

SCIENTIFIC REPORTS



OPEN

Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection

Received: 14 June 2015
Accepted: 22 October 2015
Published: 18 December 2015

Jinyan Wang¹, Wengui Yu¹, Yuwen Yang¹, Xiao Li³, Tianzi Chen¹, Tingli Liu¹, Na Ma¹, Xu Yang³, Renyi Liu² & Baolong Zhang¹

Recently, a large number of long noncoding RNAs (lncRNAs) have emerged as important regulators of many biological processes in animals and plants. However, how lncRNAs function during plant DNA virus infection is largely unknown. We performed strand-specific paired-end RNA sequencing of tomato samples infected with Tomato yellow leaf curl virus (TYLCV) with three biological replicates. Overall, we predicted 1565 lncRNAs including long intergenic ncRNAs (lincRNAs) and natural antisense transcripts (lncNATs) and definitively identified lncRNAs that are involved in TYLCV infection by virus-induced gene silencing (VIGS). We also verified the functions of a set of lncRNAs that were differentially expressed between 0 and 7 days post inoculation (dpi). More importantly, we found that several lncRNAs acted as competing endogenous target mimics (eTMs) for tomato microRNAs involved in the TYLCV infection. These results provide new insight into lncRNAs involved in the response to TYLCV infection that are important components of the TYLCV network in tomatoes.

Non-coding RNAs (ncRNAs) have emerged as major products of the eukaryotic transcriptome with regulatory importance^{1,2}. Over the last decade, significant progress has been made in our understanding of the functions and mechanisms of microRNAs (miRNAs), small interfering RNAs (siRNAs), and natural antisense siRNAs (nat-siRNAs) in the transcriptional and post-transcriptional regulation of gene expression^{3,4}. Recently, ncRNAs longer than 200 nucleotides have been defined as long non-coding RNAs (lncRNAs) and identified as new regulatory elements that are involved in many biological processes in mammals^{5–7}. Although thousands of these lncRNAs have been identified using RNA-seq and bioinformatics analyses in *Arabidopsis thaliana*, *Oryza sativa*, *Zea mays*, *Medicago truncatula*, *Solanum lycopersicum* and *Cucumis sativus*^{8–14}, the functions of lncRNAs in plants are poorly understood. The exceptions are a few lncRNAs such as the cold induced long antisense intragenic RNA (*COOLAIR*) and cold assisted intronic noncoding RNA (*COLDAIR*). *COOLAIR* and *COLDAIR* regulate vernalization in *Arabidopsis* by interacting with the polycomb-repressive complex 2 (PRC2) to modify vernalization-mediated epigenetic repression of the *Flowing Locus C (FLC)* locus and repress *FLC* expression^{15–17}.

lncRNAs can be generally classified into three groups based on their genomic regions: (i) long intergenic ncRNAs (lincRNAs), (ii) intronic ncRNAs (incRNAs) and (iii) natural antisense transcripts (NATs),

¹Jiangsu Key Laboratory for Bioresources of Saline Soils, Provincial Key Laboratory of Agrobiolgy, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China. ²Shanghai Center for Plant Stress Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 201602, China. ³College of Horticulture and Plant Protection, Yangzhou University, Yangzhou 225009, China. Correspondence and requests for materials should be addressed to B.Z. (email: zhbl2248@hotmail.com)

which are transcribed from the complementary DNA strand of their associated genes¹⁸. These lncRNAs can regulate gene expression at the transcriptional and post-transcriptional level by acting as signals, decoys, guides, and scaffolds¹⁹. Moreover, emerging evidence suggests that the expression of some lncRNAs is highly tissue-specific, and many of them are responsive to biotic and abiotic stresses^{20–22}. The application of next-generation sequencing technology greatly facilitated the discovery of lncRNAs in plants. For example, 2,224 lncRNAs were identified in rice, including lincRNAs and lncNATs, that were expressed in a tissue-specific or stage-specific manner¹¹. In *Arabidopsis*, 626 concordant and 766 discordant NATs pairs affected spatial and developmental-specific light effects²³. Using strand-specific RNA sequencing, 159 novel transcriptionally active regions (TARs) and 20 *Fusarium oxysporum*-responsive lncTARs were identified in *Arabidopsis*²⁰. Additionally, Di *et al.* (2014) identified 245 poly(A)+ and 58 poly(A)- lncRNAs that were differentially expressed under various stresses²¹. In *Populus trichocarpa*, 2,542 lincRNA candidates were identified including 504 drought-responsive lincRNAs under control and drought conditions²². In tomatoes, 3,679 lncRNAs were discovered in wild-type and ripening mutant fruit. Moreover, some lncRNAs were significantly differentially expressed in ripening mutant fruit, including two novel intergenic lncRNAs that could induce an obvious delay in fruit ripening by down-regulating these genes in wild-type tomatoes¹⁴. Tomato yellow leaf curl virus (TYLCV) causes one of the most devastating diseases of tomatoes worldwide, and ranks 3rd among plant viruses of scientific/economic importance²⁴. TYLCV belongs to the genus *Begomovirus* of the family *Geminiviridae* and is transmitted by the whitefly *Bemisia tabaci*. The rapid spread of the viral disease is caused by whitefly pressure and a high transmission efficacy. The symptoms of TYLCV infection in young plants include stunted growth, upward curling of leaf margins, marked reduction in leaf size, mottling and yellowing of young leaves, flower abscission, and severe yield loss²⁵. Breeding tomatoes resistant to TYLCV started in the mid-1970s and several commercial varieties with adequate resistance have been released. Breeding involved the introgression of resistance found in accessions of several wild tomato species (e.g., *Solanum chilense*, *S. peruvianum*, *S. pimpinellifolium* and *S. habrochaites*) into the domesticated tomato (*S. lycopersicum*)²⁶. Several loci tightly linked to TYLCV resistance (coined *Ty-1* to *Ty-5*) have been mapped to the tomato chromosomes^{26–29}. Among them, *Ty-1* and *Ty-3* were found to be allelic and were identified as RNA-dependent RNA polymerases (RDRs) that might be involved in RNA silencing³⁰. Furthermore, relative hyper-methylation of the TYLCV V1 promoter region was observed in *Ty-1* resistant tomatoes compared with susceptible tomato³¹. Despite the significant understanding that has been gained for the *Ty-1* genes, research on the *Ty-2* gene is lacking. Recently, *Ty-2* was mapped to an approximately 300 kb interval between molecular markers UP8 and M1 on chromosome 11³². However, the *Ty-2* gene has not been cloned and its regulatory mechanism is unclear. In a previous study, whole transcriptome sequencing of a TYLCV-resistant (R) tomato breeding line with *Ty-2* loci and a TYLCV-susceptible (S) tomato breeding line helped identify 209 and 809 genes, respectively, that were differentially expressed between the two tomato lines³³. Furthermore, among the 152 bHLH transcription factors genes that were identified from the whole tomato genome analysis, four were differentially expressed after TYLCV inoculation³⁴.

In previous studies, lncRNAs were found to be involved in the response to biotic and abiotic stresses^{20,22}. However, whether lncRNAs participate in the TYLCV defense network in tomatoes is unknown. In this study, we performed whole transcriptome strand-specific RNA sequencing (ssRNA-seq) of tomato leaves with and without TYLCV inoculation with three biological replicates. In our analysis, we identified lncRNAs (lincRNAs and lncNATs) and validated some differentially expressed lncRNAs by qRT-PCR and virus-induced gene silencing (VIGS). Our results indicate that a large number of lncRNAs play important roles in TYLCV infection, including some that act as endogenous miRNA target mimics (eTMs).

Materials and Methods

Plant growth conditions and viral inoculation. The TYLCV-resistant tomato breeding line CLN2777A with *Ty-2* loci was grown in a chamber under 26 °C with a 16 h light/8 h dark cycle³³. Whiteflies viruliferous for the TYLCV-IL strain were propagated and maintained with the tomato plants in an insect-proof greenhouse^{35,36}. Tomato plants at the two-leaf stage were exposed to viruliferous whiteflies in an insect-proof cage for 3 days, and subsequently treated with an insecticidal imidacloprid to kill the whiteflies³⁰.

