ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/282334242

Characterization of Small RNAs and Their Targets from Fusarium oxysporum Infected and Noninfected Cotton Root Tissues

Article *in* Plant Molecular Biology Reporter · September 2015

DOI: 10.1007/s11105-015-0945-z

CITATIONS
7

READS
339

11 authors, including:

Umidjon Shapulatov
National University of Uzbekistan
15 PUBLICATIONS **123** CITATIONS

SEE PROFILE

Zabardast T Buriev
Uzbekistan Academy of Sciences
66 PUBLICATIONS **774** CITATIONS

SEE PROFILE

Mauricio Ulloa
University of Concepción
119 PUBLICATIONS **2,908** CITATIONS

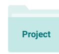
SEE PROFILE

Sukumar Saha
Agricultural Research Service
154 PUBLICATIONS **3,904** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Cotton Ovule small RNAs View project


Cotton Fair View project

All content following this page was uploaded by Umidjon Shapulatov on 19 October 2019.

The user has requested enhancement of the downloaded file.

Plant Mol Biol Rep
DOI 10.1007/s11105-015-0945-z

BRIEF COMMUNICATION

Characterization of Small RNAs and Their Targets from *Fusarium oxysporum* Infected and Noninfected Cotton Root Tissues

U. M. Shapulatov[1] · **Z. T. Buriev**[1] · **M. Ulloa**[2] · **S. Saha**[3] · **E. J. Devor**[4] · **M. S. Ayubov**[1] · **T. M. Norov**[1] · **S. E. Shermatov**[1] · **A. Abdukarimov**[1] · **J. N. Jenkins**[3] · **I. Y. Abdurakhmonov**[1]

© Springer Science+Business Media New York 2015

Abstract Genes for host-plant resistant do exist in cotton (*Gossypium* spp.) but improvement of cotton cultivars with resistance is difficult due to intensive breeding. Identifying molecular-genetic mechanisms associated with disease resistance can offer a new way to combat a serious threat such as *Fusarium oxysporum* f. sp. *vasinfectum* (FOV). Here, we captured and annotated "top-layer" of abundantly and specifically expressed cotton root small RNA (sRNA) including microRNA (miR) sequences during FOV pathogenesis using size-directed and adenylated linker-based sRNA cloning strategy. A total of 4116 candidate sRNA sequences with 16 to 30 nucleotide (nt) length were identified from four complementary DNA (cDNA) libraries of noninfected and FOV race 3-infected roots of susceptible ("11970") versus resistant ("Mebane B-1") cotton genotypes (*G. hirsutum* L.). The highest numbers of sRNA signatures were those with 19–24 nt long in all libraries, and interestingly, the number of sRNAs substantially increased during FOV infection in a resistant genotype, while it sharply decreased in a susceptible genotype. In BLAST analysis, more than 73 % of sRNAs matched *Gossypium* (G. *arboretum* L., *G. hirsutum*, and *G. barbadense* L.) ESTs. A small percentage of sRNAs matched *A. thaliana* (1.68 %), *T. cacao* (1.26 %), fungal (2 %), and other organism (21.33 %) ESTs. mirBase comparisons showed that 4 % of sRNAs were homologous to previously reported plant miRs, among which we predicted novel cotton Ghr-miR-160 that was not registered in the cotton miR database. These major representative sRNA signatures targeted proteins associated with the key biological processes and molecular functions, explaining the molecular mechanisms of the host defense response during the FOV pathogenesis in cotton.

Keywords Cotton · Small RNAs · microRNAs · Fusarium oxysporum (FOV) wilt resistance · Size-directed sRNA cloning

Introduction

Fusarium wilt [*Fusarium oxysporum* f.sp. *vasinfectum* (FOV) Atk. Sny & Hans] represents a serious threat to cotton (*Gossypium* spp.) and many other economically important crops worldwide (Kim et al. 2005; Ulloa et al. 2006, 2011). This fungus is a soil-borne organism that can survive for long periods in the absence of a host, making it impractical to eradicate from infested fields (Kim et al. 2005, Ulloa et al. 2013). Genes for Fusarium wilt resistance do exist in cotton, but cultivars with improved FOV resistance are difficult to develop due to intensive breeding and time-consuming

Electronic supplementary material The online version of this article (doi:10.1007/s11105-015-0945-z) contains supplementary material, which is available to authorized users.

✉ I. Y. Abdurakhmonov
ibrokhimabdurakhmonov@hotmail.com

[1] Center of Genomics and Bioinformatics under Academy of Sciences the Republic of Uzbekistan, Ministry of Agriculture and Water resources the Republic of Uzbekistan and "Uzpakhtasanoat" Association, University Street-2, Qibray Region, Tashkent District 111215, Uzbekistan

[2] USDA-ARS, Plant Stress and Germplasm Development Research, 3810 4th Street, Lubbock, TX 79415, USA

[3] USDA-ARS, Crop Science Research Laboratory, Mississippi State, Starkville, MS 39762, USA

[4] Department of Obstetrics and Gynecology, University of Iowa Carver College of Medicine, 3234 MERF, Iowa City, IA 52242, USA

Published online: 28 September 2015

bioassays required to identify resistant progeny. Identifying molecular mechanisms associated with Fusarium wilt resistance in cotton offers a new way to combat against the serious threat that this fungal disease poses.

First reported in 1993 (Lee et al. 1993), small RNA segments averaging 22 nucleotides (nt), now called small RNA (sRNAs) including micro RNAs (miRs), play an important role in negatively regulating gene expression of many protein-coding genes (Bartel 2004). sRNAs and miRs have been shown to be important in several biological processes, including plant resistance to biotic and abiotic stresses (Jagadeeswaran et al. 2009). The importance of sRNAs continues to expand as more of their regulatory targets and functions are discovered. For instance, double-stranded ribonucleic acid (RNA) silencing plays an important role in defense against bacterial plant pathogens in Arabidopsis (*Arabidopsis thaliana*) through modulating host defense (Ellendorff et al. 2009; Navarro et al. 2006). In these studies, gene silencing was achieved through inhibition of gene transcription (transcriptional gene silencing) or posttranscriptional degradation of RNA (post- transcriptional gene silencing). These findings correlated with the accumulation of small interfering RNAs (siRNA) or miR of 20–27-nt length.

