Analysis of Competing Endogenous RNAs and MicroRNAs in Tea (*Camellia sinensis*) Leaves During Infection by the Leaf Spot Pathogen *Pestalotiopsis trachicarpicola*

Yubin Yang,^{1,2} Qiaoxiu Yin,¹ Changlong Qiu,¹ Zhongqiu Xia,^{1,2} Hongke Huang,^{1,2} Chen Huang,¹ Xinyue Jiang,¹ Yuanyou Yang,¹ Delu Wang,^{3,†} and Zhuo Chen^{1,†}

¹ Key Laboratory of Green Pesticide and Agricultural Bioengineering, Ministry of Education, Guizhou University, Guiyang, Guizhou 550025, China

² College of Tea Science, Guizhou University, Guiyang, Guizhou 550025, China

³ College of Forestry, Guizhou University, Guiyang, Guizhou 550025, China

Genome Announcement

Pestalotiopsis trachicarpicola Y. M. Zhang & K. D. Hyde is a plant-pathogenic fungus that can cause leaf spot on tea (*Camellia sinensis*), seriously decreasing the production and quality of tea leaves. Competing endogenous RNAs (ceRNAs) include messenger RNAs (mRNAs), circular RNAs (circRNAs), and long noncoding RNAs (lncRNAs), which can regulate gene expression through binding to response elements in microRNAs (miRNAs), which, in turn, operate posttranscriptional regulation of target mRNAs. In the current study, ceRNAs and miRNAs originating from infected and uninfected leaves of tea (*C. sinensis* var. *sinensis* cv. Fuding-dabaicha) were sequenced and analyzed, and the interaction between individual ceRNAs and miRNAs was predicted, with target genes being identified. Of the ceRNA interaction pairs, 31, 24, six, and two were predicted in the form of miRNA-mRNA, miRNA-lncRNA, miRNA-circRNA, and mRNA-lncRNA interactions, respectively. Three types of interactions were predicted to occur in the regulation pairs of lncRNA/circRNA-miRNA-mRNA: i) circRNA-miRNA-mRNA (type 1 interactions), ii) lncRNA-miRNA-mRNA (type 2 interactions), and iii) circRNA-lncRNA-miRNA-mRNA (type 3 interactions). Some interaction pairs in lncRNA/circRNA-miRNA-mRNA in infected tea leaves ultimately changed the abundance of mRNAs, such as TEA015316.1:TEA015316.1, TEA020085.1:TEA020085.1, TEA027239.1:TEA027239.1, and TEA001447.1:TEA001447.1. Analysis of the types and expression abundance of ceRNAs and miRNAs originating from tea leaves during infection by *P. trachicarpicola*, and their interactive relationship will provide a valuable resource for future research into this host-pathogen interaction.

Tea [*Camellia sinensis* (L.) O. Kuntze] is a perennial, evergreen, small tree or shrub and is an important cash crop grown in plantations in tropical and subtropical regions of Asia (Li et al. 2019b). Due to its perennial nature, widespread area of cultivation, and the complex, ecological environment of tea plantations, tea suffers from a number of diseases, largely

Funding

This research was funded by National Key Research Development Program of China (grant number 2017YFD0200308), and its Post-Subsidy project (grant number 20185262), the National Natural Science Foundation of China (grant numbers 21977023 and 31860515), China Agriculture Research System of MOF and MARA (09), and the Program of Introducing Talents to Chinese Universities (111 Program, grant number D20023).

Keywords

Camellia sinensis, competing endogenous RNA, microRNA, *Pestalotiopsis trachicarpicola*, tea leaf spot

†Corresponding authors: D. L. Wang; dlwang@gzu.edu.cn, and Z. Chen; gychenzhuo@aliyun.com

*The e-Xtra logo stands for “electronic extra” and indicates there are supplementary tables published online.

The author(s) declare no conflict of interest.

Accepted for publication 16 February 2022.

Copyright © 2022 The Author(s). This is an open access article distributed under the CC BY-NC-ND 4.0 International license.

caused by fungal phytopathogens (Pandey et al. 2021). These diseases mainly occur on the leaves of tea and reduce the production and quality of tea leaves (Pandey et al. 2021).