Plant sampling, virus detection, and sample sequencing. Leaf samples were collected 7 days post inoculation (dpi) and frozen immediately in liquid nitrogen. To ensure the success of TYLCV infection in the sequencing samples, DNA was extracted from young leaves of infected tomato, and SYBR PCR assay was performed on the qTOWER 2.0/2.2 (Analytik Jena, Germany) with the AceQ qPCR SYBR Green Master Mix (Vazyme, China) using the following PCR conditions: 5 min of denaturation at 95 °C followed by 40 cycles of 95 °C for 10 s, 60 °C for 30 s. The primers for TYLCV virus content and tomato α -tubulin (Solyc04g077020.2) were listed in Supplementary Table S3. Additionally, the susceptible breeding line TMXA48-4-0 was used to detect the TYLCV content as a negative control.

Three independent biological replicates of CLN2777A leaves from tomato plants successfully infected with TYLCV and uninfected plants were used for RNA sequencing. Poly(A) RNA enrichment and strand-specific RNA-seq library were prepared using the NEBNext[®] UltraTM RNA Library Prep Kit for Illumina (NEB, USA) according to the low sample protocol guidelines. Libraries were controlled for

quality using the Nannodrop 2000 system (Thermo, USA). The resulting libraries were sequenced on an Illumina HiSeq 2500 instrument with paired-end reads of 124 bp to obtain a total yield of ~494 million reads. The data for this study have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>) with accession number SRP061792.

Pipeline for lncRNA identification. The tomato genome assembly build 2.50 and annotation ITAG2.4 used throughout this study were downloaded from http://solgenomics.net/organism/Solanum_lycopersicum/genome. Successfully prefiltered reads were quality-trimmed and quality-filtered using FASTQC tools to remove low-quality and adapter-containing sequences (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Each RNA-seq dataset was aligned to the tomato genome separately with the Tophat 2.0 program³⁷ (Tophat2, -library-type 'fr-firststrand' splice-mismatches '0' -min-intron-length '70' -max-intron-length '50000' -num-threads '6'). The transcripts from each dataset were assembled using the Cufflinks 2.0 program³⁸ (Cufflinks2, -num-threads '8' -max-intron-length '300000' -max-mle-iterations '5000' -min-frags-per-transfrag '10' -min-intron-length '50' -minisoform-fraction '0.1' -num-importance-samples '1000' -library-type 'fr-firststrand'). All transcripts were pooled and merged to generate final transcripts using Cuffmerge (Cuffmerge2, -p '6' -min-isoform-fraction '0.1' -min-isoform-fraction '0.1'). Cuffdiff was used to estimate the abundance of all transcripts from the BAM output files of Tophat 2.0 (Cuffdiff, -p '6' -min-alignment-count '10' -library-type 'fr-firststrand'). All transcripts without strand information and transcripts that overlapped with known genes were discarded. The remaining transcripts were used to identify the lincRNAs and lincNATs. The transcripts located in intergenic regions were identified as lincRNA candidates, and the transcripts that were transcribed from the antisense strands of known genes were predicted to be lincNAT candidates. The transcripts with a FPKM (fragments per kilobase of transcript per million mapped reads) score higher than 1 in a single exon or 0.5 in multiple exons in at least one sample were retained. Transcripts with a length shorter than 200 bp and an open reading frame (ORF) length longer than 120 aa were discarded (ORF Finder, <http://www.ncbi.nlm.nih.gov/gorf/orf.cgi>). The CPC³⁹ and CNCI⁴⁰ programs were used to calculate the coding potential of the remaining transcripts. Only transcripts with both CPC and CNCI scores less than 0 were used for the subsequent analysis. The remaining transcripts were searched against the NCBI non-redundant (NR) protein database, KEGG (Kyoto Encyclopedia classification of protein database), COGs (NCBI Phylogenetic classification of proteins encoded in complete genomes), and Swiss-Prot (Swiss-Protein database) by BLASTX (E-value cutoff of 1e-10, coverage >80%, and identity >90%) to exclude transcripts with significant homology to known proteins.

5' and 3' rapid amplification of cDNA ends (RACE) experiments. Total RNA was isolated from leaf tissue using Total RNA extraction kit (Tiangen, China). 5' and 3' RACE was carried out with 5' and 3' full RACE core set (Takara, Japan) according to the manufacturer's instructions. The lncRNAs primers were designed according to the known partial sequence of slylnc0049 and slylnc0761 listed in Supplementary Table S3.

miRNA mimic prediction with lncRNAs. All lncRNAs candidates were used to predict miRNA mimic sites using the psMimic software⁴¹. Mature tomato miRNAs were downloaded from miRBase database (<http://www.mirbase.org/>; release 21, June 2014)⁴². In addition, some novel tomato mature miRNAs were obtained from previous studies^{43,44}. Putative tomato target genes of the predicted miRNAs that had mimicry with lncRNAs were identified using the plant miRNA target prediction online software psRobot with moderate parameters (penalty score threshold = 2.5, five prime boundary of essential sequence = 2, three prime boundary of essential sequence = 17, maximal number of permitted gaps = 1, and position after which with gaps permitted = 17)⁴⁵.

Validation of differentially expressed lncRNAs by quantitative RT-PCR. Nine differentially expressed lncRNAs in the resistant tomato lines were selected for quantitative RT-PCR validation (Supplementary Table S3). Primers for quantitative RT-PCR were designed using the Beacon Designer 7.5 software (Premier Biosoft International, Palo Alto, California, USA). PCR amplifications were performed in a real-time thermal cycler qTOWER 2.0/2.2 (Analytik Jena, Germany) with 15 µl final volumes containing 1.0 µl of cDNA, 0.5 µl of each primer (10 µM), 6 µl of sterile water, and 7.5 µl of (2 ×) SYBR Premix ExTaq™ II Kit (TaKaRa, Japan). The conditions for amplification were as follows: 5 min of denaturation at 95 °C followed by 40 cycles of 95 °C for 10 s, 60 °C for 20 s, and 72 °C for 10 s. The expression levels of selected lncRNAs were normalized to α -Tubulin (Solyc04g077020.2) expression³³. Relative gene expression was calculated using the $2^{-\Delta\Delta CT}$ method⁴⁶. Three biological replicates were performed for each of the selected lncRNAs.

Expression level of miR166 by qRT-PCR. cDNA for the miR166 qPCR assays was prepared using the HiScript II 1st Strand cDNA Synthesis Kit (Vazyme, China) following the protocol of Varkonyi-Gasic *et al.*⁴⁷. Stem-loop primers for reverse transcription of miR166 were designed such that the 6 bp at the 5' end of the stem-loop primer were complementary to the six nucleotides at the 3' end of miR166. In addition to one stem-loop primer, the cDNA reaction contained oligo(dT) primers. The sample were loaded into the thermal cycler used for pulsed reverse transcription and incubated for 30 min at 16 °C,

	CK-1	CK-2	CK-3	T-1	T-2	T-3	Total
Raw reads	100,605,764	75,137,922	78,240,172	85,568,182	72,748,606	82,686,274	494,986,920
Clean reads	100,587,550	75,097,894	78,223,370	85,552,432	72,734,108	82,666,626	494,861,980
Mapped Unique Left Reads	41,423,191	32,149,884	33,514,764	36,006,336	30,625,791	35,427,909	209,147,875
Mapped Nonunique Left Reads	470,034	346,807	342,747	444,922	365,608	442,072	2,412,190
Mapped Unique Right Reads	49,778,308	34,239,298	36,106,985	39,476,748	33,438,329	37,779,252	230,818,920
Mapped Nonunique Right Reads	515,467	357,848	355,581	469,683	388,550	451,905	2,539,034
Total mapping pair reads	40,534,890	31,248,591	32,698,987	35,247,998	29,908,144	34,491,185	204,129,795
Overall mapping	87.90%	89.30%	89.90%	89.30%	89.10%	89.60%	89.18%

Table 1. Summary of RNA-seq data. Note: CK-1, CK-2, CK-3 were control samples from tomato plants before TYLCV infection; T-1, T-2, T-3 were TYLCV samples from tomato plants collected 7 days after TYLCV inoculation.

followed by pulsed RT of 60 cycles at 30 °C for 30 s, 42 °C for 30 s and 50 °C for 1 s. Then, the samples were incubated at 85 °C for 5 min to inactivate the reverse transcriptase. The primers used for qPCR amplification are listed in Supplementary Table S3.

Validation of lncRNAs with virus-induced gene silencing (VIGS). The tobacco rattle virus (TRV) mediated VIGS system was used to silence the lncRNAs⁴⁸. Briefly, pTRV-containing *Agrobacterium* EHA105 was cultured in liquid LB medium and resuspended in infiltration medium at an O.D. value of 2.0 and cultured at room temperature for 4 h. Three week old seedlings were infiltrated by pressure inoculation in the leaves with a needleless syringe. For the VIGS experiments, agroinfiltration was performed two weeks after TRV inoculation.