Numerous miRs and siRNAs have been reported to be involved in biotic stress tolerance in plants (Khraiwesh et al. 2011). miR-deficient *A. thaliana* mutants revealed that miRs are key components of plant basal defense (Navarro et al. 2006; Ellendorff et al. 2009; Navarro et al. 2008). For example, miR-393 was isolated from Arabidopsis following bacterial infection (Navarro et al. 2006). In Arabidopsis, two miRs, miR-160 and miR-167, were highly induced, and another, miR-825, was highly downregulated during gram-negative bacterial pathogen Pst DC3000 infection (Fahlgren et al. 2007). Lu et al. (2010) cloned 26 miRs from loblolly pine under fungal infection. Two new miRs (miR-158 and miR-1885) were induced by Turnip mosaic virus (TuMV) in *Brassica napus* (He et al. 2008). Interestingly, miRs were not induced by either Cucumber mosaic virus, Tobacco mosaic virus, or the fungal pathogen *Sclerotinia sclerotiorum* infection in *B. napus* or *B. rapa*. The target of bra-miR-1885 is a member of the TIR-nucleotide-binding site domains (NBS)-C-terminal leucine-rich repeats (LRR) class that regulates disease resistance. Other data indicate that some conserved miRs are differentially regulated in response to powdery mildew in wheat. miR-156, miR-159, miR-164, miR-171, and miR-396 were downregulated, and miR-393, miR-444, and miR-827 were upregulated in response to powdery mildew (Xin et al. 2010). Both miR-398b and miR-773 negatively regulate bacterial resistance in Arabidopsis (Li et al. 2010).

A large number of plant miRs were predicted by computational methods; however, a few of these have been identified from direct isolation and cloning (Zhang et al. 2006). Although generally understood and required that small RNA profiling studies presently involve a "large-scale" deep sequencing effort, utilizing new generation DNA/RNA sequencing platforms, these technologies are still not readily available and highly resourceful for many plant genomics laboratories under funding limitations. In our previous studies, we developed and successfully utilized (Devor et al. 2009; Abdurakhmonov et al. 2008) the size-directed sRNA cloning strategy using adenylated linker tools that is "especially suitable for the 'small-scale' plant genome laboratories worldwide" (Devor et al. 2009) and providing an opportunity to profile and capture major signatures from sRNA pool of chosen plant tissues. In this short report, using this size-directed sRNA cloning strategy, we captured and characterized major signatures of sRNAs that are abundantly and specifically expressed between FOV race 3-infected and noninfected seedling root tissues of resistant versus susceptible cotton genotypes by isolating, cloning, and sequencing. Results of analysis of target proteins of sRNA signatures should be helpful to understand sRNA-dependent FOV pathogenesis regulation in cotton and its molecular mechanisms during the host defense response. Further, being major signatures captured, these sRNA sequence signatures and their putative target proteins, characterized herein, should be useful for breeding of wilt-resistant cotton varieties to address the FOV threats in cotton farming.

Materials and Methods

Plant Materials and Growth Condition

Acid delinted cotton (*G. hirsutum* L.) seeds of accession "Mebane B-1" (FOV race 3-resistant genotype) and accession "11970" (FOV race 3 susceptible genotype) were surface sterilized in 70 % ethanol for 15 min and then germinated on MS (Sigma-Aldrich, USA) medium agar plate at 28/22 °C day/night temperature. After 7 days, young cotton seedling roots were infected with FOV race 3 (isolate no. 316; Egamberdiev et al. 2013) conidia suspension following the overall inoculation methods of Yin et al. (2012). An extremely aggressive FOV race 3 strains was provided by the Institute of Plant Genetics and Experimental Biology Germplasm Collection, the Academy of Science of Uzbekistan.

We prepared a conidial suspension solution of 10^7 spores per milliliter with sterile distilled water from FOV race 3 before inoculation. Individual cotton seedlings were infected with FOV by root-dip inoculation into a suspension of fungal conidia for 5 min and returned to the MS medium agar-plate for discrete postinoculation time intervals. Control seedlings of each line were inoculated with water following the same procedure instead of conidial suspension. The roots of both pathogen-infected and noninfected seedlings were harvested

after 3 days postinfection (dpi) and immediately frozen in liquid nitrogen. Root tissues were temporarily stored at −80 °C for isolation of total RNAs.

Total RNA Purification and Small RNA Cloning

Total RNA was extracted from 3-dpi bulked seedling root tissues of individual infected and noninfected plants following the method of Wu et al. (2002). Briefly, 0.5 M NaCl and absolute ethanol were used to precipitate and enrich RNAs with low molecular weight. Then, 100 mg of low molecular weight RNA was used for electrophoresis in 15 % denaturing polyacrylamide gel (PAGE). After gel electrophoresis, the small RNA fragments in the range of 16–30 nt were excised and eluted from the gel. The small RNAs were further purified using nuclease-free Nuc Away Spin columns (Ambion Inc., Austin, TX, USA) as per the manufacturer's protocol. The small RNAs were then ligated to 3 and 5′ linkers in two separate reactions using the miRCat-33™ microRNA cloning kit (IDT DNA Technologies, Coralville IA, USA) as per the Integrated DNA Technologies website (http://www.idtdna.com; Abdurakhmonov et al. 2008; Devor et al. 2009).