In recent years, tea has been grown on a very large scale in Guizhou Province, China. The area of the tea plantations in Guizhou Province reached 350,000 square hectometers in 2017, making this province the top tea-growing region in China (Bao et al. 2019; Li et al. 2019a; Wang et al. 2020; Yin et al. 2021). Tea plantations in Guizhou Province combine the characters of high altitude (from 800 to 1,800 m above sea level), lower temperatures during the early spring, cold spells in late spring, and a rainy climate. As a consequence, tea suffers from a number of important diseases caused by pathogenic fungi, which are widespread and cause serious damage (Wang et al. 2020; Yin et al. 2021). In previous studies, the fungal pathogens causing tea leaf spot have been identified as *Lasiodiplodia theobromae* (Li et al. 2019a), *Epicoccum sorghinum* (Bao et al. 2019), *Epicoccum nigrum* (Yin et al. 2022), *Didymella bellidis* (Wang et al. 2020), and *Alternaria longipes* (Yin et al. 2021). In addition, we found that *P. trachicarpicola* can also cause an economically important leaf spot on tea plantations in Yuqing County, Guizhou Province, China. Early symptoms of infection of tea leaves by *P. trachicarpicola* are light brown spots that gradually increase in size and develop into brown, scattered, elliptical, or irregular lesions. The lesions later become grayish at the center and become surrounded by an obvious bright yellow halo. Adjacent lesions can coalesce into larger lesions, which cover the margin of the leaves. For the leaf spots with typical symptoms, the disease incidence was estimated to range between 75 and 83%. The disease severity was estimated to range from 75 to 86, depending on the field. In tea plantations with severe disease, the loss of yield can reach more than 68% (Li et al. 2021).

P. trachicarpicola can also cause leaf spot of *Trachycarpus fortunei* (Maharachchikumbura et al. 2012; Zhang et al. 2012), fruit rot and trunk diseases of grapevine (Jayawardena et al. 2015), leaf spot of *Eucommia ulmoides* (Li et al. 2018), mango leaf spot (Shu et al. 2020), twig blight of *Pinus bungeana* (Qi et al. 2021), and gray blight disease of tea (Tsai et al. 2021).

The components of ceRNAs, namely, circRNAs, lncRNAs, and mRNAs, can bind with specific sequence elements of miRNAs, resulting in posttranscriptional regulation of the expression of genes during growth, development, and environmental response (Hansen et al. 2013; Tay et al. 2014). ceRNAs play a role in the interaction between plant hosts and fungal pathogens (Cui et al. 2020; Fan et al. 2020; Jiang et al. 2020; Sun et al. 2020). For instance, infection of rice by *Magnaporthe oryzae* can induce differential expression of circRNAs originating from the host, and a large number of circRNAs have been shown to be involved in the disease defense response of the rice host (Fan et al. 2020). The present study identified tea leaf ceRNAs that were differentially expressed (DE) in response to infection and determined potential target genes that could be affected by interactions between these DE ceRNAs and the mRNAs of the target genes.

In this study, a single-spore isolate of *P. trachicarpicola* ACCC39730 was cultured on potato dextrose agar (PDA) at 25°C in the dark, to form colonies for use as inoculum after incubation for 10 days. Three-year-old bushes of *C. sinensis* var. *sinensis* cv. Fudingdabaicha were grown in the greenhouse under day and night conditions of 25°C and 20°C, respectively, and a relative humidity of 70 to 80%. Detached tea bush twigs were placed into 50-ml conical flasks (one twig per flask) containing 30 ml of sterile water, and the third leaf on each tea twig was inoculated. Four 6-mm diameter holes were punctured into the adaxial leaf surface, two on each side of the main vein of each leaf, using a sterile syringe needle. Mycelial plugs (6 mm diameter) were cut from the edge of a 10-day-old single-spore colony of *P. trachicarpicola* ACCC39730 on PDA, using a sterile cork borer, and were then placed, mycelial face downward, on the inoculation sites, one plug per site. For the uninoculated controls, plugs of sterile PDA were used instead. Damp cotton was placed over the plugs on the inoculation sites, and the tea twigs were then covered, using self-sealing plastic bags (20 cm × 30 cm) to retain high humidity, for 48 h. The tea twigs were then placed in the plant light incubator (NingboJiangnan Instrument Factory, Ningbo City, China) and were maintained under conditions of 25°C in the day and 20°C at night (200 $\mu\text{mol m}^{-2}\cdot\text{s}^{-1}$) and a relative humidity of 70 to 80%. At 48 h postinoculation, the inoculated tea leaves showed symptoms but there were no symptoms on the control leaves. The experiment was set up with two different samples as follows. Control (uninfected) tea leaves were punctured by sterilized syringe needles, and tissue samples were harvested at 0 days postinoculation. Treatment (infected) tea leaves were punctured by sterilized syringe needles

and were inoculated with the mycelial plugs and tissue samples were harvested at 1 day postinoculation. Each control and treatment sample contained 15 technical replicates, each replicate consisting of 15 tea twigs. The third leaf from each of the 15 twigs was pooled to carry out total RNA isolation, representing one biological replicate. The tea twigs were randomly assigned to the control or treatment groups, and each experiment was carried out with three independent biological replicates.