Seven days after agroinfiltration, tomato plants were injected with a TYLCV infectious clone provided by Xueping Zhou (Zhejiang University) for a 3-day inoculation period³⁶. One month after agroinfiltration, new emerging leaves from the TYLCV infected plants were used to extract RNA and DNA, which was subsequently used to determine the expression levels of the lncRNAs and the accumulation of TYLCV DNA in the VIGS-treated plants by quantitative RT-PCR, respectively. The conditions and parameters of the quantitative RT-PCR were the same as described above.

Transient Agroinfiltration Assay in *Nicotiana benthamiana*. To construct agroinfiltration transient expression vectors, slylnc0195 was inserted into the *KpnI/XbaI*-digested pCAMBIA2301 vector. The amplification primers are provided in Supplementary Table S3.

The overexpression vector was transformed into *Agrobacterium tumefaciens* strain GV3101. *Agrobacterium* cells infiltrated into the leaves of *Nicotiana benthamiana* using pCAMBIA2301 as the control vector. The transient agroinfiltration assay was performed as described previously⁴⁹. The leaves were harvested 2 d after infiltration. The expression profiles were detected for the target genes of miR166. All primers are shown in Supplemental Table S3. The target genes of miR166 in *N. benthamiana* were predicted by psRobot⁵⁰.

Results

Genome-wide identification of lncRNAs in tomatoes. The two tomato lines showed marked differences in tolerance to TYLCV infection. Upon TYLCV infection, the leaves of the resistant line CLN2777A were normal, whereas those of the susceptible line TMXA48-4-0 were curly, mottled and yellow by 21 days post-inoculation (dpi) (Supplementary Fig. S1). In CLN2777A, the accumulation of TYLCV was barely above the level of detection whereas in TMXA48-4-0 a large quantity of TYLCV was detected (Supplementary Table S4).

We performed high-throughput strand-specific RNA-seq in the resistant tomato line CLN2777A at 0 dpi (CK) and 7 dpi of TYLCV infection, each with three biological replicates. We obtained more than 494 million clean reads that passed the quality filters (Table 1). These reads were mapped to the tomato reference genome (Assembly build 2.50), followed by transcript assembly, and differential isoform and gene expression analysis using Cufflinks (Fig. 1). Approximately 87% to 89% of clean reads were aligned to the reference genome for the CK and TYLCV samples, respectively. Approximately 87% of the reads were uniquely mapped to a single genomic locus, attesting to the high quality of the RNA-seq reads and the reference tomato genome. The annotated tomato reference genome (ITAG 2.4) comprises 34,725 protein-coding genes. Together with our transcript assemblies, the merged tomato gene annotation has 57,459 transcripts on 35,549 gene loci; out of these, ~3,558 transcripts were previously unannotated. The unannotated transcripts could be classified into two types: (1) transcriptional units mapped to previously unannotated regions of the genome (intergenic regions) and (2) natural antisense transcripts (NATs)

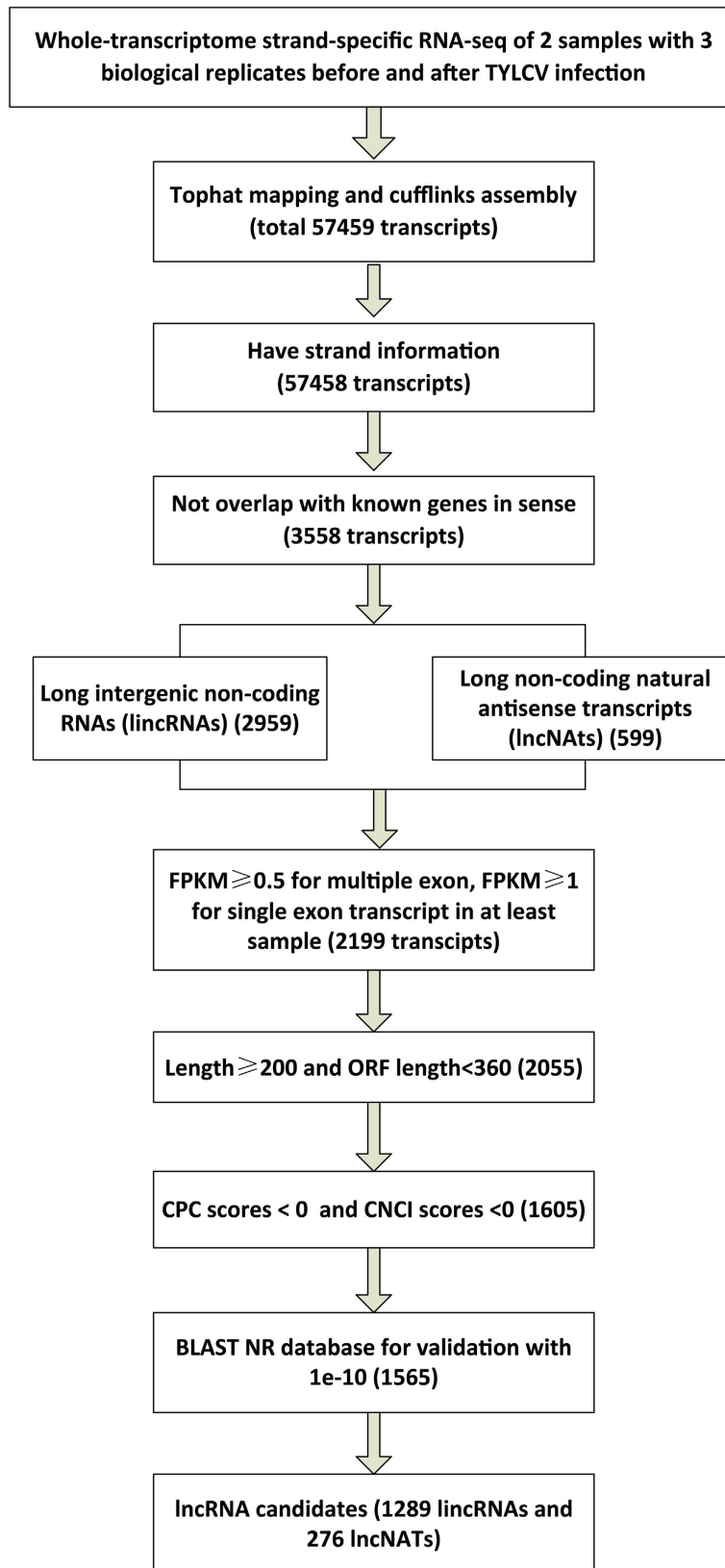


Figure 1. The bioinformatics pipeline for the systematic identification of lncRNAs in tomatoes. CPC, Coding Potential Calculator; CNCI, Coding-Non-Coding Index.