First, purified small RNAs were 3′ ligated to an adenylated cloning linker containing a 3′ block (5′-rAppCTG TAGGCACCATCAATddC-3′) using T4 RNA Ligase (1 U/μL) in the absence of ATP. Ligations were performed at 22 °C for 2 h. An internal control-miSPIKE (10 pmol) was added into the enriched small RNA extracts before 3′ linkering reaction. The 40 nt 3′ ligated RNAs were then purified from the dPAGE gel using Performa DTR gel Filtration columns (Edge BioSystems, Gaithersburg, MD, USA). Ligated RNAs were then 5′ ligated with a DNA/RNA chimeric linker (5′-TGGAATucucgggcaccaaggu-3′) using T4 RNA Ligase in the presence of 10 mg/ml ATP. The double linkered RNAs were reverse transcribed with a 3′ linker-specific reverse primer (5′-GATTGATGGTGCCTACAG-3′, T_m=50.2 °C). The 5 and 3′ ligated small RNAs were then converted to complementary DNA (cDNA) by RT-PCR using Superscript III reverse transcriptase (Invitrogen, Carlsbad, CA, USA).

PCR amplification of the reverse transcripts was carried out using the RT primer as the reverse primer and a linker-specific forward primer (5′-TGGAATTCTCGGGCACC-3′, T_m= 55.0 °C). PCR conditions were 95.0 °C for 5 min, followed by 25 cycles of 95.0 °C for 30 s, 52.0 °C for 30 s, and 72 °C for 30 s, and a final extension step of 72.0 °C for 7 min. The expected 60–65-bp PCR amplicons were obtained and gel purified using QIAQuick® PCR clean-up columns (QIAGEN, Valencia, CA, and USA). The amplicons were restriction digested with Ban I restriction endonuclease (New England Biolabs, Ipswich, MA, USA) for 1 h at 37 °C and concatenated using T4 DNA ligase (30 U/μL). Concatemers were cloned into the pCR2.1 TOPO vector and transformed into TOP10 one-shot *Escherichia coli* cells according to the manufacturer's protocol (Invitrogen, Carlsbad, CA, USA). *E. coli* transformants were spread and grown on LB plates containing 50 mg/ml kanamycin overnight. Colonies were randomly selected for colony PCR using M-13 primer sets, and positive amplicons were then purified using 26 % PEG-8000 and sequenced using the ABI auto sequencing kit (see Supplementary Fig. S1).

Sequence Analysis

We analyzed small RNA sequences using Sequencher 4.5 (Gene Codes, USA), where vector and linker sequences were trimmed and appropriate 3 and 5′ ends of inserted sRNAs were defined based on linkers. The connector sequence was either (CTGTAGGCACCAAGGT) or (ACCTTGGTGCCTACAG). Only unique sequences in the size range of 16–30 nt were searched in the Central Registry Database, miRBase (release 21, ttp://miR.sanger.ac.uk). The putative targets of small RNAs were detected using Target Finder (http://plantgrn.noble.org/psRNATarget/) using *Gossypium* DFGI gene index release 11 unigene databases (dated on November 15, 2011) with default parameters. GO molecular function and biological process of these targets were studied using Universal Protein Resources (UniProt) database (http://www.uniprot.org/, UniProt release 2011_11, November 16, 2011). Secondary structures of RNA precursors were predicted from cloned RNAs using Mfold (http://mfold.bioinfo.rpi.edu).

Prediction of Novel microRNAs

To predict novel miRs, first, we identified sRNA signatures that are highly homologous to plant miR sequences and then selected those candidate miRs that is not present in cotton miR database. Further, candidate miRs were blasted against *Gossypium* EST/GSS database (http://blast.ncbi.nlm.nih.gov/blast/), and perfectly matched ESTs were subjected for RNA folding analysis following criteria: (1) miRNAs ought to have less than six mismatches with the opposite miR sequence in the other arm, (2) the minimal folding free energy (MFE) of predicted pre-miR secondary structures has to be lower than −20 kcal/mol, and (3) the minimal folding free energy index (MEFI) of it usually must be over 0.85 (Zhang et al. 2006).

Results and Discussions

We identified 4116 candidate small RNA (sRNA) sequences in the appropriate size range (16–30 nt) from four separate cDNA libraries of cotton (*G. hirsutum*) roots: infected vs. noninfected roots of Mebane B-1 and 11970 cotton genotypes. To simplify the sequence results, all similar or duplicate sequences in each specific sRNA library were grouped, and a

consensus sequence was derived based on their sequence alignment. In all, there were 1114 sequences of Mebane B-1 genotype from noninfected and 1233 sequences from FOV race 3-infected tissues. There were 1107 sequences of 11970 from noninfected and 662 sequences from FOV race 3 infected sRNA libraries. Further, all unique RNA sequences of each library were aligned that resulted in 947 unique sRNA sequence signatures in all four libraries sequence (Fig. 1).

Although the total number of sequence signatures was approximately the same for each specific library, length distributions of these signatures were substantially different among libraries (Fig. 1). Our results revealed that the frequency of candidate signatures (16 to 30 nt) tended to vary in FOV race 3-infected vs. noninfected tissues, and in the FOV resistant vs. susceptible genotypes where 19–24-nt long sequence signatures were most abundantly represented in four libraries. In particular, in noninfected library of resistant genotype Mebane B-1, the major pick represented by 21-nt (followed by 19 and 22 nt) signatures, while in susceptible genotype 11970 major

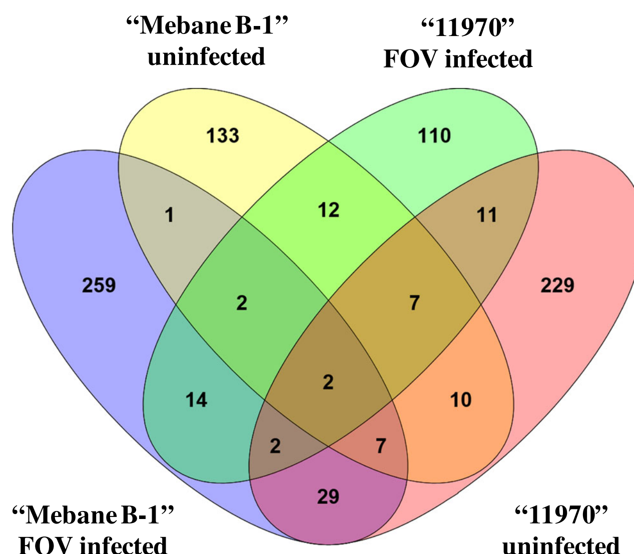


Fig. 2 *Venn diagram* showing the shared distribution of unique sRNAs from the four libraries