Fifteen tea leaves from 15 tea twigs, representing one biological replicate of a sample, were pooled and were ground into a fine powder in the presence of liquid nitrogen. Then, 100 mg of the fine powder was used to extract total RNA, using TRIzol reagent (Invitrogen, Carlsbad, CA, U.S.A.), following manufacturer procedures. The RNA amount and purity of each sample were quantified and assessed using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, U.S.A.) and an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, U.S.A.), with the samples required to meet a threshold RNA integrity number >7.0.

The cDNA libraries of all RNA species were prepared using the TruSeq Stranded Total RNA HT sample prep kit (Illumina, San Diego, CA, U.S.A.) and the TruSeq Small RNA sample prep kit (Illumina). Following the manufacturer-recommended protocol, ceRNAs were sequenced on an Illumina HiSeq 4000 at Lc-Bio Technologies Co., Ltd. (Hangzhou, China). Cutadapt software was used to remove the low-quality reads, including adapters, low-quality bases, and undetermined bases (Martin 2011). Then, the sequence quality was verified, using FastQC software to obtain clean reads for subsequent downstream analyses.

The sequence analysis was divided into four parts. First, for the mRNA component, clean reads were aligned to the reference genome of *C. sinensis* var. *sinensis* using HISAT software (version 2.0.4) (Kim et al. 2015). The mapped reads of each sample were assembled, then, their abundance was quantified by calculating the value of fragments per kilobase of transcript per million mapped reads at the level of $\log_2$ (fold change) > 1 or < -1 and with statistical significance ($P < 0.05$), using StringTie software (version stringtie-1.3.4d) (Frazee et al. 2015; Pertea et al. 2015) and Ballgown software package in R (Frazee et al. 2015). The mRNAs targeted by the DE ceRNAs were annotated in the Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases, using BLAST (Buchfink et al. 2015; Love et al. 2014).

Second, for the lncRNA component, sequence quality was verified using FastQC. Valid reads were used to map reads to the genome of *C. sinensis* var. *sinensis*, using the Bowtie 2 package in R (Langmead and Salzberg 2012) and TopHat2 (Kim et al. 2013). The mapped reads of each sample were assembled using StringTie software (version stringtie-1.3.4d) (Pertea et al. 2015), then, the transcriptomes from all replicates of both the infected and control samples were merged, to construct a comprehensive transcriptome, using the Perl program. After the final transcriptome was generated, the abundance of each transcript was estimated using StringTie software (Pertea et al. 2015) and the Ballgown package in R (Frazee et al. 2015). All unrelated transcripts, following identification according to the following criteria, were removed: transcripts that overlapped with known mRNAs or with novel mRNAs with coding potential, transcripts that were shorter than 200 bp, and transcripts with coverage of reads <3. The remaining transcripts with coding potential were used for prediction, using the Coding Potential Calculator (CPC) software (version 0.9-r2) (Kong et al. 2007) and Coding-Non-Coding Index (CNCI) software (version 2.0) (Sun et al. 2013), with transcripts with a CPC score <-1 and a CNCI score <0 being removed. The remaining transcripts were considered to be lncRNAs. The expression level of lncRNAs was calculated using StringTie software (Pertea et al. 2015; Trapnell et al. 2010) and the Ballgown package in R (Frazee et al. 2015).