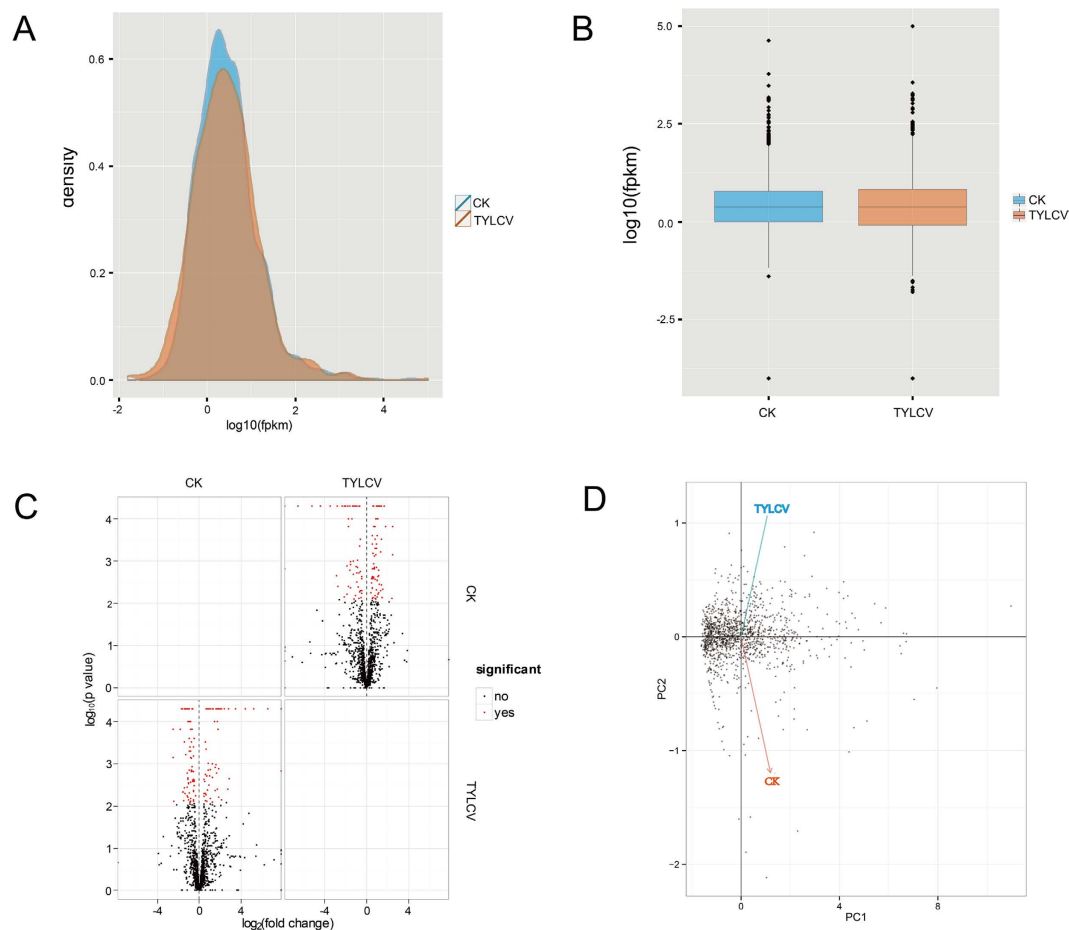


Figure 2. Global analysis of lncRNA expression in CK and TYLCV tomatoes. (A) Expression density differences among the samples. (B) Expression scatter matrix. (C) Comparison of lncRNA expression between the CK and TYLCV samples. Red dots denote differentially expressed lncRNAs whereas black dots denote lncRNAs that were expressed comparably in the CK and TYLCV samples; (D) Principal component analysis of lncRNA expression between the CK and TYLCV samples.

transcribed from the complementary DNA strand of their associated genes. These unannotated transcripts were used as the starting point to predict lncRNA candidates in tomatoes.

To identify lncRNAs, first we filtered out transcripts with lowly expressed abundance transcripts (fragments per kilobase of transcript per million mapped fragments (FPKM) < 0.5 for multiple-exon transcripts, and FPKM < 1 for single-exon transcripts), short transcripts (lengths < 200 nt) and long ORFs (length > 120 aa). Next, we evaluated the coding potential of the remaining transcripts using the Coding Potential Calculator (CPC)³⁹ and Coding-Non-Coding Index (CNCI)⁴⁰. Only transcripts with a score < 0 in both calculations were retained. We employed BLASTX against four protein databases (NR, KEGG, COGs, and Swiss-Prot) to exclude transcripts that might encode proteins. After these steps, we obtained 1,565 lncRNAs candidates, including 1,289 lncRNA candidates (Supplementary Table S1) and 276 lncNAT candidates (Supplementary Table S2).

Characteristics of tomato lncRNAs. Global inspection of the expression normalized to FPKM for all lncRNAs was performed using CummeRbund³⁸. Density and box plots of lncRNA expression ($\log_{10}\text{FPKM}$) revealed a normal overall distribution of the data points with little systematic bias among the CK and TYLCV lncRNA expression profiles (Fig. 2A,B)⁵¹. The volcano matrix plots revealed that a large number of lncRNAs were unaffected in the TYLCV-infected samples compared with the control samples, as indicated by the \log_2 fold change in the respective plots (Fig. 2C). Using principal component analysis, we found that the statistical relationship among the CK and TYLCV samples identified tight clustering of the TYLCV dataset compared with CK (Fig. 2D), indicating that the expression profiles of the lncRNAs were robust and highly reproducible.

Approximately 1074 lncRNAs were expressed in both the CK and TYLCV samples, but we also found 289 and 202 lncRNAs that were specifically expressed in the CK and TYLCV samples, respectively

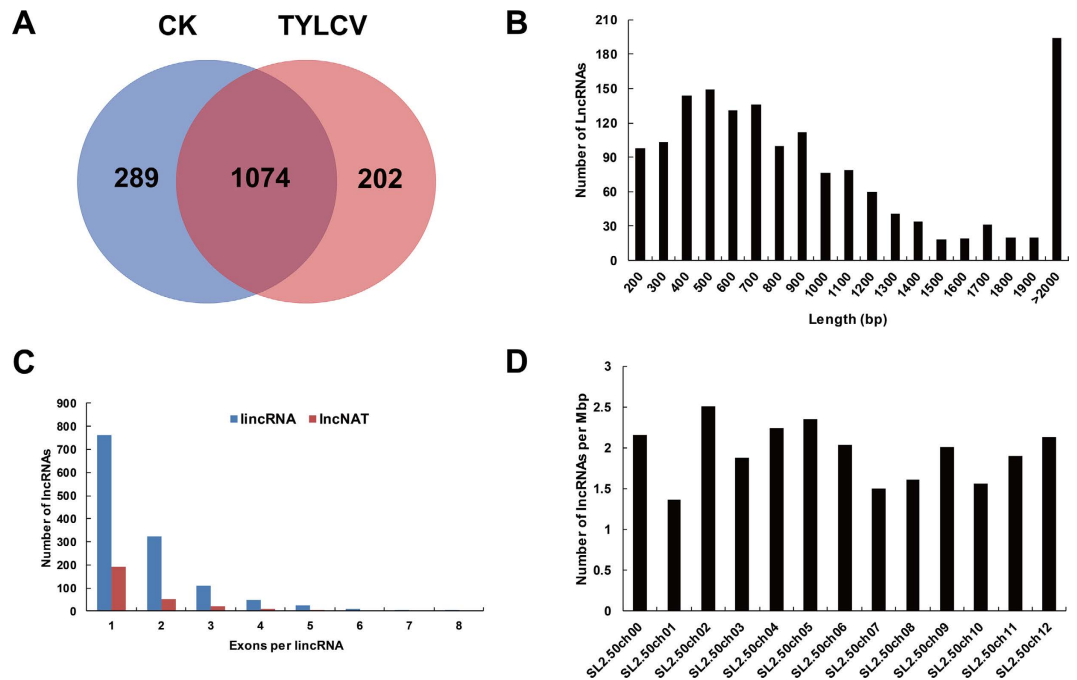


Figure 3. Characteristics of tomato lincRNAs. (A) A Venn diagram showing lincRNAs that are commonly expressed in the CK and TYLCV samples as well as those specifically expressed under one treatment but not the other. (B) Length distribution of 1573 lincRNAs. (C) Distribution of exon numbers of lincRNAs. (D) The density of lincRNAs on different tomato chromosomes.

(Fig. 3A). The lengths of the lincRNAs ranged from 201 to 5903 bp, but more than 67% of the lincRNAs were between 200 and 2000 bp in length (Fig. 3B). Approximately 60% of the tomato lincRNAs had one exon and 40% had multiple-exons (Fig. 3C). We examined the distribution of lincRNAs on the tomato chromosomes (Tomato Genome Sequence Build SL2.50) and found that it was uneven. The chromosome SL2.50ch02 had the highest lincRNA density, with 2.51 lincRNAs per 1 Mbp of nucleotides, whereas the chromosome SL2.50ch01 had the lowest density (1.37 lincRNAs per 1 Mbp of nucleotides).

Differential expression of tomato lincRNAs in response to TYLCV infection. To identify differentially expressed tomato lincRNAs between the CK and TYLCV samples, lincRNAs with a greater than 1.5-fold expression change and p -value < 0.01 were considered to be differentially expressed. A total of 529 lincRNAs were differentially expressed between the two samples. We also found more differentially expressed lincRNAs (33.6%) than mRNAs (10%) under TYLCV infection. These observations indicated that lincRNAs might have a markedly differential expression pattern compared to protein-coding genes in response to TYLCV infection.

To validate the differentially expressed lincRNA candidates, nine were randomly selected from the list of significantly regulated lincRNAs for experimental validation and expression profiling by qRT-PCR. These lincRNAs included slylnc0048, slylnc0049, slylnc0483, slylnc0531, slylnc0934, slylnc0476, slylnc0475, slylnc0673 and slylnc1052. As expected from the RNA-seq expression pattern, the qRT-PCR results mirrored of the RNA-seq data as the expression of slylnc0048, slylnc0049, slylnc0483, slylnc0531 and slylnc0934 increased substantially from 1.9 to 62.16 fold after TYLCV inoculation (Fig. 4). The expression levels of slylnc0476, slylnc0475, slylnc0673 and slylnc1052 in the TYLCV samples was suppressed to less than 0.5-fold compared with the CK samples. Additionally, the fold change in the lincRNA expression levels of the qRT-PCR and RNA-seq were closely correlated ($R^2 = 0.71$, $P < 0.05$) (Supplementary Fig. S2). These results indicated that these lincRNAs were likely to play roles in response to TYLCV infection.

