

RESEARCH ARTICLE

Genome Wide Identification and Functional Prediction of Long Non-Coding RNAs Responsive to *Sclerotinia sclerotiorum* Infection in *Brassica napus*

Raj Kumar Joshi^{1,2}, Swati Megha¹, Urmila Basu¹, Muhammad H. Rahman¹, Nat N. V. Kav^{1*}

1 Department of Agricultural Food and Nutritional Science, University of Alberta, Edmonton, Alberta, T6G 2P5, Canada, **2** Centre of Biotechnology, Siksha O Anusandhan University, Bhubaneswar-751003, India

* nat@ualberta.ca



OPEN ACCESS

Citation: Joshi RK, Megha S, Basu U, Rahman MH, Kav NNV (2016) Genome Wide Identification and Functional Prediction of Long Non-Coding RNAs Responsive to *Sclerotinia sclerotiorum* Infection in *Brassica napus*. PLoS ONE 11(7): e0158784. doi:10.1371/journal.pone.0158784

Editor: Bin Shan, Washington State University, UNITED STATES

Received: February 26, 2016

Accepted: June 22, 2016

Published: July 7, 2016

Copyright: © 2016 Joshi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The RNA-Seq reads have been deposited at the sequence read archive (SRA) of NCBI (accession numbers SRX956927 to SRX956938).

Funding: RKJ is grateful to Department of Biotechnology, Government of India for providing DBT-CREST award (Sanction letter no. BT/IN/DBT-CREST Awards/32/RKJ/2012-13) to pursue research as a Visiting Scientist at the University of Alberta. Work in NNVK's laboratory was supported in part by a Discovery Grant from the Natural Sciences and

Abstract

Sclerotinia stem rot caused by *Sclerotinia sclerotiorum* affects canola production worldwide. Emerging evidence suggests that long non-coding RNAs (lncRNAs) play important roles in the regulation of gene expression in plants, in response to both abiotic and biotic stress. So far, identification of lncRNAs has been limited to a few model plant species, and their roles in mediating responses to biotic stresses are yet to be characterized in *Brassica napus*. The present study reports the identification of novel lncRNAs responsive to *S. sclerotiorum* infection in *B. napus* at two time points after infection (24 hpi and 48 hpi) using a stranded RNA-Sequencing technique and a detection pipeline for lncRNAs. Of the total 3,181 lncRNA candidates, 2,821 lncRNAs were intergenic, 111 were natural antisense transcripts, 76 possessed exonic overlap with the reference coding transcripts while the remaining 173 represented novel lnc-isoforms. Forty one lncRNAs were identified as the precursors for microRNAs (miRNAs) including miR156, miR169 and miR394, with significant roles in mediating plant responses to fungal phytopathogens. A total of 931 differentially expressed lncRNAs were identified in response to *S. sclerotiorum* infection and the expression of 12 such lncRNAs was further validated using qRT-PCR. *B. napus* antisense lncRNA, TCONS_00000966, having 90% overlap with a plant defensin gene, showed significant induction at both infection stages, suggesting its involvement in the transcriptional regulation of defense responsive genes under *S. sclerotiorum* infection. Additionally, nine lncRNAs showed overlap with cis-regulatory regions of differentially expressed genes of *B. napus*. Quantitative RT-PCR verification of a set of *S. sclerotiorum* responsive sense/antisense transcript pairs revealed contrasting expression patterns, supporting the hypothesis that steric clashes of transcriptional machinery may lead to inactivation of sense promoter. Our findings highlight the potential contributions of lncRNAs in regulating expression of plant genes that respond to biotic stress.

Engineering Research Council (NSERC) of Canada awarded to NNVK.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Whole genome RNA sequencing (RNA-Seq), tilling arrays and large-scale cDNA cloning studies have revealed that transcription of eukaryotic genes is highly complex [1–2]. Although 90% of the eukaryotic genome is transcribed, only 1–2% has the capacity to encode proteins, suggesting that a large fraction of eukaryotic transcripts represent non-coding RNAs (ncRNAs) [3–4]. The specific functions of most of the ncRNAs has not yet been fully characterized, but recent evidence from plants and animals supports a crucial role of these ncRNAs in the regulation of a wide range of biological processes including plant responses to biotic and abiotic stresses [5–8]. Based on their length, ncRNAs are randomly grouped into short (< 200 nucleotides) and lncRNAs (> 200 nucleotides) [6]. The importance of small ncRNAs including microRNAs (miRNAs), short interfering RNAs (siRNAs), natural antisense siRNAs (nat-siRNAs), piwi interacting RNAs (piRNAs) and trans acting siRNAs (tasi-RNAs) in transcriptional and post-transcriptional regulation of gene expression in plants have been well documented [9–11]. Although, the molecular mechanisms governing the lncRNA-mediated gene regulation is still poorly understood, increasing number of studies suggest their involvement in the regulation of biological processes in plants [12–14].

lncRNAs are broadly classified as long intronic ncRNAs, long intergenic ncRNAs or natural antisense transcripts (NATs) based on their location [14]. They can be in either sense or antisense orientation in relation to the neighbouring protein-coding transcript and may overlap with exons or introns of protein coding genes [7]. However, systematic identification and classification of lncRNAs have been limited to a few model plant species such as *Arabidopsis thaliana*, rice and maize. In *A. thaliana*, 6,480 lncRNAs and 37,238 NATs have been identified and classified using genome wide tilling array and RNA-Seq [15]. Similarly, in rice, 3,819 cis-NATs have been identified, among which 503 cis-NATs seem to be associated with specific biological functions such as growth and abiotic stress tolerance [16]. An exploitation of the expressed sequence tags (EST) databases, whole genome sequence annotation and RNA-Seq datasets from 30 different experiments led to the identification of 20,163 lncRNAs in maize [17]. Both strand-specific and non-directional sequencing in rice and maize identified 22,334 lncRNAs and 6,673 pairs of NATs and genome-wide association studies highlighted their involvement in developmental agricultural traits [18]. In another study, 8,449 drought responsive transcripts were analyzed with the coding potential calculator to identify 664 drought responsive lncRNAs in maize [19]. More recently, 2,542 lncRNA candidates have been identified from *Populus trichocarpa*, 504 of which were found to be drought responsive [20]. Interestingly, a number of lncRNAs serve as precursors for small RNAs including small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and miRNAs to regulate their functions [21]. Moreover, a number of lncRNAs act as endogenous target mimics for miRNAs and inhibit their functions, thus playing a significant role in the regulation of plant growth and development [20–22].

Although a large number of lncRNAs have been identified in model plants, only a few of them have been functionally characterized. For instance, *Enod40*, a lncRNA from *Medicago truncatula* acts as molecular cargo for *M. truncatula* RNA binding protein 1 (MtrRBP1) by relocalizing the protein from nuclear speckles into cytoplasmic granules during nodulation [23]. Induced by Phosphate Starvation 1 (*IPSI*) lncRNA from *A. thaliana* modulates phosphate uptake by acting as a target mimic of miR399 and inhibiting its cleavage activity [24]. Two lncRNAs, COOLAIR and COLDAIR, transcribed from the *FLOWERING LOCUS C (FLC)* gene of *A. thaliana* play important roles in the regulation of cold stress response [12–13]. COOLAIR is induced during vernalization and transcribed in an antisense orientation to *FLC* but is not essential for *FLC* repression [12]. In contrast, COLDAIR is transcribed from the first

intron in the sense orientation relative to *FLC* and interacts with CURLY LEAF (CLF) of polycomb receptor complex 2 (PRC2) during vernalization [13]. It has been observed that the vernalization process is compromised in COLDAIR knockdown lines, suggesting its involvement in the regulation of *FLC* expression [13]. Additional lncRNAs such as *CsM10* from cucumber [25] and *zm401* from maize [26] have been found to be expressed in developing male gametophyte. A couple of recent studies have demonstrated an involvement of lncRNAs in biotic stress responses. For instance, 125 lncRNAs have been identified from wheat, which are responsive to powdery mildew infection [27]. These lncRNAs were not conserved among plant species and exhibited high tissue specific expression. In addition, 159 novel intergenic transcriptionally active regions (TARs) responsive to *Fusarium oxysporum* were characterized from *A. thaliana* [28]. Functional characterization of lncTARs through RNA-interference knockdown demonstrated direct involvement of five lncRNAs in disease development [28]. All of these reports provide evidence in support of the prominent roles of lncRNAs in the regulation of plant growth, development and stress responses.

S. sclerotiorum is a necrotrophic, non-host specific phytopathogen with over 400 different potential hosts [29]. In a wide range of host plants including the agriculturally important oil seed crop canola (*B. napus*), the fungus causes Sclerotinia stem rot [30]. Substantial yield losses from the outbreaks of Sclerotinia stem rot have been reported worldwide including in Australia, Europe and North America [31–32]. Comprehensive genome and proteome analyses of defense signal transduction pathways of both the host plant and the pathogen have provided detailed insights about the molecular mechanisms underlying disease progression in the *B. napus*-*S. sclerotiorum* pathosystem [33–35]. However, the roles of lncRNAs in the regulation of antifungal defense network in *B. napus*, specifically in response to *S. sclerotiorum*, are still unknown. In the present study, we used a high throughput stranded RNA-Seq based approach, to elucidate lncRNA expression profiles in *B. napus* subsequent to infection with *S. sclerotiorum*. We have identified numerous *S. sclerotiorum* responsive lncRNAs which can be classified into different categories and may play key roles during the *B. napus*-*S. sclerotiorum* interaction. To our knowledge, this is the first report identifying the differentially expressed (DE) lncRNAs in this pathosystem and our findings may have long-term implications for the development of novel disease resistant canola lines.

