Townsend JP, Trail F. Developmental dynamics of long noncoding RNA expression during sexual fruiting body formation in *Fusarium graminearum*. mBio. 2018;9(4):e01292–18.
60. Li R, Xue HS, Zhang DD, Wang D, Song J, Subbarao KV, Klosterman SJ, Chen JY, Dai XF. Identification of long non-coding RNAs in *Verticillium dahliae* following inoculation of cotton. Microbiol Res. 2022;257:126962.
61. Wang J, Zeng WP, Xie JT, Fu YP, Jiang DH, Lin Y, Chen WD, Cheng JS. A novel antisense long non-coding RNA participates in asexual and sexual reproduction by regulating the expression of of GzmetE in *Fusarium graminearum*. Environ Microbiol. 2022;24(10):4967–70.
62. Jiang BY, Yuan YQ, Yi T, Dang W. The roles of antisense long noncoding RNAs in tumorigenesis and development through cis-regulation of neighbouring genes. Biomolecules. 2023;13(4):684.
63. Mosca N, Russo A, Potenza N. Making sense of antisense lncRNAs in hepatocellular carcinoma. Int J Mol Sci. 2023;24(10):8886.
64. Canzio D, Nwakeze CL, Horta A, Rajkumar SM, Coffey EL, Duffy EE, Duffie R, Monahan K, O'Keeffe S, Simon MD, et al. Antisense lncRNA transcription mediates DNA demethylation to drive stochastic protocadherin alpha promoter choice. Cell. 2019;177(3):639–53.
65. Lennox KA, Behlke MA. Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. Nucleic Acids Res. 2016;44(2):863–77.
66. Romero-Barrios N, Legascue MF, Benhamed M, Ariel F, Crespi M. Splicing regulation by long noncoding RNAs. Nucleic Acids Res. 2018;46(5):2169–84.
67. Qian J, Ibrahim HMM, Erz M, Kümmel F, Panstruga R, Kusch S. Long noncoding RNAs emerge from transposon-derived antisense sequences and may contribute to infection stage-specific transposon regulation in a fungal phytopathogen. Mob DNA. 2023;14(1):17.
68. Schrettl M, Bignell E, Kragl C, Joechl C, Rogers T, Arst HN Jr., Haynes K, Haas H. Siderophore biosynthesis but not reductive iron assimilation is essential for *Aspergillus fumigatus* virulence. J Exp Med. 2004;200(9):1213–9.
69. Wallner A, Blatzer M, Schrettl M, Sarg B, Lindner H, Haas H. Ferricrocin, a siderophore involved in intra- and transcellular iron distribution in *Aspergillus fumigatus*. Appl Environ Microbiol. 2009;75(12):4194–6.
70. Rehman SU, Ullah N, Zhang ZB, Zhen YK, Din AU, Cui HM, Wang MZ. Recent insights into the functions and mechanisms of antisense RNA: emerging applications in cancer therapy and precision medicine. Front Chem. 2024;11:1335330.
71. Shehzada S, Noto T, Saksouk J, Mochizuki K. A SUMO E3 ligase promotes long non-coding RNA transcription to regulate small RNA-directed DNA elimination. Elife. 2024;13:e95337.
72. Mercer TR, Gerhardt DJ, Dinger ME, Crawford J, Trapnell C, Jeddeloh JA, Mattick JS, Rinn JL. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. Nat Biotechnol. 2012;30(1):99–147.
73. Goustin AS, Thepsuwan P, Kosir MA, Lipovich L. The growth-arrest-specific (*GAS*)-5 long non-coding RNA: a fascinating lncRNA widely expressed in cancers. Noncoding RNA. 2019;5(3):46.
74. Meng XD, Yao HH, Wang LM, Yu M, Shi S, Yuan ZX, Liu J. Knockdown of GAS5 inhibits atherosclerosis progression via reducing EZH2-mediated ABCA1 transcription in *ApoE(-/-)* mice. Mol Ther Nucleic Acids. 2020;19:84–96.
75. Renganathan A, Kresoja-Rakic J, Echeverry N, Ziltener G, Vrugt B, Opitz I, Stahel RA, Felley-Bosco E. GAS5 long non-coding RNA in malignant pleural mesothelioma. Mol Cancer. 2014;13:119.
76. Yu F, Zheng J, Mao Y, Dong P, Lu Z, Li G, Guo C, Liu Z, Fan X. Long non-coding rna growth arrest-specific transcript 5 (GAS5) inhibits liver fibrogenesis through a mechanism of competing endogenous RNA. J Biol Chem. 2015;290(47):28286–98.
77. Yuan JH, Liu XN, Wang TT, Pan W, Tao QF, Zhou WP, Wang F, Sun SH. The MBNL3 splicing factor promotes hepatocellular carcinoma by increasing PXN expression through the alternative splicing of lncRNA-PXN-AS1. Nat Cell Biol. 2017;19(7):820–32.

Publisher's note