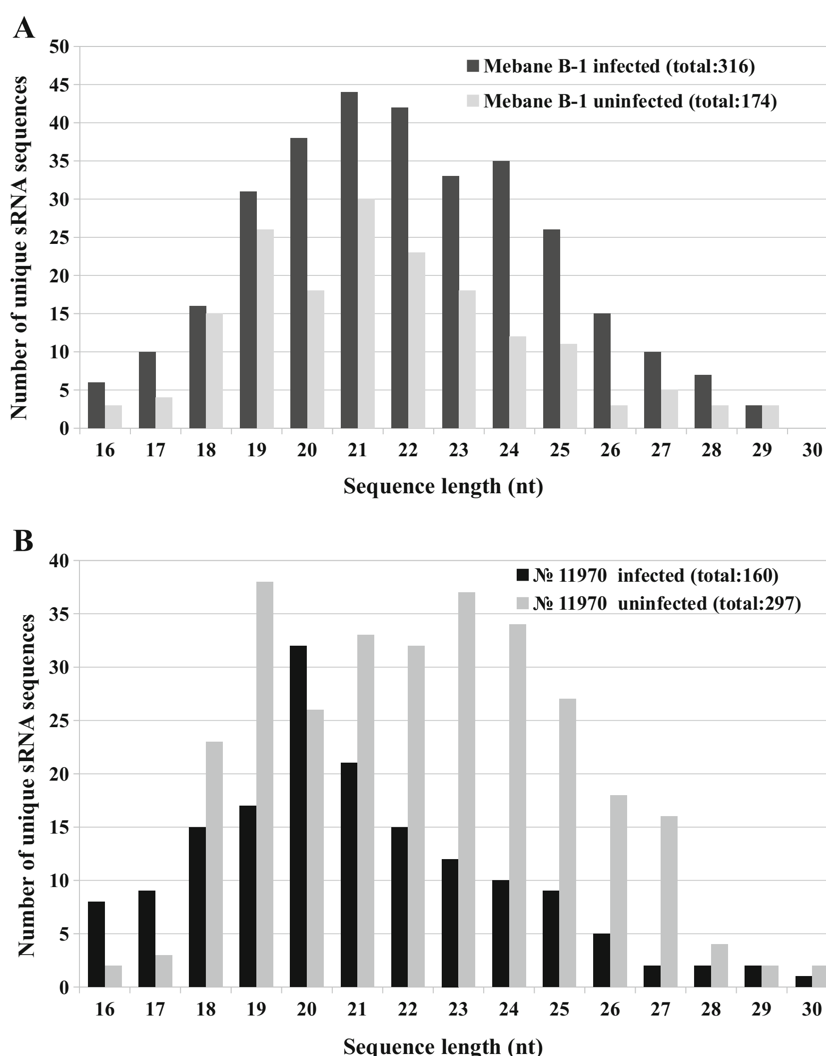


Fig. 1 Length distribution of unique small RNA (*sRNA*) sequences. **a** FOV-resistant "Mebane B-1" and **b** susceptible "11970" genotypes

Table 1 Expressed sequence tags (ESTs) matches cotton sRNA sequences from FOV infected and noninfected libraries of resistant vs. susceptible genotypes

Species	"Mebane B-1" infected	"Mebane B-1" noninfected	"11970" infected	"11970" noninfected	Total
Gossypium L.	237	121	110	230	698
A .thaliana	5	0	1	10	16
T. cacao	9	0	2	1	12
Fungal	8	0	11	0	19
Other organisms	57	53	36	56	202

sRNA class, was 19-nt sequences with secondary classes of 23-, 24-, and 21-nt sequences. After infection with FOV, resistant genotype has an increase of 21 nt sRNAs followed by a pick of 20-, 23-, and 24-nt sequence signatures, while susceptible genotype had decreased major picks of 20–21-nt signatures (Fig. 1). These results clearly demonstrate that in FOV pathogenesis and FOV defense response of cotton, there is possibly "*cross-interaction of multiple RNA silencing pathways*" functioning as observed with *Verticillium*-defense response of cotton (Yin et al. 2012).

A difference in the sequence length distribution and in the number of unique sequences was observed in our four libraries, in particular in comparisons of infected vs. noninfected tissues and/or susceptible vs. resistant genotypes that should be related to disease incidence. In a genome-wide profiling study of miRs and other small noncoding RNAs in *Verticillium dahlia*-inoculated cotton roots, Yin et al. (2012) reported distinctive distribution pattern when comparing the frequency of sequence signatures between noninfected and infected libraries of *Verticillium* resistant vs. susceptible genotypes. Yin et al. (2012) observed that the highest number of sRNAs was 21 to 24 nt in size that were significantly reduced after inoculation of both *G. hirsutum* (wilt susceptible) and *G. barbadense* (wilt resistant) genotypes with *Verticillium* except that they observed the increase of 24 nt sRNAs in infected *G. hirsutum* library. Interestingly, we observed sharply contrasting 'after-inoculation' distribution of sRNAs between resistant and susceptible libraries where the number of sRNAs of the resistant genotype (Mebane B-1) in range of 16–28 nt significantly increased after FOV infection, while sRNAs of susceptible genotype (11970) greatly decreased (except those 20 nt sRNAs) compared to noninfected library (Fig. 1). Results from capturing the major portion of sRNA signatures in these four libraries and differential after-inoculation distribution of resistant vs. susceptible genotypes collectively suggest vital role of sRNA regulation of FOV pathogenesis in cotton.