Third, for the circRNA component, Cutadapt was used to remove the reads that contained adapter contamination, low-quality bases, and undetermined bases (Martin 2011). Then, sequence quality was verified using FastQC. Bowtie2 (version 2.0.0-beta4) and TopHat2 software were used to map reads to the genome of *C. sinensis* var. *sinensis* (Kim et al. 2013; Langmead and Salzberg 2012). Any remaining unmapped reads were still mapped to the genome, using TopHat-Fusion software (Kim and Salzberg 2011). CIR-Cexplorer2 software was used to de novo assemble the mapped reads to circRNAs (Zhang et al. 2014, 2016). Then, back-splicing reads were identified in unmapped reads, using TopHat-Fusion software and CIR-Cexplorer2 software (Kim and Salzberg 2011; Zhang et al. 2014, 2016). All samples were analyzed to generate unique circRNAs, and the empirical

analysis of digital gene expression data package in R (edgeR, version 3.13) was used to analyze the DE CircRNAs under the conditions of $\log_2$ (fold change) >1 or <-1 and with statistical significance ($P < 0.05$) (Robinson et al. 2010).

Finally, for the miRNA component, raw reads were screened using ACGT101-miR software (LC Sciences, Houston, TX, U.S.A.), and adapter dimers and low-complexity sequences were removed, with the remaining reads being filtered by aligning with the databases of mRNA, RNA Families (ribosomal RNA, transfer RNA, small nuclear RNA, and small nucleolar RNA), and Repbase. Reads with lengths of 18 to 25 nt (miRNA-like) were retained, then, unique sequences with lengths in the range of 18 to 25 nt were mapped to tea species-specific precursors in miRBase 22.0 by BLAST search. Length variation at both the 3' and 5' ends and one mismatch within the sequence were allowed in the alignment. The unique sequences mapping to specific species mature miRNAs in hairpin arms were identified as known miRNAs. The unique sequences mapping to the other arm of known specific species precursor hairpin arms, opposite to the annotated mature miRNA-containing arm, were considered to be novel 5p- or 3p-derived miRNA candidates. The remaining sequences were mapped to other selected species precursors (with the exclusion of specific species) in miRBase 22.0 by BLAST search, and the mapped precursor-miRNAs were further BLAST searched against the specific species genomes to determine their genomic locations. Based on the above two steps, known miRNAs were identified. The unmapped sequences were searched for by BLAST against the specific genomes, and the hairpin RNA structures containing the sequences aligning in specific genomes were predicted from the flanking 120-nt sequences, using RNAfold software. The DE miRNAs, based on normalized deep-sequencing counts, were analyzed for significance using Student's *t* test at the threshold of *P* value <0.05 (Li et al. 2016).

The regulation mode of miRNA-mRNA interaction pairs was predicted using Target-Finder software (Allen et al. 2005; Manke et al. 2008; Mrowka et al. 2008; Wasserman and Sandelin 2004). The alignment score, as the screening criterion, was used to predict the interaction between mRNA and miRNA. When the sequences of miRNA and mRNA were aligned, the effect of the complementary sequences was evaluated using the alignment score. A mismatch or missing base was calculated as a "1" score, and a pair of "G: U" mismatches was calculated as a "0.5" score. The alignment score would double if the mismatch of "G: U" took place in the core region of miRNA; in general, the region from the second nucleotide to the thirteenth nucleotide was known as the core region. The alignment score indicated the degree of match between the target gene and miRNA. As the alignment score decreased, the degree of matching became more complete and the results became more reliable. The default maximum penalty in our research was set to a score of "4".

The regulation mode of miRNA-lncRNA/circRNA was predicted using the ssearch36 (36.3.6) software (Farrar 2007; Pearson 2016; Smith and Waterman 1981). In this study, the specific rules are as follows: i) the bump must be on the lncRNA, and the bump is in the middle of the miRNA. The middle region of the miRNA was defined as follows: #19-nt, 9 to 11 bits; #20-nt, 10 to 11 bits; #21-nt, 10 to 12 bits; #22-nt, 11 to 12 bits; #23-nt, 11 to 13 bits; #24-nt, 12 to 13 bits; #25-nt, 12 to 14 bits; #26-nt, 13 to 14 bits; ii) in addition to the middle region of the miRNA, the maximum values of mismatches were ≤ 4 ; iii) bumps are not allowed at non-middle positions of the miRNA. lncRNAs may play a *cis* role, acting on neighboring target genes. The *cis* target genes of lncRNAs were used to predict the mRNA-lncRNA interaction. In this study, coding genes in the 100,000-bp upstream and downstream base pairs were selected using Python script. Then, we carried out functional analysis of the target genes for lncRNAs by using a BLAST search (Conesa et al. 2005). Significance was expressed as $P < 0.05$.