Tomato lincRNAs involved in TYLCV infection. We selected lincRNAs slylnc0049 and slylnc0761 (which were significantly up-regulated by TYLCV infection) for functional characterization. A TRV vector carrying fragments of slylnc0049 was injected into plants by agroinfiltration at the cotyledon stage. One month after agroinfiltration, the success of the TRV silencing system was confirmed by the appearance of pTRV-PDS. The qRT-PCR assay also showed that the expression level of slylnc0049 after silencing decreased by more than 50% compared with the negative control (Fig. 5A). Total genomic DNA of TYLCV-infected tomatoes was extracted for the detection of virus accumulation after VIGS. Quantitative PCR revealed that TYLCV was barely detectable in the TRV empty vector control plants with cycle threshold (Ct) values of 30 (Supplementary Table S5). By contrast, TYLCV accumulation

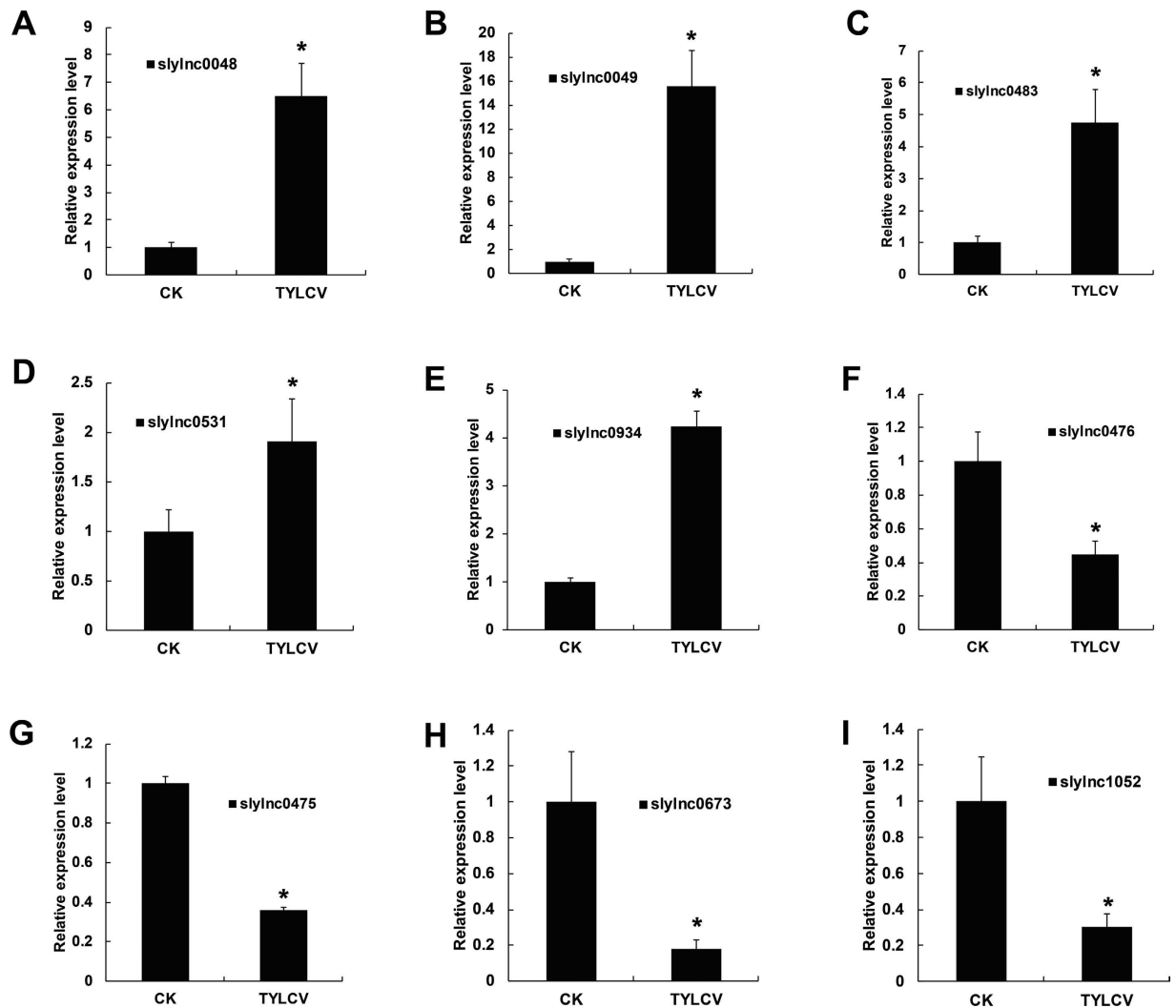


Figure 4. Confirmation of the expression patterns of differentially expressed lncRNAs using quantitative RT-PCR. Tomato α -tubulin (Solyc04g077020.2) was used as an internal reference. Error bars represented the standard error of three biological replicates. Asterisks indicate significant differences by Student's t test ($P < 0.05$).

exceeded 200-fold in the VIGS-treated tomato plants compared with the negative control (Fig. 5B). No disease symptoms of leaf curling and yellowing were observed in VIGS-treated tomato plants (Fig. 5C). VIGS was also performed with lncRNA slylnc0761 and the amount of virus accumulation in the slylnc0761-VIGS plants was six-fold higher than the level in the control (Supplementary Fig. S3). These findings indicated that the tomato lncRNAs were involved in the response to TYLCV infection and might perform some previously unknown function in the TYLCV regulatory network.

To identify the transcription start and end point of slylnc0049 and slylnc0761, 5' and 3' rapid amplification of cDNA ends (RACE) experiments were performed on mRNA from leaves of CLN2777A. As expected, the 5' and 3' RACE products of slylnc0761 were 214bp and 205bp (Supplementary Fig. S4A and S4B, Supplementary Fig. S5A and S5B), respectively, the size predicted from our RNA-seq results. However, the 5' and 3' RACE products of slylnc0049 were 714bp and 782bp, much longer than the ~200bp expected from the RNA-seq results (Supplementary Fig. S4C and S4D, Supplementary Fig. S5C and S5D). These results indicated that slylnc0049 had longer transcript than annotated by our bioinformatics analysis.

Tomato lncRNAs as putative targets of miRNAs. Plant lncRNAs may function as competing endogenous RNAs (ceRNAs) by binding to specific miRNAs via target mimicry to protect the miRNA targets^{11,22,41}. We predicted a similar mimic relationship between some tomato lncRNAs and miRNAs using the psMimic algorithm⁴¹. Two of the identified tomato lncRNAs (slylnc0195 and slylnc1077) were predicted to be 'decoys' for the conserved miRNAs, miR166 and miR399, respectively (Fig. 6A and