Further, when we compared sequence homology from noninfected and FOV race 3-infected libraries within genotypes and libraries, we observed that only one sequence aligned within Mebane B-1 noninfected vs. infected libraries, and only 11 sequences within 11970 FOV noninfected vs. infected libraries (Fig. 2). Thus, many sequences indicate specific expression among the four libraries, possibly contributing and explaining the response of cotton genome to the FOV infection. Therefore, results from our current small-scale study justify and prompt cotton research community to extend efforts toward deep sequencing that should help to capture a full profile of sRNA signatures as well as functional verification of abundantly and specifically represented sequence signatures during FOV pathogenesis.

BLAST Analysis

To examine the sRNA sequences from our four libraries and the possibility of discovering new sRNAs, a total 947 sRNA sequences were screened against Expressed Sequence Tags (ESTs) using BLAST. EST analysis composition is an economical alternative for sRNA discovery in species such as cotton, which do not yet have a whole genome sequence

Table 2 List of candidate mature miRs of cotton from FOV infected and noninfected small RNA libraries

Clone ID	Sequence 5–3′	Length (nt)	miRBase hit	*e*-value
"Mebane B-1" noninfected-P-10_48	ugaacugguuuguuggcugc	21	ghr-miR3476-5p	0.001
"Mebane B-1" noninfected-P-11_39	ugccuggcucccuguaugcca	21	ath-miR160a	0.001
"Mebane B-1" noninfected-P-20_79	uggcauccagggagccaggca	21	ptc-miR160g	0.001
"Mebane B-1" noninfected-P-13_35	agaaucuugaugaugcugcau	21	ath-miR172a	0.001
"Mebane B-1" noninfected-P-10_81	ggcauacagggagccaggca	20	ath-miR160c	0.003
"Mebane B-1" noninfected-P-13_43	uggcauacagggagccaggc	20	ath-miR160a	0.004
"11970" noninfected- P-1_48	ggaauguuggcuggcucgaug	21	zma-miR166m-5p	0.007

Table 3 New predicted Gh-miR-160 from EST and GSS of *G. hirsutum* L.

sRNA	Query miR	Mature sequence	Strand	Precursor length	G+C%	MFE	MFEI	EST/GSS
"Mebane B-1" noninfected-P-11_39	ath-miR160a osa-miR160a zma-miR160	ugccuggcucccuguaugcca	+	87	0.517	−45.8	1.01	GR924310

MFE minimal folding free energies, *MFEI* minimal folding free energy indexes

(Yin et al. 2012). Results indicated that there were several kinds of RNA fragments. More than 90 % of the cloned RNAs represented breakdown products of abundant coding and noncoding RNAs such as messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), and other unknown short fragments. About 3.8 % of the cloned sRNAs to be homologous to previously reported plant miRNAs. More than 73 % of sRNAs matched to *Gossypium* (*G. arboretum* L., *G. hirsutum*, and *G. barbadense* L.) ESTs. A small percentage of sRNAs matched *A. thaliana* (1.68 %), *T. cacao* (1.26 %), fungal (2 %), and other organism (21.33 %) ESTs. The two noninfected sequence libraries from the two genotypes (Mebane B-1 and 11970) did not match any fungal ESTs (see Table 1 and Supplementary Table S1) suggesting that FOV infection introduces its own sequence signatures into cotton tissues.

Identification of Cotton microRNAs

Small RNA sequence signatures were compared/screened with Expressed Sequence Tags (EST), Genomic Survey Sequence (GSS), and the mRNA database in NCBI. We chose all candidate sequences with <four mismatches. We then used mFold to determine RNA secondary structure. Homology analysis in mirBase (www.mirbase.org) showed that 3.8 % of the cloned sequences matched known miRNAs (*e*-value 0.0–0.007) of plant species (Table 2). For instance, Ghr-miR-3476 was cloned from the noninfected library that was initially identified from cotton ovules and fiber tissues (Kwak et al. 2009).

Results also showed only one sequence signature (Mebane B-1 noninfected-P-11_39, named as Ghr-miR-160; Table 3; Fig. 4), aligning with known and highly conserved plant miR-160, which was not registered in the cotton database (www.leonxie.com, EST Database of Cotton). The candidate miR-160 was sequenced from FOV noninfected Mebane B-1 young root tissues. Previous studies indicate that more than 90 % of miR precursors have a minimal folding free energy index (MFEI) greater than 0.85 (MFEI=[(MFE/length of the RNA sequence)×100]/(G+C)%) (Zhang et al. 2006). In our study, minimal folding free energy indexes of cotton Ghr-miR-160 have showed 1.01 (Table 3). This was similar to MFE index of ath-miR-160a precursor (1.01) and higher than MFEI of tRNAs (0.64), rRNAs (0.59), and mRNAs (0.62–0.66). We also evaluated G+C% and minimal folding free energy (MFE) for prediction of miR precursors confirming particular sequence signature structure as a cotton miR-160 (Table 3; Fig. 3).

Small RNA Target Analysis

Target analysis of the sRNA sequences from the four libraries was performed using the Plant Target Finder database (http://

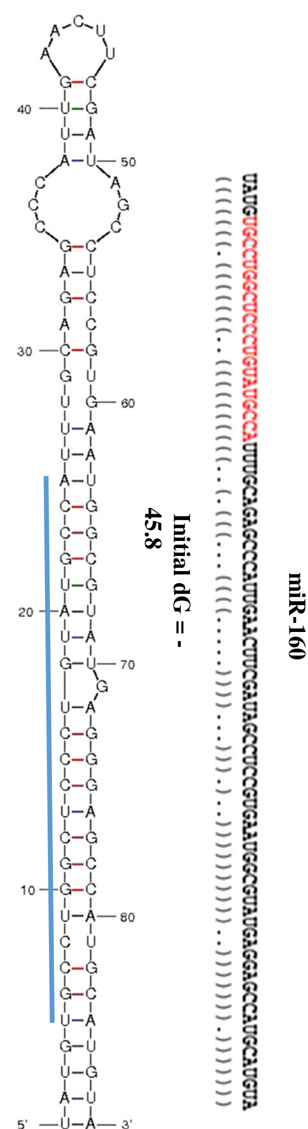


Fig. 3 Predicted stem-loop structures of the newly cloned Ghr-miR-160 from young root of *G. hirsutum*. The mature miR sequence has been *underlined*