Materials and Methods

Plant material, pathogen inoculation and RNA sequencing

A strain of *S. sclerotiorum* was kindly provided by Dr. Stephen Strelkov (University of Alberta) and seeds of *B. napus* cv. DH12075 were kindly provided by Dr. Gerhard Rakow of Agriculture and Agri-Food Canada (AAFC). *B. napus* cv. DH12075 (susceptible to *Sclerotinia* stem rot) seeds were surface sterilized and sown in plastic trays containing Metro Mix 290 (Grace Horticultural products, Ajax, ON, Canada). Plants were grown in the greenhouse with a 16 h photoperiod (light intensity of $130 \mu\text{mol m}^{-2} \text{s}^{-1}$) with a day ($21 \pm 1^\circ\text{C}$)/ night ($18 \pm 1^\circ\text{C}$) temperature cycle for 20 days and transferred to a humidity chamber 24 h prior to inoculation. Mature sclerotia of a wild-type *S. sclerotiorum* collected from an infected canola plant was used as the source of infection. Sclerotia were sub-cultured on potato dextrose agar medium (PDA; Becton Dickinson, Columbia, MD) and incubated at room temperature (RT; $21 \pm 2^\circ\text{C}$) for 3 days. Agar plugs (0.2 cm) were removed with a sterile cork borer from the leading edge of the mycelia and were sub-cultured on PDA agar plate, grown for an additional 3 days at RT and used for inoculation of plants. Mycelial plugs (5 mm) excised from the culture plates with actively growing fungal cultures were placed on the first and second true leaves of *B. napus* which were slightly wounded prior to exposure to the pathogen. Leaves of un-inoculated plants were wounded and

inoculated with sterile PDA plugs excised from plates without fungal cultures and will be referred to as control plants. Pathogen inoculated and control plants were placed in a humidity chamber for 24 h, after which they were returned to the greenhouse. The entire leaf samples (both first and second true leaves) from control and inoculated plants were harvested across all replicates and flash frozen in liquid nitrogen for RNA isolation at 24 and 48 hours post-infection (hpi).

Total RNA was extracted from *B. napus* leaves using TRIzol reagent (Invitrogen, USA) and treated with DNase (Promega, Madison, WI) according to manufacturer's instructions. The quality and quantity of total RNA was determined using NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, USA). Integrity of isolated RNA was determined on Agilent 2100 Bioanalyzer using Plant RNA Nano chip assay in accordance with manufacturer's instructions (Agilent Technologies, Santa Clara, CA). RNA samples with 260/280nm ratio between 2.0 to 2.1, and RNA integrity number (RIN) greater than 8.0, were used for the analysis. Total RNA from 12 samples, six un-inoculated (three each for 24 hpi and 48 hpi) and six *S. sclerotiorum* infected (three each for 24 hpi and 48 hpi) were used to construct RNA-Seq libraries using the TruSeq stranded total RNA preparation kit with RiboZero plant (Illumina, San Diego CA) according to the manufacturer's instructions. The sequencing of libraries was performed on Illumina HiSeq 2500 platform at the Beijing Genomics Institute (BGI Americas, San Diego, CA).

Assembling RNA transcripts and lncRNA detection

The basic approach used for transcript assembly and detection of lncRNAs is represented in Fig 1. Raw reads in the FASTAQ files were processed to remove adaptors, low-quality tags and contaminants. To remove the remaining rRNA reads, clean reads were mapped to the ribosomal RNA (rRNA) reference using soft read alignment software SOAPAligner/SOAP2 [36]. *B. napus* (AACC, 2n = 38) is an amphidiploid species which originated from interspecies crosses between *B. rapa* (AA, 2n = 20) and *B. oleracea* (CC, 2n = 18). As the complete sequence of the *B. napus* genome was not available at the time of this study, filtered reads were mapped to *B. rapa* (AA genome) [37] (<http://www.brassicadb.org>) and *B. oleracea* (CC genome) [38] (<http://www.ocri-genomics.org/bolbase/>) genomes respectively, using TopHat2 [39]. Also, the sequence reads were mapped to the *S. sclerotiorum* genome [40] (<http://www.genome.jgi.doe.gov/Scslsc1/>) and the mapped reads were removed from further analysis. Subsequently, the reads mapped onto the *B. rapa* (AA) and *B. oleracea* (CC) genomes were assembled by Cufflinks [41]. A reference Annotation Based Transcripts (RABT) assembly [42] was generated with the reference gene annotation, to compensate for incompletely assembled transcripts caused by read coverage gaps in the regions of reference gene. Novel transcripts were detected and categorized into different categories according to their position compared with reference genes by utilizing Cuffcompare [41]. Low quality assemblies were filtered according to transcript length ≥ 200 nucleotides, transcripts with open reading frames (ORFs) ≤ 300 nucleotides and an optimum RPKM (reads per kilobase transcriptome per million mapped) (RPKM ≥ 2.0) value. Cuffmerge was used to merge the replicate assemblies and filter out transfrags [41]. Finally, novel transcripts merged from different assemblies were evaluated using coding potential calculator (CPC) [43] and iSeeRNA [44] to predict transcript with non-coding functions. Using iSeeRNA, the unannotated transcripts were categorized into candidate categories such as 'i', 'j', 'o', 'u' and 'x', which may contain novel lncRNAs. The 'i' category may include lncRNAs within the introns of known genes, the 'j' category may contain alternative lncRNAs isoforms of known genes, the 'u' category may include the intergenic lncRNAs, the 'o' category may contain lncRNAs with exonic overlap

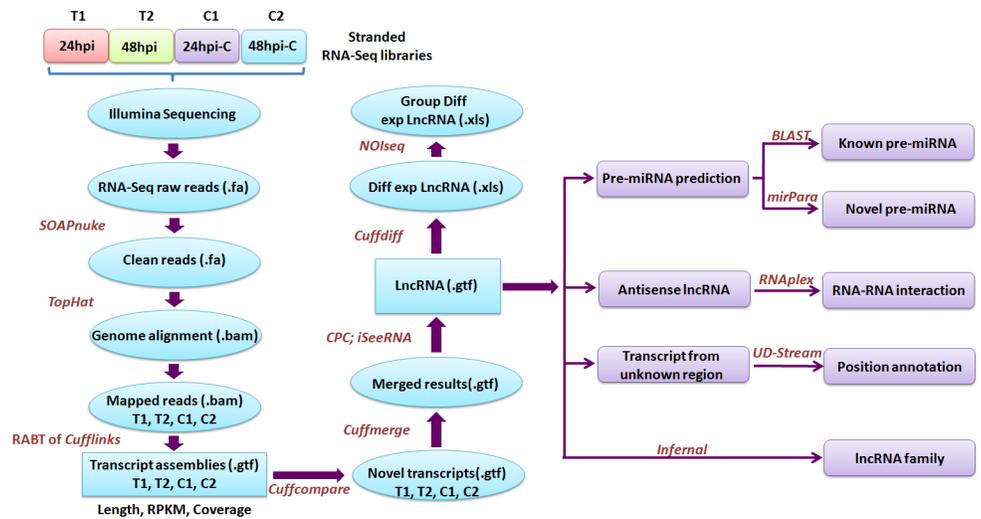


Fig 1. Pipeline for identification, annotation and classification of lncRNAs. Raw reads generated through stranded RNA-Seq were cleaned and mapped onto the reference genome of *B. napus*. Mapped reads were subsequently assembled by *Cufflinks*. A RABT assembly was performed with the reference gene annotation to compensate incompletely assembled transcripts. Novel set of transcripts were detected and categorized by utilizing *Cuffcompare*. *Cuffmerge* was used to merge the replicate assemblies and filter out transfrags. Low quality assemblies were filtered based on transcript length ≥ 200 nucleotides, open reading frames (ORFs) ≤ 300 nucleotides and RPKM ≥ 2.0 . Finally, novel transcripts merged from different assemblies were evaluated using coding potential calculator (CPC) and *iSeeRNA* to identify transcripts with non-coding functions. *INFERNAL* was used to classify lncRNAs into various ncRNA families. miRNA precursors were identified using *miRPara*. Potential sense/antisense interactions were determined using *RNAplex*. Intergenic lncRNAs were annotated as upstream or downstream lncRNAs using *UD stream*.

doi:10.1371/journal.pone.0158784.g001

with a known transcript, while the ‘x’ category may include exonic lncRNAs present on the opposite strand.