A total of 31, 24, six, and two of the ceRNA or miRNA interaction pairs were predicted in the miRNA-mRNA, miRNA-lncRNA, miRNA-circRNA, and mRNA-lncRNA interactions, respectively. For alignment scores of 2 and 2.5, miRNA-mRNA interactions were predicted for PC-3p-97705_24/TEA004183.1:TEA004183.1 and PC-3p-125521_18/TEA022449.1:TEA022449.1, respectively. A *probable ribose-5-phosphate isomerase 2* (TEA004183.1) was annotated with the GO term "transmembrane transport" and the KEGG pathway "biosynthesis of amino acids". The *uncharacterized gene LOC100251837* (TEA022449.1) was annotated with the GO term "trehalose-phosphatase activity" and the KEGG pathway "starch and sucrose metabolism". From the prediction of miRNA-lncRNA interactions,

Table 1. The targeted prediction analysis of CeRNA from tea [*Camellia sinensis* (L.) O. Kuntze] leaves during challenge by the fungal pathogen *Pestalotiopsis trachicarpicola*^a

Targeted prediction analysis		Number of correlation analyses
miRNA-mRNA	31 pairs	
	23 miRNAs/20 mRNAs	
miRNA-lncRNA	24 pairs	
	11 miRNAs/16 lncRNAs	
miRNA-circRNA	6 pairs	
	5 miRNAs/3 circRNAs	
mRNA-lncRNA	2 pairs	
	2 mRNA/2 lncRNA	
All lncRNA/circRNA/miRNA/mRNA	4/0/1/1 pairs	
Type 1: circRNA-miRNA-mRNA	circRNA292/PC-5p-4673_314/ TEA020085.1:TEA020085.1	
	circRNA791/PC-5p-4673_314/ TEA020085.1:TEA020085.1	
Type 2: lncRNA-miRNA-mRNA	..MSTRG.5527.1/nta-MIR172e-p3_2ss15CG19TA/ TEA015316.1:TEA015316.1	
	..MSTRG.1706.1/nta-MIR172e-p3_2ss15CG19TA/ TEA015316.1:TEA015316.1	
	TEA024896.1:MSTRG.5459.1/nta-MIR172e-p3_2ss15CG19TA/TEA015316.1:TEA015316.1	
	..MSTRG.30017.1/nta-MIR172e-p3_2ss15CG19TA/ TEA015316.1:TEA015316.1	
Type 3: circRNA-lncRNA-miRNA-mRNA	circRNA292/vvi-MIR396b-p3_2ss7AG20AG/ TEA027239.1:TEA027239.1	
	circRNA292/vvi-MIR396b-p3_2ss7AG20AG/ TEA001447.1:TEA001447.1	

^a miRNA = microRNA, mRNA = messenger RNA, lncRNA = long noncoding RNA, and circRNA = circular RNA.

nta-MIR172e-p3_2ss15CG19TA/..MSTRG.5527.1, nta-MIR172e-p3_2ss15CG19TA/..MSTRG.30017.1, nta-MIR172e-p3_2ss15CG19TA/..MSTRG.1706.1, nta-MIR172e-p3_2ss15CG19TA/TEA024896.1:MSTRG.5459.1, and PC-3p-103122_23/..MSTRG.1183.1 exhibited significant interactions, and Smith-Waterman (SW) scores of these pairs were at the top of the list. Interestingly, nta-MIR172e-p3_2ss15CG19TA may target the four lncRNAs described above. In the prediction of miRNA-circRNA interactions, vvi-MIR396b-p3_2ss7AG20AG and circRNA292 may exhibit an interaction, with a SW score of 117 and an E-value of 1.10. The interaction of PC-5p-189883_10 and circRNA776 exhibited a higher SW score of 77 but a lower E-value of 0.00. For the prediction of mRNA-lncRNA interactions, TEA033765.1/MSTRG.24999.1, and TEA030398.1/MSTRG.12855.1 may exhibit interactions, with the Pearson's correlation coefficients being 0.87 and 0.97, respectively. Integrated analysis of interactions among these four kinds of noncoding RNAs with mRNAs showed that the number of miRNA-mRNA and miRNA-lncRNA interactions was significantly higher than that of miRNA-circRNA interactions (Table 1).