Table 4 List of cotton sRNA and miR-targeted genes from FOV-infected and noninfected root tissues

Library	small RNA sequence	siRNA targets	Protein annotation
Mebane-1_I	CCAGGGAUCGGCGGAUGUUG	ES836880	Superoxide dismutase [Cu-Zn]
Mebane-1_I	AAGGGGCGCCGGAGAGGGUG	TC235482	Probable fructokinase-3
Mebane-1_I	CAGACUACAAUUCGGACGCCG	TC273598	Fb27
Mebane-1_I	CGGCGAUGCGCCCCGGUCGGAUG	TC248496	Xyloglucan endotransglycosylase
Mebane-1_I	UGAACCGGGACGCGGCGGCU	TC234269	Sulfite oxidase
Mebane-1_I	GCACUUUGAUUUCUCGUAAGGUGC	DT462219	Mitogen-activated protein kinase 19
Mebane-1_I	CAAGACGGGCCGAAUGGGGA	FF403833	Extensin-like protein
Mebane-1_I	GCAGCCGACGGGUUCGGGACUG	AJ513840	LL-diaminopimelate aminotransferase
Mebane-1_I	GACACGUGCCCUUGGGGGCC	TC231794	Chalcone synthase
Mebane-1_I	GACCAAUUCUCAAGCUGGGC	EV483778	Fiber annexin
Mebane-1_I	GACUUGCGCUUACUAGGGAUUC	BF278402	Fiddlehead-like protein
Mebane-1_I	GAGCCAAUCCUUUUCCCGAGG	TC248549	Elongation factor 1-alpha
Mebane-1_I	GAGCCAAUCCUUUUCCCGAG	ES837750	Ethylene response factor 1
Mebane-1_I	UGGAGCGGAAGGAGGGACUUG	CO086780; EX166061	SNAP25 homologous protein SNAP33; Zinc finger A20 and AN1 domain-containing stress-associated protein 5
Mebane-1_I	UCACGUCUGCCUGGGUGUCA	ES847499	MADS box protein GHMADS-1
Mebane-1_NI	UUGGUGUGAAUUGCAGAAUCC	EX167305	Thaumatin-like protein
Mebane-1_NI	AGCCCGCUCCGCCAUCCUGAG	ES791930	Enolase
Mebane-1_NI	GGGGCGCCGGAGAGGGUGAG	ES797338; TC234920	Gibberellin 20-oxidase 1; glutathione S-transferase
Mebane-1_NI	GCUGAGCAGUGUGGUGGAGGUG	DT466371	BHLH transcription factor
Mebane-1_NI	GCCCCCAUCCGCUUCCCUCCCG	ES850508	E3 ubiquitin-protein ligase At1g63170
Mebane-1_NI	CCUUAGCUACUCGAGCCUCAG	BF276096	Fiber protein Fb7
Mebane-1_NI	UGGCAUCCAGGGAGCCAGGC	TC273654	Proline-rich protein precursor
Mebane-1_NI	UCCCAACUGAGCUAUCCCCG	TC246587	Nucleobase-ascorbate transporter 4
Mebane-1_NI	AGAAUCUUGAUGAUGCUGCA	TC251459	Floral homeotic protein APETALA 2
Mebane-1_NI	GAGGGAGCCGUUCCCAGUCC	TC247373	Auxin response factor 1
Mebane-1_NI	UGAUCCUCCGUGGCUAGCUU	TC259572	Histone H2B
“11970”_I	GGGCCAGUCCGGACGCCUCACU	TC230218	Chitinase-like protein
“11970”_I	GCGACCGCCUGGGUGUCACGC	ES829849	Shaggy-related protein kinase alpha
“11970”_I	GCCAAGGGCAUCAACACUCGG	ES822278	Ubiquitin carrier protein
“11970”_I	AGACCGCUCCGCCAUCCUUU	TC252466	Probable inosine-5′-monophosphate dehydrogenase
“11970”_I	AGUCCGGAGACGUCGGCGGG	TC241619	F6N23.26 protein
“11970”_I	CCCAGCCAAACUCCCCACCU	TC272269	Phospholipase D delta isoform 1a
“11970”_NI	UGGGUUCGAUUCCUACUGCCC	ES843586	Similarity to receptor protein kinase
“11970”_NI	CCAACCCUUGGGACCGACUACAG	CO128532	Ulp1 protease family protein
“11970”_NI	CCGACUCGUAGACAGCGCCU	ES842843	Myristoyl-acyl carrier protein thioesterase
“11970”_NI	GGAGCAGCUUGAGCAGUGCGCCG	CO108709	ATPase 8, plasma membrane-type
“11970”_NI	UCAGUUCCCCGGGUCUGCCU	BQ402255	Lipid transfer protein 3 precursor
“11970”_NI	CCAUACUCCCCCCGGAACCC	TC273008	E6
“11970”_NI	GUGAAGUGUUCGGAUCGCGG	TC233952	Leucoanthocyanidin reductase 2
“11970”_NI	CCUACCGGGUUACCGAAUUCA	EV484549	Cp10-like protein
“11970”_NI	ACCAAACAAGGCCACACAUC	ES830500	HMG protein
“11970”_NI	UCUGCCAAGCCCGUUCCCUU	TC258657	3-ketoacyl-CoA synthase
“11970”_NI	CACACCAAGUAUCGCAUUUCGC	CD4866566; TC235961	Lipase, Polygalacturonase-like protein

Protein annotations were performed in the Plant Target Finder database (http://bioinfo3.noble.org/psRNATarget/) using the *Gossypium* DFGI gene index release 11 (dated on November 15, 2011)

I FOV infected, *NI* FOV noninfected

bioinfo3.noble.org/psRNATarget; the Gossypium DFGI gene index release 11, dated on November 15, 2011). We chose less than three mismatches between sRNA to identify best targets. One hundred and seventy target genes (40 to 50 per library) from *Gossypium* L. and *Arabidopsis thaliana* genomes were identified using the sRNA sequences in our study (see Supplementary Table S2). Several sRNA sequences targeted more than two genes. These targeted genes may play important roles in biological process, cellular component, and metabolic process in plants (Table 4). Specifically, as expected, novel cotton Ghr-miR-160 sequence targeted *G. hirsutum* auxin response factor 3 (ARF3) as expected because plant miR-160 targets the auxin signaling protein that plays an important role in plant growth and development, lateral root formation, and shoot elongation (Hongyan et al. 2011). This target can regulate a variety of developmental and physiological processes. In addition, plant miR-160 family targets are highly conservative in plants and include ARF10, ARF16, and ARF17 that regulate leaf development and floral identity (Mallory and Vaucheret 2006). *G. hirsutum* auxin responsive genes play key roles in plant development that might be involved in seed germination.