Screening of differentially expressed lncRNAs

lncRNA sequence reads from control and pathogen infected libraries for 24 hpi and 48 hpi were normalized to FPKM (fragments per kilobase of transcript per million mapped reads) [41]. Differential expression test between the two comparisons was performed using *Cuffdiff* [41]. Fold changes for differentially expressed lncRNAs were calculated as: $\log_2 \text{ratio} = \log_2 (\text{FPKM}_{\text{infected}} / \text{FPKM}_{\text{control}})$. False discovery rate (FDR) was determined after Benjamini-Hochberg correction [45] to identify false positive and false negatives. lncRNAs were considered DE, significant and pathogen responsive when \log_2 ratios of pathogen-inoculated samples compared to control were ≥ 2 , and FDR adjusted *P* value (q-value) was ≤ 0.05 . Furthermore, the DE lncRNAs between two time points were screened using *NOISeq* [46] by developing noise distribution models and Pearson’s correlation coefficients were calculated to determine the correlation of the DE lncRNAs between control and infected samples.

lncRNA classification and functional prediction

Predicted lncRNAs were classified into different ncRNA families using *INFERNAL*, which categorizes ncRNAs and their conserved primary sequence and RNA secondary structure through the use of multiple sequence alignments (MSAs), consensus secondary structure annotation and covariance models (CMs) [47]. Potential miRNA precursors (pre-miRNAs) were predicted by subjecting the lncRNAs to BLAST search against the miRBase (www.mirbase.org) [48] and

selecting hits with sequence coverage more than 90%. Subsequently, a support vector machine (SVM) based software mirPara [49] was used to detect novel miRNAs. All antisense lncRNA-mRNA duplexes exhibiting complementary base pairing were screened using RNAplex [50] to determine the potential antisense lncRNA-mRNA interaction. Finally, lncRNAs belonging to "unknown region" in former analyses were either classified as upstream or downstream, in order to determine their involvement in transcriptional regulation of corresponding genes.

Quantitative real time polymerase chain reaction (qRT-PCR) analysis

Twelve randomly selected, pathogen responsive lncRNAs and six corresponding sense transcripts from *B. napus* were validated using qRT-PCR. Sequences of primers used in qRT-PCR analyses are provided in [S1 Table](#). All qRT-PCR analyses were performed on the StepOne Plus real time PCR system (Life Technologies, Burlington, ON, Canada) using FASTSYBR green mix from Kappa Biosystem (D Mark, Toronto, ON, Canada). RNA isolated from pathogen-inoculated and control samples were treated with RNase-free DNase I. Total RNA (1 µg) was reverse transcribed using random primers and SuperScript III reverse transcriptase (Invitrogen). Control reactions without reverse transcriptase were performed for each primer combination in order to ensure the absence of genomic DNA contamination. Cycle threshold (C_T) values were determined for each sample based on three biological replicates each with two technical replicates. Constitutively expressed *UBC9* (Ubiquitin conjugating enzyme 9) gene was used as the reference. Relative expression of the target transcripts was calculated using the $2^{-\Delta\Delta C_T}$ method [51].

Results and Discussion

B. napus disease symptoms in response to *S. sclerotiorum*

The progression of infection was monitored in the susceptible *B. napus* cv. DH12075 to correlate the host responses with pathogen invasion. Symptoms of leaf necrosis and pathogen invasion have been described in our previous study [52]. The results demonstrated that, the area of leaf necrosis and amount of fungal mycelia detectable in the host tissue increased from 24 hpi to 48 hpi and was consistent with the severity of the observed, macroscopic disease symptoms.

Deep sequencing and identification of lncRNAs

Strand specific sequencing reveals antisense transcription activities and assists in the identification of novel lncRNAs. Recently, this approach has been utilized by plant biologists for analyses of transcriptomes and also towards characterization of a wide range of ncRNAs in different plant species [16, 20, 28]. To identify the lncRNAs involved in mediating responses to *S. sclerotiorum* infection of *B. napus*, a strand specific Illumina sequencing of transcripts from canola leaves at 24hpi and 48hpi was performed. A total of 317 million clean reads were obtained from the 12 libraries ([S2 Table](#)) and approximately 204 million reads (64.56%) were successfully mapped on to the reference genomes (AA genome of *B. rapa* and CC genome of *B. oleracea*). The RNA-Seq reads have been deposited at the sequence read archive (SRA) of NCBI (accession numbers SRX956927 to SRX956938). Among transcripts, 62,071 multi-exon transcripts were assembled and 43,714 of the assembled transcripts were completely annotated, 393 transcripts were partially annotated while the remaining 17,964 transcripts ([S3 Table](#)) were not annotated and are potentially novel in nature. The assembly of 17,964 transcripts was further categorized into initial categories of lncRNAs ([Table 1](#)). Among the 17,964 sequences, 12,542 transcripts were classified with classcode 'u', 4,127 transcripts were classified in the 'j' category, 716 transcripts represented the 'x' category while the remaining 579 were grouped under the 'o'

Table 1. Categories of initial assemblies and classification of novel lncRNAs from *B. napus* post infection with *S. sclerotiorum*.

Class code	After initial assemblies		After final processing		Description
	Transcript no.	%	Transcript no.	%	
=	43,714	70.42	-	-	Complete match of intron chain
C	393	0.63	-	-	Contained by a reference transcript
J	4,127	6.64	173	5.43	Potentially novel isoform (Atleast one splice junction is shared with a reference transcript)
O	579	0.93	76	2.38	Generic exonic overlap with a reference transcript
U	12,542	20.20	2,821	88.68	Intergenic transcript
X	716	1.15	111	3.48	Exonic overlap with reference on the opposite strand
Total	62,071	100%	3,181	100%	

doi:10.1371/journal.pone.0158784.t001

category. Based on the FPKM ≥ 2.0 as cut off, we identified a total of 3,181 transcripts from the previous four categories as *B. napus* lncRNAs responsive to *S. sclerotiorum* (S4 Table) (Table 1). In total, 2,821 lncRNAs were assigned classcode ‘u’ as they were located in the intergenic regions. The remaining lncRNAs had an exonic overlap with known genes. For instance, 111 lncRNAs with classcode ‘x’ had exonic overlap with known genes in the opposite strand, while 76 lncRNAs with classcode ‘o’ had generic exonic overlap with the reference coding transcripts. Additionally, 173 lncRNAs categorized with classcode ‘j’ represented novel lnc isoforms with at least one splice junction shared with the reference genes (Table 1). Thus, we could predict and classify *B. napus* lncRNAs into four biotypes as per the categories laid down in the NONCODE ver. 4.0 database [53]. Earlier, this pipeline was applied to a mouse Krueppel-like factor 1 (*Klf1*) knockout single-end RNA-Seq dataset to identify 308 novel lncRNA candidates [54]. We could predict a higher number of lncRNAs in our current study using the same pipeline, perhaps due to the fact that paired end sequencing and strand specific sequencing of *B. napus* transcript improved the prediction accuracy of the lncRNA detection tool. This premise is also supported by a recent report where a genome wide analysis of lncRNAs using a strand specific paired-end RNA-Seq experiment identified 2,542 lncRNAs in *P. trichocarpa* [20]. These observations suggest that the use of high quality paired-end reads generated through strand specific deep sequencing can act in an unbiased way to capture maximum long non coding transcriptome.

Distribution of *B. napus* lncRNAs

Based on the above results, 3,181 novel *B. napus* lncRNAs were selected for further analysis. The size distribution of these lncRNAs ranged from 200 bases to 5,782 bases, with more than 90% lncRNAs ranging from 200 to 350 bases (Fig 2A). Characterization of the genomic structure revealed that 1.3% of *B. napus* lncRNAs have only one exon, 43% have two exons and 26% have three exons while the rest have more than three exons (Fig 2B). From the category of lncRNAs with multiple exons, 67 lncRNAs had more than eight exons. Size distribution of the exons showed that 82% of the lncRNA exons had sizes ranging from 100 to 200 bases (median 146 bases; t-test, P-value ≤ 0.00011) (Fig 2C). Similarly, *B. napus* lncRNAs possessed longer intronic regions with sizes ranging from 1500 to 2500 bases (median 2198 bp; t-test, P-value ≤ 0.00037) for 72% of the introns (Fig 2D). Gene structure analysis of human lncRNAs has shown that the lncRNA exons are of similar length to protein coding exons and introns are longer than those of the protein coding transcripts [55]. However, since the majority of human lncRNAs have two to three exons as in case of *B. napus* lncRNA in this study, the overall lncRNA transcripts are shorter than protein coding transcripts.

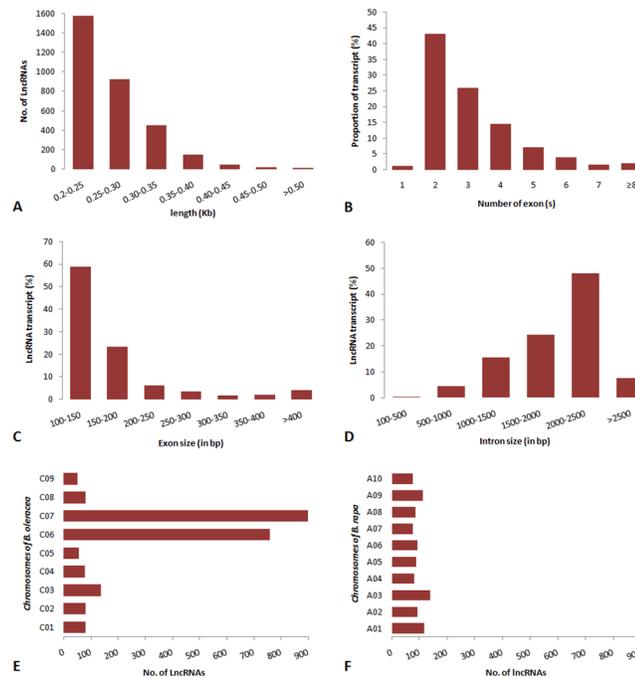


Fig 2. Features and distribution of lncRNA structures in *B. napus*. (A) Length distribution of 3,181 *B. napus* lncRNAs. (B) Number of exons per transcripts for all lncRNA transcripts. Size distribution of exons (C) and introns (D) for lncRNAs. Number of lncRNAs on each chromosome of the reference genomes of *B. oleracea* (E) and *B. rapa* (F).