For analysis of the interaction pairs in lncRNA/circRNA-miRNA-mRNA, the regulation modes were divided into three types, namely circRNA-miRNA-mRNA (type 1), lncRNA-miRNA-mRNA (type 2), and circRNA-lncRNA-miRNA-mRNA (type 3). For type 1, two ceRNA pairs were predicted in the regulation modes and are listed as circRNA292/PC-5p-4673_314/TEA020085.1:TEA020085.1 and circRNA791/PC-5p-4673_314/TEA020085.1:TEA020085.1. For type 2, four pairs of ceRNAs, composed of four lncRNAs, one miRNA, and one mRNAs, were predicted and are listed as ..MSTRG.5527.1/nta-MIR172e-p3_2ss15CG19TA/TEA015316.1:TEA015316.1, ..MSTRG.1706.1/nta-MIR172e-p3_2ss15CG19TA/TEA015316.1:TEA015316.1, TEA024896.1:MSTRG.5459.1/nta-MIR172e-p3_2ss15CG19TA/TEA015316.1:TEA015316.1, and ..MSTRG.30017.1/nta-MIR172e-p3_2ss15CG19TA/TEA015316.1:TEA015316.1. GO annotation of TEA015316.1:TEA015316.1 indicated that it was involved with GO:0006810 (transport), GO:0005524 (ATP binding), GO:0016021 (integral component of membrane), and GO:0016887 (ATPase activity). For type 3, two ceRNA pairs were predicted in the regulation modes, listed as circRNA292/vvi-MIR396b-p3_2ss7AG20AG/TEA027239.1:TEA027239.1 and circRNA292/vvi-MIR396b-p3_2ss7AG20AG/TEA001447.1:TEA001447.1 (Table 1).

Based on the analysis of interactions between ceRNAs and miRNAs, the abundance of mRNAs originating from tea leaves was predicted. For instance, lncRNAs ..MSTRG.5527.1, ..MSTRG.1706.1, TEA024896.1:MSTRG.5459.1, and ..MSTRG.30017.1 competitively bound miRNA nta-MIR172e-p3_2ss15CG19TA, and the abundance of mRNA TEA015316.1:TEA015316.1 was significantly upregulated through the comparison of the mRNA abundance of infected and uninfected samples. CircRNA292 and circRNA791 competitively bound miRNA PC-5p-4673_314, so that the abundance of mRNA TEA020085.1:TEA020085.1 was significantly upregulated. CircRNA292 competitively bound the miRNA vvi-MIR396b-p3_2ss7AG20AG,

resulting in the abundance of mRNAs TEA027239.1:TEA027239.1 and TEA001447.1:TEA001447.1 being significantly down- and upregulated, respectively. These three mRNAs were not annotated in the GO database (Table 1).

Data Availability

In this study, the sequences and the interaction analysis of ceRNAs and miRNAs originating from infected tea leaves have provided a significant data resource for use in understanding the pathogenicity of *P. trachicarpicola* and the response mechanism of tea. The sequences of ceRNAs and miRNA of tea during infection by *P. trachicarpicola* have been deposited in GenBank under SRA accession number PRJNA756805 for lncRNAs and circRNAs, PRJNA756832 for mRNAs, and PRJNA742222 for small RNAs. In addition, the bioinformatics analysis of ceRNAs is presented in the Supplementary Tables.

Author-Recommended Internet Resources

CIRCexplorer2 software: <https://github.com/YangLab/CIRCexplorer2>
 Cutadapt: <https://github.com/marcelm/cutadapt>
 edgeR: <http://bioconductor.org/packages/release/bioc/html/edgeR.html>
 FastQC: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>
 HISAT software (version 2.0.4): <http://www.ccb.jhu.edu/software/hisat/index.shtml>
 Perl program: <https://www.perl.org>
 Repbase: <https://www.girinst.org/repbase>
 RNA Families database: <http://rfam.janelia.org>
 StringTie software: <https://ccb.jhu.edu/software/stringtie>
 TargetFinder software: <https://github.com/carringtonlab/TargetFinder>
 TopHat-Fusion software: <http://tophat-fusion.sourceforge.net>