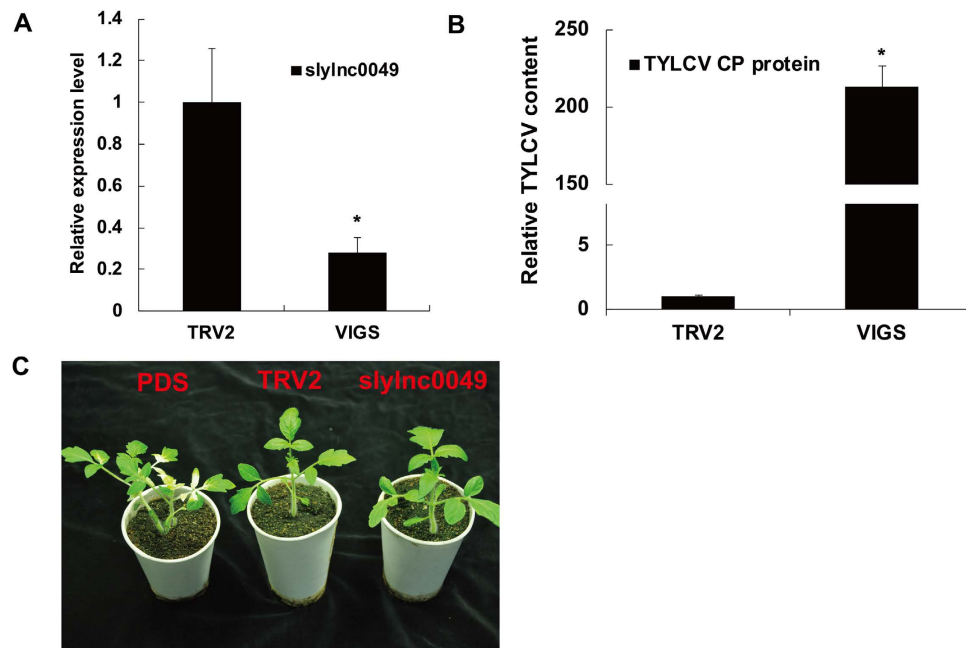


Figure 5. Validation of differentially expressed lncRNAs with virus-induced gene silencing. (A) Relative expression levels of slylnc0049 using real-time RT-PCR analysis in the VIGS-treated tomato plants 20 days after agroinfiltration with TRV2 vectors. Tomato α -tubulin (Solyc04g077020.2) was used as an internal reference. Error bars represented standard errors of three biological replicates, and asterisks indicate significant differences based on the Student's t test ($P < 0.05$). (B) TYLCV accumulation in the slylnc0049 silenced plants was estimated from the total genomic DNA by quantitative RT-PCR. Values were normalized using the tomato α -tubulin (Solyc04g077020.2) as an internal reference. Error bars represented standard errors of three biological replicates and asterisk indicates significant difference based on the Student's t test ($P < 0.05$). (C) Cotyledon agroinfiltration of TRV vectors was performed in the resistant tomato at the cotyledon stage. Tomato plants treated with the phytoene desaturase (PDS) gene silencing constructs pTRV1 and pTRV2-PDS showed bleached areas in the leaflets (left). Plants treated with the pTRV1 and pTRV2 vectors showed the normal phenotype (middle). Resistant plantlets treated with the slylnc0049 gene silencing constructs pTRV1 and pTRV2-slylnc0049 showed the normal phenotype (right).

Fig. S2A). We investigated the correlation between the expression of these lncRNAs, miR166a and its miRNA targets after TYLCV inoculation by qRT-PCR. As expected from the negative correlation, the expression of slylnc0195 was dramatically increased after TYLCV inoculation, whereas miR166a was down-regulated (Fig. 6B,C). Using psRobot we predicted the targets of miR166a using moderate parameters⁴⁵. Among the many predicted targets of miR166a, we concentrated on the class III homeodomain-leucine zipper (class III HD-Zip) family genes because they encoded transcription factors known to play a role in plant development^{52,53}. Using qRT-PCR, we found that the expression of five class III HD-Zip transcription factor genes targeted by miR166a increased after TYLCV inoculation (Fig. 6D). We also used VIGS to silence slylnc0195 and then analyzed the expression of the lncRNAs and miR166a targets. The expression of slylnc0195 was dramatically suppressed to approximately 40% after VIGS (Fig. 6E), and the amount of virus accumulation in slylnc0195-VIGS plants was increased 70-fold compared with the control (Fig. 6F). Notably, the class III HD-Zip genes of the miR166a targets showed decreased abundance (Fig. 6G). These data suggest the existence of a specific crosstalk between slylnc0195 and the class III HD-Zip through competitive miR166a binding. Furthermore, similar expression correlation patterns were observed between the slylnc1077 and sly-miR399 targets (Supplementary Fig. S6B and S6C). These results indicate that the miRNA-lncRNA pairs might be important novel regulatory components in tomato TYLCV resistance.

Next, we used a transient agroinfiltration assay to test whether slylnc0195 was functional. We constructed the expression vector pCAMBIA2301 containing slylnc0195 and overexpressed it in the leaves of *N. benthamiana*. The sequences of miR166 and its targets were conserved between the tomato and *N. benthamiana*. The slylnc0195 dramatically increased the mRNA abundance of the corresponding miR166 targets in their transiently expressed leaves 2 days after agrobacteria infiltration, suggesting that slylnc0195 indeed inhibited the functions of the corresponding miR166 (Fig. 7).

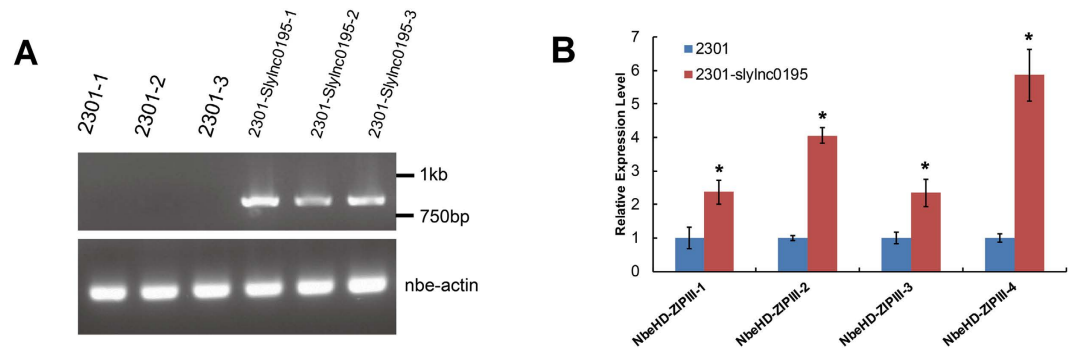


Figure 7. Functional analysis of slylnc0195 by the transient agroinfiltration assay. (A) Detection of slylnc0195 expression in the control vector pCOMBIA2301 and pCOMBIA2301-slylnc0195 infiltrated leaves by semi-quantitative RT-PCR. (B) Quantitative RT-PCR analysis of miR166 target genes of *N. benthamiana* in the control vector (2301) and 2301-slylnc0195 vector infiltrated leaves. Error bars represented standard errors of three biological replicates, and asterisks indicate significant differences based on the Student's t test ($P < 0.05$).

on the transfer of resistance genes from wild tomato relatives. Resistance gene *Ty-2* was derived from *S. habrochaites* f. *glabratum* accession “B6013”⁵⁶ and was previously fine mapped to the long arm of chromosome 11 near markers UP8 (51.344 Mb) and M1 (51.645 Mb)³². Using genome-wide analysis and VIGS, some basic/helix-loop-helix (bHLH) transcription factors were shown to be involved in TYLCV infection³⁴. However, *Ty-2* has not been cloned and the regulatory pathways that mediate the resistance to TYLCV are far from being illustrated. The recent discovery of lncRNAs has opened up a new field in the investigation of novel regulatory pathways. Although an increasing number of reports demonstrated that lncRNAs functioned in gene regulation in mammals, lncRNAs were reported to play roles in regulation of the biotic and abiotic stress responses in only a few plants^{20,22}. Recently, a comprehensive set of 3679 putative lncRNAs from wild-type and ripening inhibitor (*rin*) mutant were identified using paired-end strand-specific RNA sequencing. Many lncRNAs showed significantly differential expression in the *rin* mutant. Furthermore, the down-regulation of the expression of some novel lncRNAs in the wild-type tomato fruit induced an obvious delay in fruit ripening¹⁴. In this study, we used a strand-specific RNA-seq approach to investigate transcriptomic changes in response to TYLCV infection in tomatoes and systematically identified and analyzed the tomato lncRNAs associated with TYLCV resistance. Moreover, we identified several lncRNAs that were specifically or differentially expressed between mock and TYLCV inoculation, and found that some lncRNAs that acted as miRNA mimics participated in the tomato TYLCV resistance regulatory process.

Most lncRNAs identified in plants were most likely transcribed by RNA polymerase II (pol II) in a process, that harbored characteristic features of mRNA, such as the addition of a 5'-7^mGTP-cap and 3'-polyadenylated tail⁵⁷. Additionally, a combination of RNA-sequencing approaches with mutants with defects in pol IV-dependent transcription was used to identify more than 20,000 pol IV-dependent lncRNAs⁵⁸. In contrast to the pol II and pol IV products, pol V transcripts are 5'-triphosphorylated or 7^mGTP-capped, but do not have 3'-poly(A) tails⁵⁸. Zhu *et al.* discovered 3679 lncRNAs from wild-type tomatoes and ripening mutant fruits, including lncRNAs both with and without a poly(A) tail¹⁴. Due to the poly(A) RNA enrichment step in library preparation, all of the lncRNAs in our study might be evolved from pol II, and thus the number was less than the previous study.

lncRNAs have been shown to be involved in the response to several biotic and abiotic stresses^{20–23}. In *Arabidopsis*, many lncRNAs could be reproducibly detected by different technical platforms, including strand-specific tiling arrays, Agilent custom expression arrays, strand-specific RNA-seq, and qRT-PCR experiments. More than 1000 NAT pairs were regulated by light in a spatial and developmental-specific manner²³. Additionally, fifteen lncRNAs responsive to *Fusarium oxysporum* infection were identified using the strand-specific RNA-seq approach. Only one sense-antisense pair was observed to be co-regulated²⁰. By analyzing the poly(A)⁺ and poly(A)⁻ RNA-seq of *Arabidopsis* under four stress conditions, a total of 245 poly(A)⁺ and 58 poly(A)⁻ lncRNAs were identified to be differentially expressed. Many lncRNAs exhibited more stress-specific expression than coding genes, particularly for the poly(A)⁻ lncRNAs²¹. In our study, we found 1565 lncRNAs including some lncRNAs by strand-specific RNA-seq profiling of CK and TYLCV inoculation samples. A higher percentage of lncRNAs exhibited TYLCV-specific expression than coding genes, particularly for the lncRNAs. This finding is similar to previous studies showing that lncRNAs have highly specific temporal and spatial expression profiles^{59,60}.