doi:10.1371/journal.pone.0158784.g002

The availability of the AA (*B. rapa*) and CC (*B. oleracea*) genomes as references provides the possibility to determine the origin of *B. napus* lncRNAs and understand their evolution in *B. napus*. Among the identified lncRNAs, 2,200 and 961 lncRNAs were assigned positions across nine and ten chromosomes of CC (*B. oleracea*) and AA (*B. rapa*) genomes, respectively. This suggests that *B. oleracea* CC genome has contributed a larger number of lncRNAs to *B. napus*. Out of 2,220 lncRNAs in *B. oleracea*, 1,660 were located only on chromosome 6 (C06) and chromosome 7 (C07) and constituted the majority of stress responsive *B. napus* lncRNAs (Fig 2E). The remaining 560 lncRNAs were uniformly distributed amongst other chromosomes. Chromosome C09 (containing only 50 lncRNA loci) had the lowest lncRNA density. In the *B. rapa* genome, chromosome 3 (A03) had the highest number of lncRNA loci (138), followed by A01 (116), A09 (112) and A06 (94) (Fig 2F). In addition, 532 lncRNAs were conserved between both the genomes and were evenly distributed throughout the chromosomes suggesting that they are ancient and well-conserved. Moreover, the majority of intergenic lncRNAs were AA or CC genome specific suggesting that *B. napus* intergenic lncRNAs are probably more recent and may have evolved after speciation of *B. rapa* and *B. oleracea* separately or very recently after the speciation of *B. napus*. Future investigation into the genomic synteny analysis of the predicted lncRNAs between the AA and CC sub genomes will provide additional insights into the origin and evolution of *B. napus* lncRNAs.

B. napus lncRNAs differentially expressed in response to *S. sclerotiorum*

Recent reports have shown that many lncRNAs actively participate in the regulation of various stress responses including abiotic stresses and pathogen infection [56]. In *Arabidopsis*, 20 *F.*

oxysporum responsive novel lncRNAs were identified using a strand specific RNA sequencing, including five lncRNAs related to disease development [28]. Similarly, 125 putative lncRNAs responsive to powdery mildew infection and heat stress have been identified in wheat [27]. In the present study, the normalized expression pattern of *B. napus* lncRNA reads of control and inoculated samples were calculated using FPKM (fragment per Kilobase of transcript per million mapped reads) and compared at different time points to identify *S. sclerotiorum* responsive lncRNAs. Based on the statistical analysis of our results, 931 *B. napus* lncRNAs were significantly ($P \leq 0.05$) differentially expressed in response to *S. sclerotiorum* infection. Of these, 386 and 307 lncRNAs were induced in response to the pathogen and 276 and one lncRNA were repressed in the 24hpi and 48hpi samples, respectively (Fig 3A) (S5 Table). We also observed that the abundance of 39 lncRNAs was significantly ($P \leq 0.05$) increased at both the time points. The number of DE *B. napus* lncRNAs was higher at 24 hpi than 48 hpi. This is further supported by the fact that, the correlation of the lncRNA expression levels between 24 hpi control vs.24 hpi pathogen challenged ($R = 0.939$; $P < 0.01$) were higher when compared 48hpi control vs.48 hpi pathogen challenged ($R = 0.82$; $P < 0.01$) (Fig 3B). This suggests that *B. napus* lncRNA mediated post transcriptional changes might be occurring at 24 hpi of *S. sclerotiorum* challenge which gradually decreased by 48 hpi. This could be due to increase in virulence and high necrosis at 48 hpi which might have resulted in very little tissue for isolating lncRNAs.

Classification of lncRNAs into ncRNA families in *B. napus*

To annotate lncRNAs from an evolutionary point of view, we classified the predicted lncRNAs into different ncRNA families using INFERNAL [47]. Based on a consensus secondary structure annotation using a covariance model, we identified 415 unique non-coding sequences belonging to 126 conserved lncRNA families from the libraries at both the time points (S6 Table). Among the conserved lncRNA families, four families (tRNA, snoR71, mir156 and

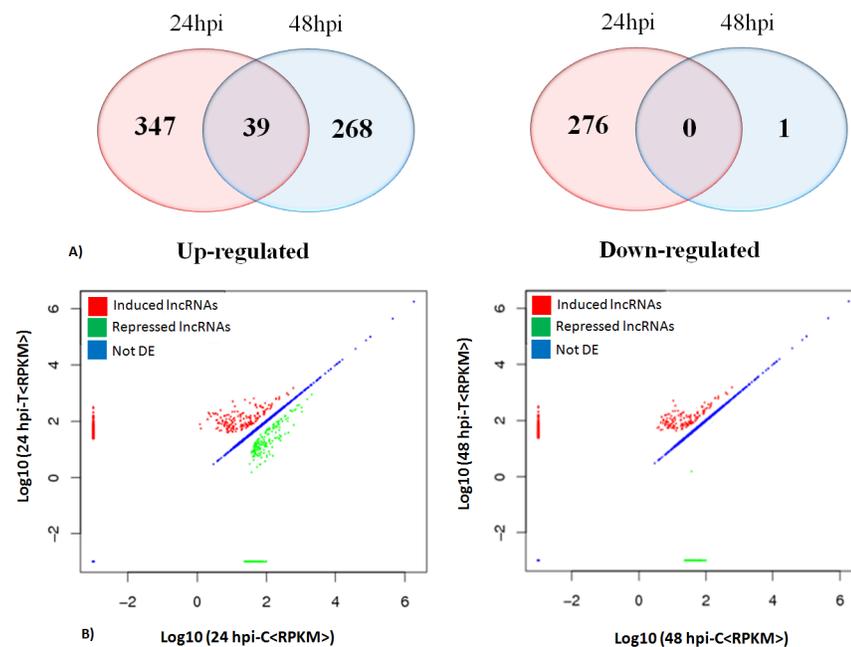


Fig 3. Differential expression of *B. napus* lncRNAs post inoculation with *S. sclerotiorum*. Venn diagram representation of differentially expressed lncRNAs at 24 hpi (A) and 48 hpi (B). Scatter plot representing the correlation analysis of differentially expressed lncRNAs at 24 hpi (C) and 48 hpi (D).

doi:10.1371/journal.pone.0158784.g003

mir169) accounted for more than 10 members. Moreover, several small nucleolar RNA families (SNORD14, SNORD18, SNORD25, snoR71, snoR116) constituted the major categories of lncRNAs in *B. napus* with more than 5 members. A recent report on the comparison of the genomic position of human lncRNAs with small RNAs (sRNAs) on the same strand revealed that the lncRNA exons are enriched for all classes of small RNAs with the exception of snRNAs and snoRNAs [55]. A similar study in wheat reported 26 lncRNAs as being putative precursors for 97 siRNAs and snoRNAs involved in resistance to powdery mildew infection [27]. Consistent with other results, it is possible that *B. napus* lncRNAs might serve as precursors for different types of functional sRNAs with or without having intrinsic functionality within them.

Additionally, many non-plant specific lncRNA families including the human nuclear enriched abundant transcript 1 conserved region 2 (NEAT1_2), X inactive specific transcript (XIST) and small cajal body-specific RNA 7 (SCARNA7) [56–59], bacterial small transcript non coding 430 (STnc430) and fungal small nucleolar RNA (Fungi_U3) [60] were also represented by one or more lncRNA members in *B. napus*. This suggests that, lncRNA transcripts may be evolutionarily conserved across the living systems. As it is difficult to compare divergent nucleic acid sequences, it may be assumed that lncRNAs are relatively conserved as gene units.

B. napus lncRNAs as potential miRNA precursors

MicroRNAs (miRNAs) are small RNAs of 20–22nt in length and act as an important class of ncRNAs involved in the regulation of gene expression at both transcriptional and post-transcriptional levels in plants [61]. These miRNAs are usually generated via sequential cleavage of long precursor transcripts by Dicer-like enzyme [62]. A single long non-coding transcript can be processed into multiple miRNAs, each with a distinct sub-cellular localization and unique function [5–6]. Since the mature miRNAs target multiple sites on mRNAs resulting in translational repression and gene silencing, identification of miRNAs and their targets could aid in our understanding of regulatory processes involved in mediating stress response in plants. Recent genome-wide studies have revealed that a major fraction of long non-coding transcripts in eukaryotes function as precursors for miRNAs [1, 14, 63]. In human keratinocytes and neonatal mice, miR675 has been derived from the lncRNAH19 and is endogenously expressed [64]. Similarly, four lncRNAs responsive to powdery mildew infection were found to be precursors for miR2004, miR2010 and miR2066 and were simultaneously up-regulated post pathogen challenge [27]. In the present study, alignment of 3,181 *B. napus* lncRNAs against the miRBase resulted in the identification of 28 lncRNAs having high homology (>90%) with known miRNA precursors from *B. napus*, *B. rapa*, *A. thaliana* and *A. lyrata* (S7 Table). Additionally, miRPara analysis of the lncRNA datasets revealed 13 lncRNAs with precursor sequences for 96 probable novel miRNAs from *B. napus*. Five lncRNAs (TCONS_00012499; TCONS_00004577; TCONS_00015411; TCONS_00004034 and TCONS_00009614) served as precursor for miR156, while another five lncRNAs (TCONS_00006568; TCONS_00018692; TCONS_000017152; TCONS_00008591; TCONS_000010926) were precursors for miR169 (Fig 4). Both miR156 and miR169 have been earlier reported to be significantly up-regulated post infection with fungal phytopathogens such as *Dothiorella gregaria* and *Botrytis cinerea* [65–66]. Three lncRNAs (TCONS_00027961; TCONS_00027962; TCONS_00011200) carried the precursor sequence for miR394 (Fig 4), whose potential targets are the F-box family proteins. Furthermore, eleven novel *B. napus* miRNAs identified through miRPara analysis also targeted F-box proteins. A recent report has shown that the expression of *Botrytis cinerea* responsive miR394 in *Lycopersicon esculentum* is negatively correlated with that of its target F-box protein [66]. The negative regulation of COI-F-box proteins in host plant is often associated with compromised Jasmonic Acid (JA) signaling