Literature Cited

- Allen, E., Xie, Z., Gustafson, A. M., and Carrington, J. C. 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121: 207-221.
- Bao, X.-T., Dharmasena, D. S. P., Li, D.-X., Wang, X., Jiang, S.-L., Ren, Y.-F., Wang, D.-L., Song, B.-A., and Chen, Z. 2019. First report of *Epicoccum sorghinum* causing leaf spot on tea in China. *Plant Dis.* 103:3282.
- Buchfink, B., Xie, C., and Huson, D. H. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12:59-60.
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676.
- Cui, J., Jiang, N., Hou, X., Wu, S., Zhang, Q., Meng, J., and Luan, Y. 2020. Genome-wide identification of lncRNAs and analysis of ceRNA networks during tomato resistance to *Phytophthora infestans*. *Phytopathology* 110:456-464.
- Fan, J., Quan, W., Li, G.-B., Hu, X.-H., Wang, Q., Wang, H., Li, X.-P., Luo, X., Feng, Q., Hu, Z.-J., Feng, H., Pu, M., Zhao, J.-Q., Huang, Y.-Y., Li, Y., Zhang, Y., and Wang, W.-M. 2020. circRNAs are involved in the rice-*Magnaporthe oryzae* interaction. *Plant Physiol.* 182:272-286.
- Farrar, M. 2007. Striped Smith-Waterman speeds database searches six times over other SIMD implementations. *Bioinformatics* 23:156-161.
- Frazee, A. C., Perte, G., Jaffe, A. E., Langmead, B., Salzberg, S. L., and Leek, J. T. 2015. Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat. Biotechnol.* 33:243-246.
- Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., and Kjems, J. 2013. Natural RNA circles function as efficient microRNA sponges. *Nature* 495:384-388.
- Jayawardena, R. S., Zhang, W., Liu, M., Maharachchikumbura, S. S., Zhou, Y., Huang, J., Nilthong, S., Wang, Z., Li, X., Yan, J., and Hyde, K. D. 2015. Identification and characterization of *Pestalotiopsis*-like fungi related to grapevine diseases in China. *Fungal Biol.* 119:348-361.
- Jiang, N., Cui, J., Hou, X., Yang, G., Xiao, Y., Han, L., Meng, J., and Luan, Y. 2020. Sl-lncRNA15492 interacts with Sl-miR482a and affects *Solanum lycopersicum* immunity against *Phytophthora infestans*. *Plant J.* 103:1561-1574.
- Kim, D., Langmead, B., and Salzberg, S. L. 2015. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* 12:357-360.
- Kim, D., Perte, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. 2013. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14:R36.
- Kim, D., and Salzberg, S. L. 2011. TopHat-Fusion: An algorithm for discovery of novel fusion transcripts. *Genome Biol.* 12:R72.
- Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., and Gao, G. 2007. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 35: W349-W349.
- Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9:357-359.
- Li, D.-X., Bao, X.-T., Ren, Y.-F., Wang, Y., Song, B.-A., and Chen, Z. 2019a. First report of *Lasiodiplodia theobromae* causing leaf spot on tea plant in Guizhou Province of China. *Plant Dis.* 103:374.
- Li, H., Teng, R.-M., Liu, J.-X., Yang, R.-Y., Yang, Y.-Z., Lin, S.-J., Han, M.-H., Liu, J.-Y., and Zhuang, J. 2019b. Identification and analysis of genes involved in auxin, abscisic acid, gibberellin, and brassinosteroid metabolisms under drought stress in tender shoots of tea plants. *DNA Cell Biol.* 38:1292-1302.
- Li, S.-J., Li, S.-Y., Zhang, Z.-Y., Zhu, H.-M.-Y., Zhu, T.-H., Qiao, T.-M., and Han, S. 2018. First report of *Pestalotiopsis trachicarpicola*, a causal agent of leaf spot of *Eucommia ulmoides* in Southwest China. *Plant Dis.* 102: 2656.
- Li, T., Yin, Q.-X., Wu, X., Jiang, S.-L., Wang, Y., Dharmasena, D. S. P., Wei, X.-L., and Chen, Z. 2021. Studies on identification and biology of the pathogenic fungi *Pestalotiopsis trachicarpicola* isolated from tea leaf spot in Yuqing County. *China Plant Prot.* 41:24-30.
- Li, X., Shahid, M. Q., Wu, J., Wang, L., Liu, X., and Lu, Y. 2016. Comparative small RNA analysis of pollen development in autotetraploid and diploid rice. *Int. J. Mol. Sci.* 17:499.
- Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15: 550.
- Maharachchikumbura, S. S. N., Guo, L. D., Cai, L., Chuksatirote, E., Wu, W. P., Sun, X., Crous, P. W., Bhat, D. J., McKenzie, E. H. C., Bahkali, A. H., and Hyde, K. D. 2012. A multi-locus backbone tree for *Pestalotiopsis*, with a polyphasic characterization of 14 new species. *Fungal Divers.* 56:95-129.
- Manke, T., Roider, H. G., and Vingron, M. 2008. Statistical modeling of transcription factor binding affinities predicts regulatory interactions. *PLOS Comput. Biol.* 4:e1000039.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17:10-12.