Endogenous target mimicry (eTM) is a recently identified regulatory mechanism for miRNA functions in plants in which the decoy RNAs bind to miRNAs via complementary sequences and subsequently block the interaction between miRNAs and their authentic targets^{41,61}. Some lncRNAs that contain miRNA-binding sites have been shown to regulate corresponding miRNA target genes by competing

specifically for the shared miRNAs. For example, two reproduction-related rice lncRNAs were confirmed to be target mimics of miR160 and miR164¹¹. In our study, we identified lncRNAs that might act as eTMs for conserved miRNAs in tomatoes. After experimental verification, two of these TYLCV-responsive lncRNAs were confirmed to be target mimics of miR166 and miR399. miR166 is a well-studied plant miRNA involved in various aspects of plant development⁶², abiotic stresses such as drought and cold stresses^{63,64}, and biotic stress such as fungal invasion^{65,66}. In wheat, miR166a and miR166d were significantly altered after powdery mildew infection⁶⁵. In *Populus trichocarpa* plantlets, three members of miR166 were upregulated by induced with the poplar stem canker pathogen, *Botryosphaeria dothidea*⁶⁶. miR166 negatively regulated its target class III homeodomain leucine-zipper (HD-ZIP III or class III HD-ZIP) transcription factors^{62,67}, which were demonstrated to be important for lateral root development, axillary meristem initiation, leaf polarity and abiotic stresses such as salt and drought^{68–70}. In our study, we found that slylnc0195 acted as a ‘sponge’ for sly-miR166 which significantly reduced by TYLCV infection and regulated the expression of sly-miR166 targets, including the class III HD-ZIP transcription factors, by competing specifically for shared miRNAs. In *Arabidopsis*, miR166 and its eTM ath-eTM-166-1 had a bulge in the middle of the sequence, and transgenic plants overexpressing ath-eTM-166-1 had abnormal rosette leaf shapes. The expression of miR166 targets was significantly increased in the overexpressing plants⁴¹. Therefore, these results provide strong evidence that the eTMs of miR166 are functional target mimics that not only play important roles in plant development but also regulate tomato TYLCV resistance.

We found that slylnc1077 and sly-miR399 could function as mimics and that the expression pattern of the sly-miR399 target was similar to that of slylnc1077. Thus, slylnc1077 acts as an eTM for sly-miR399, suggesting that slylnc1077 might be involved in TYLCV infection response networks. The eTMs between lncRNA and miR399 was also found in a previous study¹⁴. These data suggest that these lncRNAs are functional candidates involved in the TYLCV signaling pathways. More elaborate experiments such as overexpressing or RNAi transgenic lncRNAs *in vivo* need to be performed to elucidate the detailed mechanisms.

References

- Rymarquis, L. A., Kastenmayer, J. P., Hüttenhofer, A. G. & Green, P. J. Diamonds in the rough: mRNA-like non-coding RNAs. *Trends in Plant Science* **13**, 329–334 (2008).
- Brosnan, C. A. & Voinnet, O. The long and the short of noncoding RNAs. *Current Opinion in Cell Biology* **21**, 416–425 (2009).
- Cuperus, J. T., Fahlgren, N. & Carrington, J. C. Evolution and Functional Diversification of MIRNA Genes. *Plant Cell* **23**, 431–442 (2011).
- Chen, X. Small RNAs in development – insights from plants. *Current Opinion in Genetics & Development* **22**, 361–367 (2012).
- Cheetham, S. W., Gruhl, F., Mattick, J. S. & Dinger, M. E. Long noncoding RNAs and the genetics of cancer. *Br J Cancer* **108**, 2419–2425 (2013).
- Leung, A. *et al.* Novel Long Noncoding RNAs Are Regulated by Angiotensin II in Vascular Smooth Muscle Cells. *Circulation Research* **113**, 266–278 (2013).
- Ng, S.-Y., Bogu, Gireesh K., Soh, Boon S. & Stanton, Lawrence W. The Long Noncoding RNA RMST Interacts with SOX2 to Regulate Neurogenesis. *Molecular Cell* **51**, 349–359 (2013).
- Liu, J. *et al.* Genome-Wide Analysis Uncovers Regulation of Long Intergenic Noncoding RNAs in Arabidopsis. *Plant Cell* **24**, 4333–4345 (2012).
- Boerner, S. & McGinnis, K. M. Computational Identification and Functional Predictions of Long Noncoding RNA in Zea mays. *PLoS ONE* **7**, e43047 (2012).
- Li, L. *et al.* Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biology* **15**, R40 (2014).
- Zhang, Y.-C. *et al.* Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. *Genome Biology* **15**, 512 (2014).
- Wen, J., Parker, B. J. & Weiller, G. F. In Silico Identification and Characterization of mRNA-Like Noncoding Transcripts in *Medicago truncatula*. *In Silico Biology* **7**, 485–505 (2007).
- Hao, Z. *et al.* Genome-Wide Identification, Characterization and Evolutionary Analysis of Long Intergenic Noncoding RNAs in Cucumber. *PLoS ONE* **10**, e0121800 (2015).
- Zhu, B. *et al.* RNA sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening. *Journal of Experimental Botany* **66**, 4483–4495 (2015).
- Swiezewski, S., Liu, F., Magusin, A. & Dean, C. Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature* **462**, 799–802 (2009).
- Tibor, C., I., Q. J., Qianwen, S. & Caroline, D. Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *PNAS* **111**, 16160–16165 (2014).
- Heo, J. B. & Sung, S. Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. *Science* **331**, 76–79 (2011).
- Liu, J., Wang, H. & Chua, N.-H. Long noncoding RNA transcriptome of plants. *Plant Biotechnology Journal* **13**, 1–10 (2015).
- Wang, Kevin C. & Chang, Howard Y. Molecular Mechanisms of Long Noncoding RNAs. *Molecular Cell* **43**, 904–914 (2011).
- Zhu, Q.-H., Stephen, S., Taylor, J., Helliwell, C. A. & Wang, M.-B. Long noncoding RNAs responsive to Fusarium oxysporum infection in Arabidopsis thaliana. *New Phytologist* **201**, 574–584 (2014).
- Di, C. *et al.* Characterization of stress-responsive lncRNAs in Arabidopsis thaliana by integrating expression, epigenetic and structural features. *Plant Journal* **80**, 848–861 (2014).
- Shuai, P. *et al.* Genome-wide identification and functional prediction of novel and drought-responsive lincRNAs in *Populus trichocarpa*. *Journal of Experimental Botany* **65**, 4975–4983 (2014).
- Wang, H. *et al.* Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. *Genome Research* **24**, 444–453 (2014).
- Scholthof, K.-B. G. *et al.* Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology* **12**, 938–954 (2011).
- Glick, E., Levy, Y. & Gafni, Y. The viral etiology of tomato yellow leaf curl disease—a review. *Plant Protect Sci* **45**, 81–97 (2009).
- Anbinder, I. *et al.* Molecular dissection of Tomato leaf curl virus resistance in tomato line TY172 derived from Solanum peruvianum. *Theoretical and Applied Genetics* **119**, 519–530 (2009).