Biological Process and Molecular Function Annotation of Targets

To better understand putative functions of our sRNA sequence signatures, we examined identified targets using Gene Ontology (GO) analysis (http://www.uniprot.org/uniprot/). GO analysis revealed that a few target genes were associated with only biological process or molecular function, while most were associated with both terms (see Supplementary Table S2 and Fig. S2). Our annotation results show that target genes from the four libraries are playing important roles in catabolic, biosynthetic, and metabolic process in the cell. We identified a number of target gene functions of the FOV race 3-resistant Mebane B-1 and susceptible 11970 genotypes under two infected and noninfected conditions (Fig. 4), where GO terms connected to many important biological processes such as metabolic, catabolic, transport, ox/reduction and biosynthesis, and/or key molecular functions including enzymatic and binding properties with incorporation or reduction of molecular oxygen or hydrolyzing *O*-glycosyl compounds (Supplementary Table S2 and Fig. S2).

Many of these metabolic and biological processes, regulated with small RNAs characterized from FOV infection process, explain the pathogenesis and suggest the critical role of these processes in cotton and FOV pathogen interaction. Particularly, FOV-infected cotton plants have wilting symptoms accompanying with loosing water, where compounds controlling osmotic and water deprivation stress (e.g., A20/AN1 zinc-finger domain-containing SAP5 proteins; Vij and Tyagi 2008; Hozain et al. 2012) as well as key enzymes (e.g., CHS), facilitating the production of important metabolites (such as flavanoids, isoflavonoid-type phytoalexins; Dao et al. 2011) and protecting plant from various environmental stresses (e.g., salycilic acid defense pathway; Zheng et al. 2015), were targeted by sRNA/miR signatures characterized in our study. Interestingly, only Mebane B-1 which is a FOV-resistant genotype GO terms was connected to resistance process including response to abscisic acid stimulus, response to mechanical stimulus, response to other organism, response to water deprivation, and response to sulfur dioxide under FOV infection, which are known important factors in plant disease resistance (Dong 1998). Results on involvement in key biological processes and molecular function of small RNA and microRNA-targeted proteins demonstrate important molecular compounds conditioning the cotton resistance to pathogenic fungi Fusarium that add to our understanding of the possible mechanisms of FOV resistance in cotton. Similar molecular factors have been reported as associated compounds in other studies (Li et al. 2012), implying vital role of small RNAs in FOV infection.

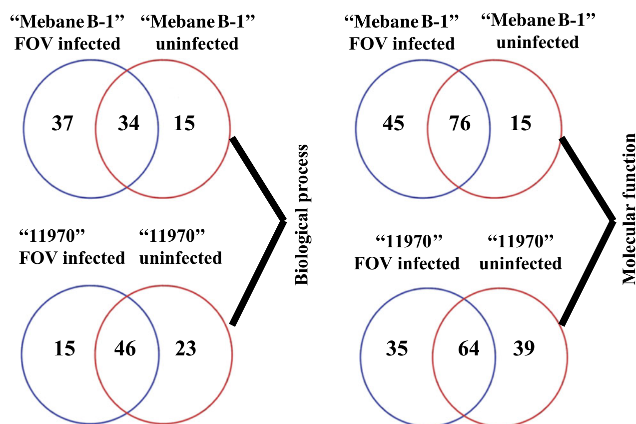


Fig. 4 The total number of potential target genes by GO assignments in four libraries

Conclusions

Thus, in this novel study, we investigated abundantly represented portion of small noncoding RNAs in FOV-infected and noninfected conditions using FOV-resistant and susceptible upland cotton genotypes. Using both small-scale experimental and bioinformatic tools, we characterized a pool of cotton root tissue-expressed sRNAs including miRs associated with FOV pathogenesis in Upland cotton. Although there is a need for full profiling of sRNA/miRs from our four libraries using new generation DNA sequencing tools as well as for functional verification of the expression of specific sRNAs from each library and condition/genotype, a snapshot of sequence characterization of sRNA/miR signatures and their putative targets in four transcriptome libraries, reported herein, highlighted the

importance of sRNA/miR signatures in posttranscriptional gene regulation of host-pathogen defense response during FOV pathogenesis in cotton. These results should be helpful to advance our understanding of molecular mechanisms underlying the cotton defense response against FOV, and results will help to develop innovative genetic engineering and biotechnology tools to improve FOV resistance of cotton cultivars.

Funding This project was funded by the Office of International Research Programs (OIRP) of the United States Department of Agriculture (USDA)—Agricultural Research Service (ARS) in the frame of USDA-ARS—Former Soviet Union (FSU) cooperation programs with the research grant of UZB2-31016-TA-09. We acknowledge the Academy of sciences of Uzbekistan and US Civilian Research & Development Foundation (CRDF) for project coordination. Mention of trade names or commercial products in this article is solely for providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The US Department of Agriculture is an equal opportunity provider and employer.