- Mrowka, R., Blüthgen, N., and Föhling, M. 2008. Seed-based systematic discovery of specific transcription factor target genes. *FEBS J.* 275:3178-3192.
- Pandey, A. K., Sinniah, G. D., Babu, A., and Tanti, A. 2021. How the global tea industry copes with fungal diseases - challenges and opportunities. *Plant Dis.* 105:1868-1879.
- Pearson, W. R. 2016. Finding protein and nucleotide similarities with FASTA. *Curr. Protoc. Bioinform.* 53:3.9.1-3.9.25.
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33:290-295.
- Qi, M., Xie, C. X., Chen, Q. W., and Yu, Z. D. 2021. *Pestalotiopsis trachicarpicola*, a novel pathogen causes twig blight of *Pinus bungeana* (Pinaceae: Pinoideae) in China. *Antonie van Leeuwenhoek* 114:1-9.
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. 2010. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139-140.
- Shu, J., Yu, Z., Sun, W., Zhao, J., Li, Q., Tang, L., Guo, T., Huang, S., Mo, J., Hsiang, T., and Luo, S. 2020. Identification and characterization of *Pestalotioid* fungi causing leaf spots on mango in southern China. *Plant Dis.* 104:1207-1213.
- Smith, T. F., and Waterman, M. S. 1981. Identification of common molecular subsequences. *J. Mol. Biol.* 147:195-197.
- Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., Liu, Y., Chen, R., and Zhao, Y. 2013. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 41:e166.
- Sun, Y., Zhang, H., Fan, M., He, Y., and Guo, P. 2020. Genome-wide identification of long non-coding RNAs and circular RNAs reveal their ceRNA networks in response to cucumber green mottle mosaic virus infection in watermelon. *Arch. Virol.* 165:1177-1190.
- Tay, Y., Rinn, J., and Pandolfi, P. P. 2014. The multilayered complexity of ceRNA crosstalk and competition. *Nature* 505:344-352.
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., and Pachter, L. 2010. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28:511-515.
- Tsai, I., Chung, C.-L., Lin, S.-R., Hung, T.-H., Shen, T.-L., Hu, C.-Y., Hozzein, W. N., and Ariyawansa, H. A. 2021. Cryptic diversity, molecular systematics, and pathogenicity of genus *Pestalotiopsis* and allied genera causing gray blight disease of tea in Taiwan, with a description of a new *Pseudopestalotiopsis* species. *Plant Dis.* 105:425-443.
- Wang, X., Yin, Q.-X., Jiang, S.-L., Wu, X., Wang, D.-L., Song, B.-A., and Chen, Z. 2020. First report of *Didymella bellidis* causing tea leaf spot in China. *Plant Dis.* 104:1254.
- Wasserman, W. W., and Sandelin, A. 2004. Applied bioinformatics for the identification of regulatory elements. *Nat. Rev. Genet.* 5:276-287.
- Yin, Q., An, X., Wu, X., Dharmasena, D. S. P., Li, D., Jiang, S., Wang, Y., Wang, D., and Chen, Z. 2021. First report of *Alternaria longipes* causing leaf spot on tea in China. *Plant Dis.* 105:4167.
- Yin, Q. X., Jiang, S. L., Li, D. X., Huang, H. L., Wang, Y., Wang, D. L., and Chen, Z. 2022. First report of *Epicoccum nigrum* causing brown leaf spot in tea in Guizhou Province, China. *Plant Dis.* 106:321.
- Zhang, X.-O., Dong, R., Zhang, Y., Zhang, J.-L., Luo, Z., Zhang, J., Chen, L.-L., and Yang, L. 2016. Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* 26:1277-1287.
- Zhang, X.-O., Wang, H.-B., Zhang, Y., Lu, X., Chen, L.-L., and Yang, L. 2014. Complementary sequence-mediated exon circularization. *Cell* 159:134-147.
- Zhang, Y.-M., Maharachchikumbura, S. S. N., McKenzie, E. H. C., and Hyde, K. D. 2012. A novel species of *Pestalotiopsis* causing leaf spots of *Trachycarpus fortunei*. *Cryptogamie Mycol.* 33:311-318.