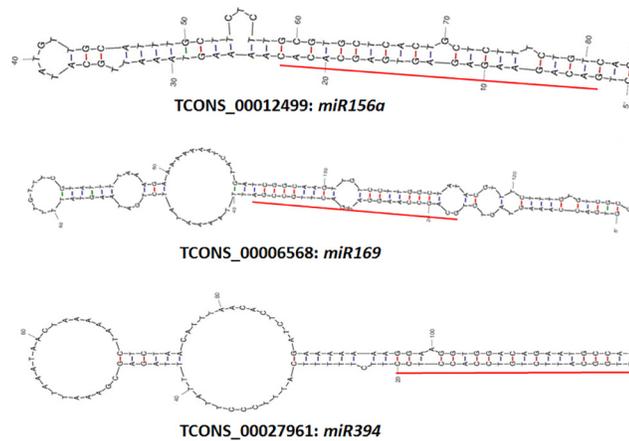


Fig 4. Representative potential miRNA precursor sequences in differentially expressed lncRNAs of *B. napus*. The red lines indicate the mature miRNA sequences.

doi:10.1371/journal.pone.0158784.g004

and high sensitivity to necrotrophic pathogens [67]. Therefore, it can be suggested that, biosynthesis of miR394 by three *B. napus* lncRNAs might be responsible for cleavage of F-box encoding transcripts which in turn may alter the JA dependent defense response against *S. sclerotiorum*. However, this suggestion needs to be confirmed through experiments, which are beyond the scope of this current study.

Analysis of interaction between complementary lncRNA-mRNA in *B. napus*

lncRNAs control various aspects of post-transcriptional mRNA processing through complementary base pairing with coding transcripts [68]. This sense-antisense base pairing is critical for the regulation of transcription and stability of the target mRNA [50]. To determine the role of antisense lncRNA in modulating transcription and mRNA stability, we used RNAPlex [50] to perform comparative interaction analysis between *B. napus* antisense lncRNA and the corresponding sense mRNA. Out of the 15 antisense lncRNAs identified in the present study, 10 lncRNAs showed more than 60% complementary base pairing with the sense mRNA. Specifically, antisense lncRNA TCONS_00000966 exhibited more than 90% complementary base pairing with the corresponding sense gene *BnaNo39233486*, which encodes for a plant defensin (PDF) protein (Fig 5). Diverse antisense transcript expression was also detected at a region with three defensin like genes in *Arabidopsis*, post inoculation with *F. oxysporum* [28]. Plant defensins are small basic peptides similar to those found in animals, produced in plants as part of an innate immune response and exhibit antifungal activity against a broad range of plant

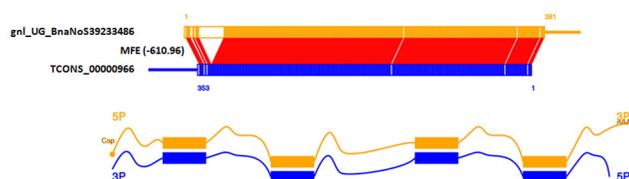


Fig 5. Complementary interaction analysis of a novel up-regulated lncRNA and sense coding transcripts of *B. napus* in response to *S. sclerotiorum* using RNAPlex. “gnl_UG_BnaNoS39233486” is the gene ID of a messenger RNA and “TCONS_00000966” is the ID of correspond lncRNA. Minimum free energy (MFE) is shown in the bracket at the centre.

doi:10.1371/journal.pone.0158784.g005

pathogens [69]. It can be assumed that TCONS_00000966 might have a significant role in the modulation of *BnaNo39233486* by complementary base pairing during pathogen infection. Therefore, future investigation towards development and characterization of knockout mutants for lncRNA TCONS_00000966 will be useful to determine its functional role in *B. napus* defense response to *S. sclerotiorum* infection.

Upstream/downstream lncRNAs regulating *B. napus* genes

Transcription from an upstream non-coding promoter can positively or negatively affect expression of the downstream coding gene by inducing chromatin remodelling or by inhibiting RNA polymerase II recruitment [5]. In humans, a ncRNA transcribed from an upstream region of the Dihydrofolate reductase (*DHFR*) locus forms a stable complex with the major promoter of *DHFR* and prevents the binding of the transcriptional co-factor transcription factor II D (TFIID) leading to the repression of *DHFR* gene [70]. Similarly, downstream lncRNA may bind to the 3' UTRs of the coding sequence and could be involved in intergenic regulatory mechanisms [71]. In the present study, of the 2,821 *B. napus* lncRNAs located in the intergenic regions, 500 DE lncRNAs were annotated to the upstream and 512 DE lncRNAs were annotated to 3' UTR and downstream of the coding transcripts (S8 Table). Specifically, nine lncRNAs from the intergenic regions of *B. napus* defense responsive genes were significantly differentially expressed in response to *S. sclerotiorum* (Table 2). For instance, lncRNA TCONS_00021690 located in the positive strand upstream of the *B. napus* gene, *BnaNoS38899684* encoding F-box protein was found repressed (-1.3 fold at 24 hpi; -0.59 fold at 48 hpi) in response to *S. sclerotiorum* infection. Similarly, lncRNA TCONS_00021619 located in the upstream of the *B. napus* gene, *BnaNoS55191345* encoding *Lipoxygenase 4* was significantly induced (6.69 fold at 24 hpi; 14.48 fold at 48 hpi) post infection with *S. sclerotiorum*. Our results suggest that *B. napus* lncRNAs in the intergenic regions might overlap with the cis-regulatory elements of the coding transcripts but exhibit contrasting roles towards transcriptional regulation of the defense responsive genes. Further research on the specific role(s) of these lncRNAs will provide additional information about their detailed roles in disease development or pathogen defense.

Validation of lncRNA expression using qRT-PCR

To validate the lncRNA expression results obtained from the RNA-Seq experiment, we used qRT-PCR to investigate the expression of *B. napus* lncRNAs (TCONS_00023722, TCONS_00023671, TCONS_00019243, TCONS_00006969, TCONS_00025169, TCONS_00011551, TCONS_00026914, TCONS_00014179, TCONS_00022944, TCONS_00000966, TCONS_00013795 and TCONS_00001004) with a range of expression levels in response to *S.*

Table 2. List of selected DE lncRNAs located upstream or downstream of defense responsive genes of *B. napus*.

lncRNA ID	Chr	Bracketing gene	Up/down stream	Bracketing gene annotation	Expression of gene
TCONS_00003548	A03	BnaNoS37273850	UPSTREAM_2K	Ethylene-responsive element binding factor 5	Induced
TCONS_00021619	C06	BnaNoS55191345	UPSTREAM_2K	Lipoxygenase 4	Induced
TCONS_00027151	C07	BnaNoS53250997	UPSTREAM_2K	Allene oxidase cyclase	Induced
TCONS_00021690	C06	BnaNoS38899684	UPSTREAM_2K	F-box protein	Repressed
TCONS_00018913	C03	BnaNoS31705296	UPSTREAM_2K	Glutathione S-transferase TAU 12	Induced
TCONS_00012312	A09	BnaNoS52591963	DOWNSTREAM_2K	Glucan endo-1,3-beta-glucosidase 10	Repressed
TCONS_00026803	C07	BnaNoS39174956	DOWNSTREAM_2K	Transcription factor subunit NF-YB3A	Repressed
TCONS_00026707	C07	BnaNoS25441206	DOWNSTREAM_2K	Glycine decarboxylase P-protein	Repressed
TCONS_00013960	A10	BnaNoS39270037	DOWNSTREAM_2K	Tetratricopeptide repeat domain protein	Repressed

doi:10.1371/journal.pone.0158784.t002

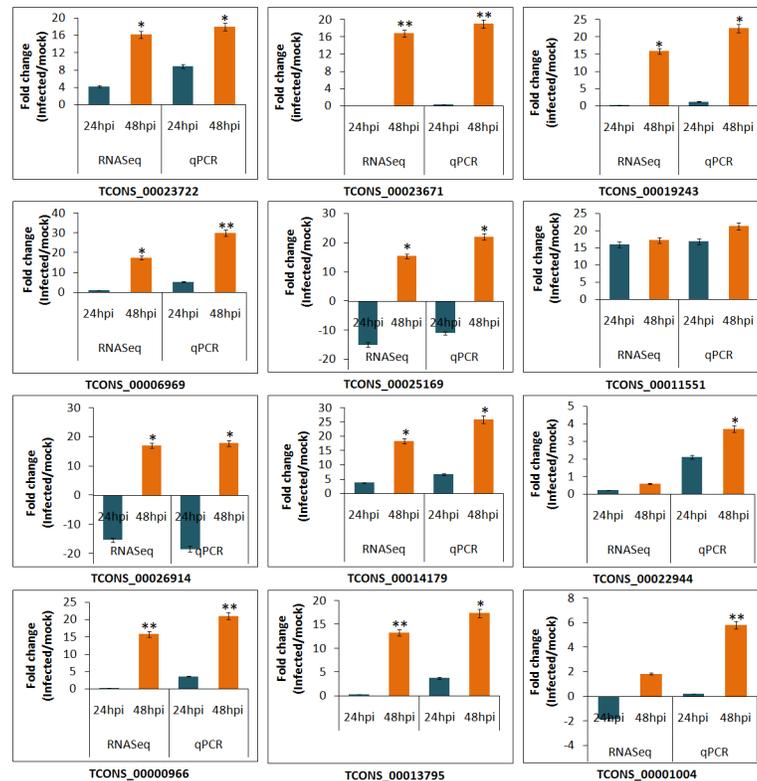


Fig 6. qRT-PCR validation of pathogen induced fold changes in 12 selected *B. napus* lncRNAs detected by RNA-seq. At each time point, the expression level in *S. sclerotiorum* infected sample was normalized to that in its corresponding mock sample. Standard error bars for the fold changes determined by qRT-PCR are shown.