27. Ji, Y., Scott, J. W. & Schuster, D. J. Toward Fine Mapping of the Tomato Yellow Leaf Curl Virus Resistance Gene Ty-2 on Chromosome 11 of Tomato. *HortScience* **44**, 614–618 (2009).
28. Ji, Y., Schuster, D. & Scott, J. Ty-3, a begomovirus resistance locus near the Tomato yellow leaf curl virus resistance locus Ty-1 on chromosome 6 of tomato. *Molecular Breeding* **20**, 271–284 (2007).
29. Ji, Y., Scott, J. W., Schuster, D. J. & Maxwell, D. P. Molecular Mapping of Ty-4, a New Tomato Yellow Leaf Curl Virus Resistance Locus on Chromosome 3 of Tomato. *Journal of the American Society for Horticultural Science* **134**, 281–288 (2009).
30. Verlaan, M. G. *et al.* The Tomato Yellow Leaf Curl Virus Resistance Genes Ty-1 and Ty-3 Are Allelic and Code for DFDGD-Class RNA-Dependent RNA Polymerases. *PLoS Genet* **9**, e1003399 (2013).
31. Butterbach, P. *et al.* Tomato yellow leaf curl virus resistance by Ty-1 involves increased cytosine methylation of viral genomes and is compromised by cucumber mosaic virus infection. *PNAS* **111**, 12942–12947 (2014).
32. Yang, X. *et al.* Fine mapping of the tomato yellow leaf curl virus resistance gene Ty-2 on chromosome 11 of tomato. *Molecular Breeding* **34**, 749–760 (2014).
33. Chen, T. *et al.* Comparative Transcriptome Profiling of a Resistant vs. Susceptible Tomato (*Solanum lycopersicum*) Cultivar in Response to Infection by Tomato Yellow Leaf Curl Virus. *PLoS ONE* **8**, e80816 (2013).
34. Wang, J. *et al.* Genome-wide analysis of bHLH transcription factor and involvement in the infection by yellow leaf curl virus in tomato (*Solanum lycopersicum*). *BMC Genomics* **16**, 39 (2015).
35. Zhang, H., Gong, H. & Zhou, X. Molecular characterization and pathogenicity of tomato yellow leaf curl virus in China. *Virus Genes* **39**, 249–255 (2009).
36. Wu, J. B., Dai, F. M. & Zhou, X. P. First Report of Tomato yellow leaf curl virus in China. *Plant Disease* **90**, 1359–1359 (2006).
37. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**, R36 (2013).
38. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protocols* **7**, 562–578 (2012).
39. Kong, L. *et al.* CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Research* **35**, W345–W349 (2007).
40. Sun, L. *et al.* Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Research* **41**, e166 (2013).
41. Wu, H.-J., Wang, Z.-M., Wang, M. & Wang, X.-J. Widespread Long Noncoding RNAs as Endogenous Target Mimics for MicroRNAs in Plants. *Plant Physiology* **161**, 1875–1884 (2013).
42. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research* **42**, D68–D73 (2014).
43. Xu, T. *et al.* Small RNA and degradome sequencing reveals microRNAs and their targets involved in tomato pedicel abscission. *Planta*, 1–22 (2015).
44. Luan, Y. *et al.* High-throughput sequencing reveals differential expression of miRNAs in tomato inoculated with *Phytophthora infestans*. *Planta* **241**, 1405–1416 (2015).
45. Wu, H.-J., Ma, Y.-K., Chen, T., Wang, M. & Wang, X.-J. PsRobot: a web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Research* **40**, W22–W28 (2012).
46. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402–408 (2001).
47. Varkonyi-Gasic, E., Wu, R., Wood, M., Walton, E. F. & Hellens, R. P. Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods* **3**, 12 (2007).
48. Liu, Y., Schiff, M. & Dinesh-Kumar, S. P. Virus-induced gene silencing in tomato. *Plant Journal* **31**, 777–786 (2002).
49. Liu, T. *et al.* Unconventionally secreted effectors of two filamentous pathogens target plant salicylate biosynthesis. *Nat Commun* **5**, 4686 (2014).
50. Wu, H.-J., Ma, Y.-K., Chen, T., Wang, M. & Wang, X.-J. PsRobot: a web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Research* **40**, W22–W28 (2012).
51. Dillies, M.-A. *et al.* A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings in Bioinformatics* **14**, 671–683 (2013).
52. Prigge, M. J. *et al.* Class III Homeodomain-Leucine Zipper Gene Family Members Have Overlapping, Antagonistic, and Distinct Roles in Arabidopsis Development. *Plant Cell* **17**, 61–76 (2005).
53. Itoh, J.-I., Hibara, K.-I., Sato, Y. & Nagato, Y. Developmental Role and Auxin Responsiveness of Class III Homeodomain Leucine Zipper Gene Family Members in Rice. *Plant Physiology* **147**, 1960–1975 (2008).
54. Cohen, S. & Lapidot, M. Appearance and Expansion of TYLCV: a Historical Point of View. in *Tomato Yellow Leaf Curl Virus Disease* (ed. Czosnek, H.) 3–12 (Springer Netherlands, 2007).
55. Horowitz, R., Denholm, I. & Morin, S. Resistance to Insecticides in the TYLCV vector, Bemisia Tabaci. in *Tomato Yellow Leaf Curl Virus Disease* (ed. Czosnek, H.) 305–325 (Springer Netherlands, 2007).
56. Kalloo & Banerjee, M. K. Transfer of Tomato Leaf Curl Virus Resistance from *Lycopersicon hirsutum* f. *glabratum* to *L. esculentum*. *Plant Breeding* **105**, 156–159 (1990).
57. Kim, Y. J. *et al.* The role of Mediator in small and long noncoding RNA production in *Arabidopsis thaliana*. 814–822 (2011).
58. Li, S. *et al.* Detection of Pol IV/RDR2-dependent transcripts at the genomic scale in Arabidopsis reveals features and regulation of siRNA biogenesis. *Genome Research* **25**, 235–245 (2015).
59. Nam, J.-W. & Bartel, D. P. Long noncoding RNAs in *C. elegans*. *Genome Research* **22**, 2529–2540 (2012).
60. Pauli, A. *et al.* Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Research* **22**, 577–591 (2012).
61. Franco-Zorrilla, J. M. *et al.* Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet* **39**, 1033–1037 (2007).
62. Jung, J.-H. & Park, C.-M. MIR166/165 genes exhibit dynamic expression patterns in regulating shoot apical meristem and floral development in Arabidopsis. *Planta* **225**, 1327–1338 (2007).
63. Valiollahi, E., Farsi, M. & Kakhki, A. Sly-miR166 and Sly-miR319 are components of the cold stress response in *Solanum lycopersicum*. *Plant Biotechnology Reports* **8**, 349–356 (2014).
64. Kantar, M., Lucas, S. & Budak, H. miRNA expression patterns of *Triticum dicoccoides* in response to shock drought stress. *Planta* **233**, 471–484 (2011).
65. Xin, M. *et al.* Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.). *BMC Plant Biology* **10**, 123 (2010).
66. Zhao, J.-P., Jiang, X.-L., Zhang, B.-Y. & Su, X.-H. Involvement of microRNA-Mediated Gene Expression Regulation in the Pathological Development of Stem Canker Disease in *Populus trichocarpa*. *PLoS ONE* **7**, e44968 (2012).
67. Sakaguchi, J. & Watanabe, Y. miR165/166 and the development of land plants. *Development, Growth & Differentiation* **54**, 93–99 (2012).

68. Boualem, A. *et al.* MicroRNA166 controls root and nodule development in *Medicago truncatula*. *Plant Journal* **54**, 876–887 (2008).
69. Hawker, N. P. & Bowman, J. L. Roles for Class III HD-Zip and KANADI Genes in Arabidopsis Root Development. *Plant Physiology* **135**, 2261–2270 (2004).
70. Chen, X. *et al.* Genome-Wide Analysis of Soybean HD-Zip Gene Family and Expression Profiling under Salinity and Drought Treatments. *PLoS ONE* **9**, e87156 (2014).

Acknowledgements

We thank Dr. Yinghua Ji (Institute of Plant Protection, Jiangsu Academy of Agricultural Sciences) for providing the TYLCV infection clone. And this work was supported by Natural Science Foundation of Jiangsu Province (BK20150543), National Natural Science Foundation of China (No. 31471873, No. 31301775, No. 31371930) and Jiangsu Key Laboratory for Bioresources of Saline Soils (JKLBS2014004).

Author Contributions

J.Y.W., R.Y.L., B.L.Z. and W.G.Y. performed the data analysis and drafted the manuscript. J.Y.W., R.Y.L., T.Z.C., Y.W.Y. and T.L.L. participated in the analysis of the data. X.Y., N.M. and X.L. performed the experiments. All authors have read and approved the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Wang, J. *et al.* Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection. *Sci. Rep.* **5**, 16946; doi: 10.1038/srep16946 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>