References

Abdurakhmonov IY, Devor EJ, Buriev ZT, Huang L, Makamov A, Shermatov SE, Bozorov T, Kushanov FN, Mavlonov GT, Abdukarimov A (2008) Small RNA regulation of ovule development in the cotton plant, *G. hirsutum* L. BMC Plant Biol 8:93. doi:10.1186/1471-2229-8-93

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297

Dao TT, Linthorst HJ, Verpoorte R (2011) Chalcone synthase and its functions in plant resistance. Phytochem Rev 10:397–412

Devor EJ, Huang L, Abdukarimov A, Abdurakhmonov IY (2009) Methodologies for in vitro cloning of small RNAs and application for plant genome(s). Int J Plant Genomics. doi:10.1155/2009/915061

Dong X (1998) SA, JA, ethylene, and disease resistance in plants. Curr Opin Plant Biol 1:316–323

Egamberdiev SS, Ulloa M, Saha S, Salakhutdinov IB, Abdullaev A, Glukhova LA, Adylova AT, Scheffler BE, Jenkins JN, Abdurakhmonov IY (2013) Molecular characterization of Uzbekistan isolates of *Fusarium oxysporum* f. sp. *Vasinfectum*. J Plant Sci Mol Breed. doi:10.7243/2050-2389-2-3

Ellendorff U, Fradin EF, de Jonge R, Thomma BPHJ (2009) RNA silencing is required for Arabidopsis defense against Verticillium wilt disease. J Exp Bot 60:591–602

Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangl JL, Carrington JC (2007) High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of miR genes. PLoS One 2, e219

He X-F, Fang Y-Y, Feng L, Guo H-S (2008) Characterization of conserved and novel microRNAs and their targets, including a TuMV-induced TIR-NBS-LRR class R gene-derived novel miR in Brassica. FEBS Lett 582:2445–2452

Hongyan X, Pudake RN, Guo G, Xing G, Hu Z, Zhang Y, Sun Q, Ni Z (2011) Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. BMC Genomics 12: 178

Hozain M, Abdelmageed H, Lee J, Kang M, Fokar M, Allen RD, Holaday AS (2012) Expression of AtSAP5 in cotton up-regulates putative stress-responsive genes and improves the tolerance to rapidly developing water deficit and moderate heat stress. J Plant Physiol 169:1261–1270

Jagadeeswaran G, Saini A, Sunkar R (2009) Biotic and abiotic stress down-regulate miR398 expression in Arabidopsis. Planta 229: 1009–1014

Khraiwesh B, Zhu JK, Zhu J (2011) Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. BBA Gene Regul Mech 2:137–148

Kim Y, Davis RM, Hutmacher RB (2005) Characterization of California isolates of *Fusarium oxysporum* f. sp. *Vasinfectum*. Plant Dis 89: 366–372

Kwak PB, Wang QQ, Chen XS, Qiu CX, Yang ZM (2009) Enrichment of a set of microRNAs during the cotton fiber development. BMC Genomics 10:457. doi:10.1186/1471-2164-10-457

Lee RC, Feinbaum RL, Anbros V (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75:543–854

Li Y, Zhang QQ, Zhang J, Wu L, Qi Y, Zhou JM (2010) Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity. Plant Physiol 152:2222–2231

Li C, Deng G, Yang J, Viljoen A, Jin Y, Kuang R, Zuo C, Lv Z, Yang Q, Sheng O, Wei Y, Hu C, Dong T, Yi G (2012) Transcriptome profiling of resistant and susceptible Cavendish banana roots following inoculation with *Fusarium oxysporum* f. sp. *cubense* tropical race 4. BMC Genomics 13:374

Lu S, Sun YH, Amerson H, Chiang VL (2010) MicroRNAs in loblolly pine (*Pinus taeda* L.) and their association with fusiform rust gall development. Plant J 51:1077–1098

Mallory AC, Vaucheret H (2006) Functions of microRNAs and related small RNAs in plants. Nat Genet 38:S31–S37

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N et al (2006) A plant miR contributes to antibacterial resistance by repressing auxin signaling. Science 312:436–439

Navarro L, Jay F, Nomura K, He SY, Voinnet O (2008) Suppression of the microRNA pathway by bacterial effectors proteins. Science 321: 964–967

Ulloa M, Hutmacher RB, Davis RM, Wright SD, Percy R, Marsh B (2006) Breeding for Fusarium wilt race 4 resistance in cotton under field and greenhouse conditions. J Cotton Sci 10:114–127

Ulloa M, Wang C, Hutmacher RB, Wright SD, Davis RM, Saski CA, Roberts PA (2011) Mapping Fusarium wilt race 1 resistance genes in cotton by inheritance, QTL, and sequencing composition. Mol Genet Genomics 286:21–36

Ulloa M, Hutmacher RB, Roberts PA, Wright SD, Nichols RL, Davis RM (2013) Inheritance and QTL mapping of Fusarium wilt race 4 resistance in cotton. Theor Appl Genet 126:1405–1418

Vij S, Tyagi AK (2008) A20/AN1 zinc-finger domain-containing proteins in plants and animals represent common elements in stress response. Funct Integr Genomics 8:301–307

Wu Y, Llewellyn DJ, Dennis ES (2002) A quick and easy method for isolating good-quality RNA from Cotton (*G.hirsutum* L.) tissues. Plant Mol Biol Report 20:213–218

Xin M, Wang Y, Yao Y, Xie C, Peng H, Ni Z, Sun Q (2010) Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (Triticum aestivum L.). BMC Plant Biol 10:123

Yin Z, Li Y, Han X, Shen F (2012) Genome-wide profiling of miRNAs and other small non coding RNAs in the Verticillium dahliae-inoculated cotton roots. PLoS One 7(4), e35765. doi:10.1371/journal.pone.0035765

Zhang B, Pan X, Cobb GP, Anderson TA (2006) Plant microRNA: a small regulatory molecule with big impact. Dev Biol 289:3–6

Zheng XY, Zhou M, Yoo H, Pruneda-Paz JL, Spivey NW, Kay SA, Dong X (2015) Spatialand temporal regulation of biosynthesis of the plant immune signal salicylic acid. Proc Natl Acad Sci U S A 112:9166–9173

View publication stats