doi:10.1371/journal.pone.0158784.g006

sclerotiorum. As shown in Fig 6, the expression patterns of the pathogen responsive lncRNAs as investigated by RNA-Seq and qRT-PCR were relatively consistent with similar trends. However, the relative expression levels of all lncRNAs as evaluated by qRT-PCR were slightly higher than indicated by RNA-Seq. For example, *B. napus* lncRNA TCONS_00014179 showed an increased abundance of 3.74 fold at 24 hpi and 18.37 fold at 48 hpi with RNA-Seq while the observed increase in abundance was 6.65 fold at 24 hpi and 25.92 fold at 48 hpi with qRT-PCR. This difference in transcript abundance can be attributed to differential sensitivity of RNA-Seq and qRT-PCR towards normalization of gene expression data [72]. Thus, the validation of the RNA-Seq results by qRT-PCR supported the fact that the size normalized expression values (RPKM) obtained with RNA-Seq data can be efficiently used to determine the lncRNA transcript expression, and further confirmed a significant role for these lncRNAs in *B. napus*-*S. sclerotiorum* interaction.

Verification of sense/antisense transcript expression in *B. napus*

Antisense lncRNAs exhibit different modes of actions including inhibition of the synthesis of sense transcript, modulation of RNA splicing and controlling the abundance of sense RNA [73]. In the present study, nine pairs of the sense and antisense transcripts were selected for further verification using qRT-PCR to determine the involvement of the antisense lncRNAs on the expression of coding genes in *B. napus* (Fig 7). For example, antisense lncRNA TCONS_00001004 revealed low transcript abundance (0.703 fold) at 24 hpi and significant induction

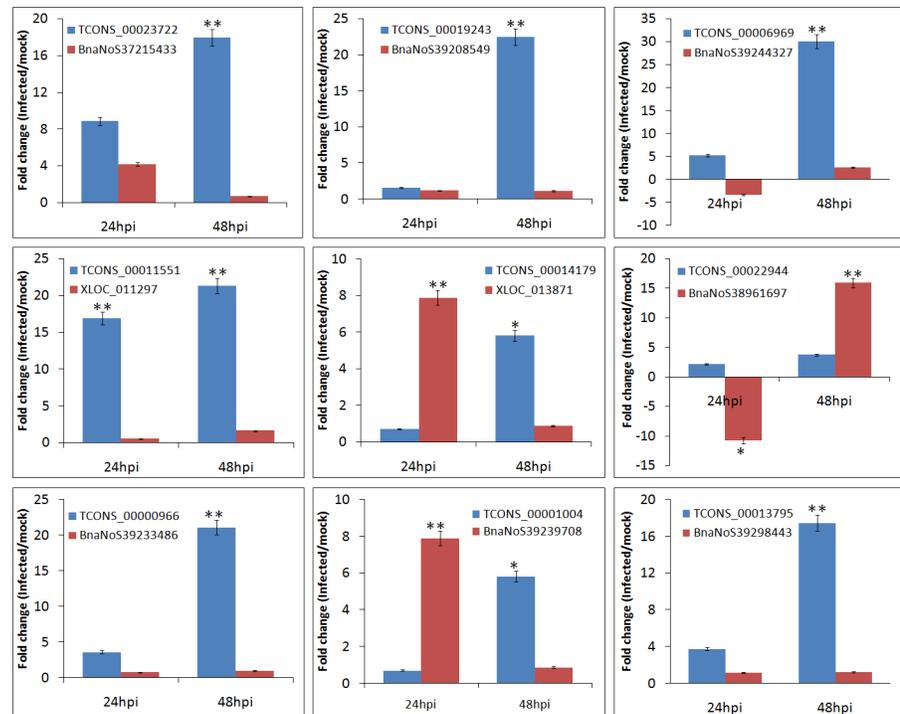


Fig 7. qRT-PCR verification of selected lncRNAs and the corresponding sense coding transcripts from *B. napus* detected by RNA-Seq. At each time point, the expression level in *S. sclerotiorum* infected sample was normalized to that in its corresponding mock sample. Two internal controls (*Actin* and *UBC9*) were used for data normalization.

doi:10.1371/journal.pone.0158784.g007

(7.89 fold) at 48 hpi while the corresponding *B. napus* sense transcript BnaNoS39239708 showed higher transcript abundance (5.81 fold) at 24 hpi and lower expression (0.875) at 48 hpi. Similar results were also observed for six pairs of sense/antisense transcripts in *Neurospora crassa* [73]. The influence of antisense transcript on the expression of sense gene largely depends upon the nature of overlap between them. The expression levels of sense transcript have been reported to decrease significantly with increase in overlap between the sense and the antisense pairs in human and mouse [74]. This is in accordance with the hypothesis that steric clashes of transcriptional machinery through collision between RNA polymerase II leads to depleted expression levels of the transcripts [75]. Further, our study also revealed that the other eight *B. napus* antisense transcripts with significant expression post infection with *S. sclerotiorum* had no associated expression of the sense transcripts (data not shown). This may be due to the absence of sense transcriptional activators or presence of strong antisense promoters resulting in steric clashes of the transcriptional machinery and subsequently leading to the suppression or termination of transcription from the sense promoter.

Conclusions

By applying a strand specific approach we have systematically identified, for the first time, 3,181 novel lncRNAs in *B. napus*. Moreover, our study revealed 931 *B. napus* lncRNAs responsive to *S. sclerotiorum* infection suggesting that they play important roles in mediating host responses to the pathogen. The lncRNAs were classified into 126 families with broad conservation across eukaryotes and some of them appeared to serve as precursors for major stress responsive miRNAs. Additionally, we also identified many lncRNAs overlapping with the cis-

regulatory regions of *B. napus* DEGs with possible role in defense response, which needs further elucidation. Overall, our study has demonstrated that lncRNAs may be important in mediating responses of *B. napus* to *S. sclerotiorum* and will provide a starting point for future investigation into the molecular mechanisms and regulatory functions of lncRNAs.

Supporting Information

S1 Table. List of primers used for qRT-PCR.

(XLSX)

S2 Table. Map summary of the RNA-Seq reads from 12 stranded *B. napus* libraries post infection with *S. sclerotiorum*.

(XLSX)

S3 Table. List of 17,964 un-annotated transcripts obtained through initial assembly with Cufflinks.

(XLSX)

S4 Table. List of 3,181 novel *B. napus* lncRNAs.

(XLSX)

S5 Table. List of differentially expressed *B. napus* lncRNAs post inoculation with *S. sclerotiorum*.

(XLSX)

S6 Table. Classification of predicted lncRNAs into different non coding RNA families.

(XLSX)

S7 Table. *B. napus* lncRNAs serving as potential miRNA precursors.

(XLSX)

S8 Table. *B. napus* differentially expressed lncRNAs located upstream or downstream of coding transcripts.

(XLSX)

Acknowledgments

RKJ is grateful to Department of Biotechnology, Government of India for providing DBT-CREST award (Sanction letter no. BT/IN/DBT-CREST Awards/32/RKJ/2012-13) to pursue research as a Visiting Scientist at the University of Alberta.

Author Contributions

Conceived and designed the experiments: RKJ SM UB MHR NNVK. Performed the experiments: RKJ SM. Analyzed the data: RKJ SM UB NNVK. Contributed reagents/materials/analysis tools: NNVK. Wrote the paper: RKJ SM UB MHR NNVK.

References

1. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, et al. (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316: 1484–1488. PMID: [17510325](https://pubmed.ncbi.nlm.nih.gov/17510325/)
2. Sanchez-Leon N, Arteaga-Vazquez M, Alvarez-Mejia C, Mendiola-Soto J, Duran-Figueroa N, Rodriguez-Leal D, et al. (2012) Transcriptional analysis of the *Arabidopsis* ovule by massively parallel signature sequencing. *J Expt Bot.* 63: 3829–3842.

3. Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, et al. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447: 799–816. PMID: [17571346](#)
4. Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Penkett CJ, et al. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single nucleotide resolution. *Nature Lett* 453: 1239–1245.
5. Wilusz JE, Sunwoo H, Spector DL (2009) Long noncoding RNAs: functional surprises from the RNA world. *Gen Dev* 23: 1494–1504.
6. Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. *Nature Rev Genet* 10: 155–159. doi: [10.1038/nrg2521](#) PMID: [19188922](#)
7. Kim ED, Sung S (2012) Long non-coding RNA: unveiling hidden layer of gene regulatory networks. *Trends Plant Sci* 17: 16–21. doi: [10.1016/j.tplants.2011.10.008](#) PMID: [22104407](#)
8. Zhu QH, Wang MB (2012) Molecular functions of long non-coding RNAs in plants. *Genes* 3: 15.
9. Ghildiyal M, Zamore PD (2009) Small silencing RNAs: An expanding universe. *Nat Rev Genet* 10:94–108. doi: [10.1038/nrg2504](#) PMID: [19148191](#)
10. Chitwood DH, Timmermans MC (2010) Small RNAs are on the move. *Nature* 467: 415–419. doi: [10.1038/nature09351](#) PMID: [20864994](#)
11. Seo JK, Wu J, Lii Y, Li Y, Jin H (2013) Contribution of small RNA pathway components in plant immunity. *Mol Plant Microbe Interact* 26: 617–625. doi: [10.1094/MPMI-10-12-0255-IA](#) PMID: [23489060](#)
12. Swiezewski S, Liu F, Magusin A, Dean C (2009) Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature* 462:799–802. doi: [10.1038/nature08618](#) PMID: [20010688](#)
13. Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* 331: 76–79. doi: [10.1126/science.1197349](#) PMID: [21127216](#)
14. Liu J, Jung C, Xu J, Wang H, Deng S, Bernad L, et al. (2012) Genome-wide analysis uncovers regulation of long intergenic non-coding RNAs in *Arabidopsis*. *Plant Cell* 24: 4333–4345. doi: [10.1105/tpc.112.102855](#) PMID: [23136377](#)
15. Wang H, Chung PJ, Liu J, Jang IC, Kean MJ, Xu J, et al. (2014) Genome wide identification of long non-coding natural antisense transcripts and their responses to light in *Arabidopsis*. *Genome Res* 24: 444–453. doi: [10.1101/gr.165555.113](#) PMID: [24402519](#)
16. Lu T, Zhu C, Lu G, Zhou Y, Zhang Z, Zhao Y, et al. (2012) Strand specific RNA-Seq reveals widespread occurrence of novel cis-natural antisense transcripts in rice. *BMC Genomics* 13:721–735. doi: [10.1186/1471-2164-13-721](#) PMID: [23259405](#)
17. Li L, Eichten SR, Shimizu R, Petsch K, Yeh CT, et al. (2014) Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biol* 15: R40. doi: [10.1186/gb-2014-15-2-r40](#) PMID: [24576388](#)
18. Wang H, Niu QW, Wu HW, Liu J, Ye J, Yu N, Vhuan NH (2015) Analysis of non-coding transcriptome in rice and maize uncovers roles of conserved lncRNAs associated with agricultural traits. *The Plant Journal* 84: 404–416. doi: [10.1111/tpj.13018](#) PMID: [26387578](#)
19. Zhang W, Han Z, Guo Q, Liu Y, Zheng Y, et al. (2014) Identification of maize long non coding RNAs responsive to drought stress. *PloS ONE* 9(6): e98958. doi: [10.1371/journal.pone.0098958](#) PMID: [24892290](#)
20. Shuai P, Liang D, Tang S, Zhang Z, Ye CY, Su Y, et al. (2014) Genome wide identification and functional prediction of novel and drought-responsive lincRNAs in *Populus trichocarpa*. *J Expt Bot*. doi: [10.1093/jxb/eru256](#)
21. Ben Amor B, Wirth S, Merchan F, Laporte P, d'Aubenton-Carafa Y, Hirsch J, et al. (2009) Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. *Genome Res* 19: 57–69. doi: [10.1101/gr.080275.108](#) PMID: [18997003](#)
22. Wu HJ, Wang ZM, Wang M, Wang XJ (2013) Widespread long non-coding RNAs as endogenous target mimics for microRNAs in plants. *Plant Physiol* 161:1875–1884. doi: [10.1104/pp.113.215962](#) PMID: [23429259](#)
23. Campalans A, Kondorosi A, Crespi M (2004) Enod40, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*. *Plant Cell* 16: 1047–1059. PMID: [15037734](#)
24. Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, et al. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet* 39:1033–1037. PMID: [17643101](#)
25. Cho J, Koo DH, Nam YW, Han CT, Lim HT, Bang JW, et al. (2005) Isolation and characterization of cDNA clones expressed under male sex expression conditions in a monoecious cucumber plant (*Cucumis sativus* L. cv Winter Long). *Euphytica* 146: 271–281.

26. Ma J, Yan B, Qu Y, Qin F, Yang Y, Hao X, et al. (2008) *Zm401*, a short-open reading frame mRNA or non-coding RNA, is essential for tapetum and microspore development and can regulate the floret formation in maize. *J Cell Biochem.* 105: 136–146. doi: [10.1002/jcb.21807](https://doi.org/10.1002/jcb.21807) PMID: [18465785](https://pubmed.ncbi.nlm.nih.gov/18465785/)
27. Xin M, Wang Y, Yao Y, Song N, Hu Z, Qin D, et al. (2011) Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. *BMC Plant Biol.* 7: 11–61.
28. Zhu QH, Stephen S, Taylor J, Helliwell CA, Wang MB (2014) Long non-coding RNAs responsive to *Fusarium oxysporum* infection in *Arabidopsis thaliana*. *New Phytol.* 201: 574–584. doi: [10.1111/nph.12537](https://doi.org/10.1111/nph.12537) PMID: [24117540](https://pubmed.ncbi.nlm.nih.gov/24117540/)
29. Bolton MD, Thomma BPHJ, Nelson BD (2006) *Sclerotinia sclerotiorum* (Lib.) de Bary: biology and molecular traits of a cosmopolitan pathogen. *Mol Plant Pathol.* 7:1–16. doi: [10.1111/j.1364-3703.2005.00316.x](https://doi.org/10.1111/j.1364-3703.2005.00316.x) PMID: [20507424](https://pubmed.ncbi.nlm.nih.gov/20507424/)
30. Del Rio LE, Bradley CA, Henson RA, Endres GJ, Hanson BK, McKay K, et al. (2007) Impact of *Sclerotinia* stem rot on yield of canola. *Plant Dis.* 91:191–194.
31. Koch S, Dunker S, Kleinhenz B, Rohrig M, von Tiedemann A (2007) Crop loss-related forecasting model for *Sclerotinia* stem rot in winter oilseed rape. *Phytopathol.* 97(9):1186–1194.
32. Malvárez G, Carbone I, Grünwald NJ, Subbarao KV, Schafer MR (2007) New populations of *Sclerotinia sclerotiorum* from lettuce in California and peas and lentils in Washington. *Phytopathol.* 97: 470–483.
33. Yang B, Srivastava S, Deyholos MK, Kav NNV (2007) Transcriptional profiling of canola (*Brassica napus* L.) responses to the fungal pathogen *Sclerotinia sclerotiorum*. *Plant Sci.* 173:156–171.
34. Liang Y, Srivastava S, Rahman MH, Strelkov SE, Kav NNV (2008) Proteome changes in the leaves of *Brassica napus* L. as a result of *Sclerotinia sclerotiorum* challenge. *J Agril Food Chem.* 56: 1963–1976.
35. Zhao J, Buchwaldt L, Rimmer SR, Sharpe A, McGregor L, et al. (2009) Patterns of differential gene expression in *Brassica napus* cultivars infected with *Sclerotinia sclerotiorum*. *Mol Plant Pathol.* 10(5): 635–649. doi: [10.1111/j.1364-3703.2009.00558.x](https://doi.org/10.1111/j.1364-3703.2009.00558.x) PMID: [19694954](https://pubmed.ncbi.nlm.nih.gov/19694954/)
36. Li R, Yu C, Li Y, Lam TW, Yiu SM, et al. (2009) SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25(15): 1966–1967. doi: [10.1093/bioinformatics/btp336](https://doi.org/10.1093/bioinformatics/btp336) PMID: [19497933](https://pubmed.ncbi.nlm.nih.gov/19497933/)
37. Wang X, Wang H, Wang J, Sun R, Wu J, et al. (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genet.* 43:1035–1039. doi: [10.1038/ng.919](https://doi.org/10.1038/ng.919) PMID: [21873998](https://pubmed.ncbi.nlm.nih.gov/21873998/)
38. Liu S, Liu Y, Yang X, Tong C, Edwards D, et al. (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Comm.* 5: 3930.
39. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley A, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomics in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14: R36. doi: [10.1186/gb-2013-14-4-r36](https://doi.org/10.1186/gb-2013-14-4-r36) PMID: [23618408](https://pubmed.ncbi.nlm.nih.gov/23618408/)
40. Amselem J, Cuomo CA, van Kan JA, Viaud M, Benito EP, Coulox A, et al (2011) Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS Genet.* 7(8): e1002230. doi: [10.1371/journal.pgen.1002230](https://doi.org/10.1371/journal.pgen.1002230) PMID: [21876677](https://pubmed.ncbi.nlm.nih.gov/21876677/)
41. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. (2012) Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks. *Nature Protocols* 7: 562–578. doi: [10.1038/nprot.2012.016](https://doi.org/10.1038/nprot.2012.016) PMID: [22383036](https://pubmed.ncbi.nlm.nih.gov/22383036/)
42. Roberts A, Pimentel H, Trapnell C, Pachter L (2011) Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* 27: 2325–2329. doi: [10.1093/bioinformatics/btr355](https://doi.org/10.1093/bioinformatics/btr355) PMID: [21697122](https://pubmed.ncbi.nlm.nih.gov/21697122/)
43. Kong L, Zhang Y, Ye ZQ, Zhao SQ, Wei L, Gao G (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic acids research* 35: W345–W349. PMID: [17631615](https://pubmed.ncbi.nlm.nih.gov/17631615/)
44. Sun K, Chen X, Jiang P, Song X, Wang H, Sun H (2013) iSeeRNA: identification of long intergenic non-coding RNA transcripts from transcriptome sequencing data. *BMC Genom.* 14:S7.
45. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics.* 29:1165–1188.
46. Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A (2011) Differential expression in RNA-seq: a matter of depth. *Genome Res.* 21: 2213–2223. doi: [10.1101/gr.124321.111](https://doi.org/10.1101/gr.124321.111) PMID: [21903743](https://pubmed.ncbi.nlm.nih.gov/21903743/)
47. Nawrocki E P, Kolbe D L, Eddy S R (2009) Infernal 1.0: inference of RNA alignments. *Bioinformatics* 25: 1335–1337. doi: [10.1093/bioinformatics/btp157](https://doi.org/10.1093/bioinformatics/btp157) PMID: [19307242](https://pubmed.ncbi.nlm.nih.gov/19307242/)
48. Griffiths-Jones S, Grocock R J, Van Dongen S, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34: D140–D144. PMID: [16381832](https://pubmed.ncbi.nlm.nih.gov/16381832/)

49. Wu Y, Wei B, Liu H, Li T, Rayner S (2011) MiRPara: a SVM-based software tool for prediction of most probable microRNA coding regions in genome scale sequences. *BMC Bioinformatics* 12(1): 107.
50. Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, et al. (2012) Long non-coding anti-sense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* 491(7424): 454–457. doi: [10.1038/nature11508](https://doi.org/10.1038/nature11508) PMID: [23064229](https://pubmed.ncbi.nlm.nih.gov/23064229/)
51. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* 25: 402–408. PMID: [11846609](https://pubmed.ncbi.nlm.nih.gov/11846609/)
52. Joshi RK, Megha S, Rahman MH, Basu U, Kav NNV (2016) A global study of transcriptome dynamics in canola (*Brassica napus* L.) responsive to *Sclerotinia sclerotiorum* infection using RNA-Seq. *Gene* 590: 57–67. doi: [10.1016/j.gene.2016.06.003](https://doi.org/10.1016/j.gene.2016.06.003) PMID: [27265030](https://pubmed.ncbi.nlm.nih.gov/27265030/)
53. Xie C, Yuan J, Li H, Zhao G, Bu D, et al. (2014) NONCODEv4: exploring the world of long non coding RNA genes. *Nucleic Acids Res.* 42: D98–D103. doi: [10.1093/nar/gkt1222](https://doi.org/10.1093/nar/gkt1222) PMID: [24285305](https://pubmed.ncbi.nlm.nih.gov/24285305/)
54. Sun L, Zhang Z, Bailey TL, Perkins AC, Tallack MR, et al. (2012) Prediction of novel long non-coding RNAs based on RNA-Seq data of mouse *Klf1* knockout study. *BMC Bioinformatics*. 13: 331 doi: [10.1186/1471-2105-13-331](https://doi.org/10.1186/1471-2105-13-331) PMID: [23237380](https://pubmed.ncbi.nlm.nih.gov/23237380/)
55. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. (2012) The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res* 22: 1775–1789. doi: [10.1101/gr.132159.111](https://doi.org/10.1101/gr.132159.111) PMID: [22955988](https://pubmed.ncbi.nlm.nih.gov/22955988/)
56. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. (2011) Integrative annotation of human large intergenic non-coding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25: 1915–1927. doi: [10.1101/gad.17446611](https://doi.org/10.1101/gad.17446611) PMID: [21890647](https://pubmed.ncbi.nlm.nih.gov/21890647/)
57. Zhang YC, Chen YQ (2013) Long non-coding RNAs: new regulators in plant development. *Biochem Biophys Res Comm.* 436: 111–114.
58. Darzacq X, Jady BE, Verheggen C, Kiss AM, Bertrand E, Kiss T (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *The EMBO Journal* 21(11): 2746–2756. PMID: [12032087](https://pubmed.ncbi.nlm.nih.gov/12032087/)
59. Gontan C, Jonkers I, Gribnau J (2011) Long non-coding RNAs and X chromosome inactivation. *Prog Mol Subcell Biol.* 51: 43–64. doi: [10.1007/978-3-642-16502-3_3](https://doi.org/10.1007/978-3-642-16502-3_3) PMID: [21287133](https://pubmed.ncbi.nlm.nih.gov/21287133/)
60. Zwieb C (1997) The uRNA database. *Nucleic Acids Res.* 25: 102–103. PMID: [9016512](https://pubmed.ncbi.nlm.nih.gov/9016512/)
61. Chen X (2012) Small RNAs in development—insights from plants. *Curr Opin Genet Dev.* 22: 361–367. doi: [10.1016/j.gde.2012.04.004](https://doi.org/10.1016/j.gde.2012.04.004) PMID: [22578318](https://pubmed.ncbi.nlm.nih.gov/22578318/)
62. Sunkar R, Li YF, Jagadeeswaran G (2012) Functions of microRNAs in plant stress responses. *Trends Plant Sci.* 17: 196–203. doi: [10.1016/j.tplants.2012.01.010](https://doi.org/10.1016/j.tplants.2012.01.010) PMID: [22365280](https://pubmed.ncbi.nlm.nih.gov/22365280/)
63. Fejes-Toth K, Sotirova V, Sachidanandam R, Assaf G et al. (2009) Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. *Nature* 457:1028–1032. doi: [10.1038/nature07759](https://doi.org/10.1038/nature07759) PMID: [19169241](https://pubmed.ncbi.nlm.nih.gov/19169241/)
64. Cai X, Cullen BR (2007) The imprinted H19 non-coding RNA is a primary micron precursor. *RNA* 13: 313–316. PMID: [17237358](https://pubmed.ncbi.nlm.nih.gov/17237358/)
65. Chen L, Ren Y, Zhang Y, Xu J, Zhang Z, Wang Y (2011) Genome-wide profiling of novel and conserved *Populus* microRNAs involved in pathogen stress response by deep sequencing. *Planta* 235: 873–883. doi: [10.1007/s00425-011-1548-z](https://doi.org/10.1007/s00425-011-1548-z) PMID: [22101925](https://pubmed.ncbi.nlm.nih.gov/22101925/)
66. Jin W, Wu F (2015) Characterization of miRNAs associated with *Botrytis cinerea* infection of tomato leaves. *BMC Plant Biol.* 15:1 (doi: [10.1186/s12870-014-0410-4](https://doi.org/10.1186/s12870-014-0410-4)) PMID: [25592487](https://pubmed.ncbi.nlm.nih.gov/25592487/)
67. Lorenzo O, Piqueras R, Sanchez-Serrano JJ, Solano R (2003) ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense. *Plant Cell* 15: 165–178. PMID: [12509529](https://pubmed.ncbi.nlm.nih.gov/12509529/)
68. Geisler S, Collier J (2013) RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol.* 14: 699–712. doi: [10.1038/nrm3679](https://doi.org/10.1038/nrm3679) PMID: [24105322](https://pubmed.ncbi.nlm.nih.gov/24105322/)
69. Stotz HU, Thomson JG, Wang Y (2009) Plant defensins: defense, development and application. *Plant Signal Behav.* 4: 1010–1012. PMID: [20009545](https://pubmed.ncbi.nlm.nih.gov/20009545/)
70. Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A (2007) Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* 445: 666–670. PMID: [17237763](https://pubmed.ncbi.nlm.nih.gov/17237763/)
71. Carninci P, Kasukawa T, Katayama S, Gough J, Frith M, et al. (2005) The transcriptional landscape of the mammalian genome. *Science* 309: 1559–1563. PMID: [16141072](https://pubmed.ncbi.nlm.nih.gov/16141072/)
72. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 10: 57–63. doi: [10.1038/nrg2484](https://doi.org/10.1038/nrg2484) PMID: [19015660](https://pubmed.ncbi.nlm.nih.gov/19015660/)

73. Arthanari Y, Heintzen C, Griffiths-Jones S, Crosthwaite SK (2014) Natural antisense transcripts and long non-coding RNA in *Neurospora crassa*. PLoS ONE 9:e91353. doi: [10.1371/journal.pone.0091353](https://doi.org/10.1371/journal.pone.0091353) PMID: [24621812](https://pubmed.ncbi.nlm.nih.gov/24621812/)
74. Osato N, Suzuki Y, Ikeo K, Gojobori T (2007) Transcriptional interferences in cis natural antisense transcripts of humans and mice. Genetics 176: 1299–1306. PMID: [17409075](https://pubmed.ncbi.nlm.nih.gov/17409075/)
75. Hobson DJ, Wei W, Steinmetz LM, Svejstrup JQ (2012) RNA polymerase II collision interrupts convergent transcription. Mol Cell. 48: 365–374. doi: [10.1016/j.molcel.2012.08.027](https://doi.org/10.1016/j.molcel.2012.08.027) PMID: [23041286](https://pubmed.ncbi.nlm.nih.gov/23041286/)