Transcriptome analysis of host-pathogen interaction between sugarcane and *Colletotrichum falcum* by Suppression Subtractive Hybridization and Illumina sequencing

R Viswanathan, M Sathyabhama, P Malathi and A Ramesh Sundar

*Plant Pathology Section, ICAR-Sugarcane Breeding Institute, Coimbatore 641007, India; rasaviswanathan@yahoo.co.in*

**Abstract**  A high throughput analyses of interaction between sugarcane and *Colletotrichum falcum*, the hemibiotrophic pathogen causing red rot, was made by comparing transcriptomes of compatible and incompatible interactions in a tropical sugarcane cultivar Co7805 exhibiting differential reaction to the pathotypes Cf94012 and Cf87012, respectively. Suppression subtractive hybridization (SSH) combined with Illumina 2000 high throughput sequencing was used to identify the differential transcripts in resistance response library (RRL) and susceptible response library (SRL). After processing and filtering the raw reads from Illumina 2000 high throughput sequencing of subtracted products, we obtained 10,038 and 4,022 high quality transcripts, from RRL and SRL, respectively. Based on the transcripts mapping to KEGG-KASS database, the presence of a CEBiP receptor and the signals ROS, Ca^{2+}, BR, JA and ABA were identified in both the responses. However, MAPK, ET, PI signals and JA amino conjugation were found only in the incompatible interaction and expression of 10 transcripts involved in these pathways was validated using qRT-PCR. Our study concludes that perception of PAMPs occurs in both systems, but downstream signaling through MAPK, ET, PI and JA amino conjugation and activation of R genes occurs only in the incompatible interaction. This is the first detailed transcriptomic analysis of compatible and incompatible interactions in sugarcane with two different *C. falcum* pathotypes through SSH and the next generation sequencing (NGS) platform.

**Key words** Sugarcane, *Colletotrichum falcum*, Illumina sequencing, resistance and susceptible responses, KEGG-KASS, PRR, signaling pathways, qRT-PCR

**INTRODUCTION**

Red rot, caused by the ascomycete fungal pathogen *Colletotrichum falcum* Went (Teleomorph: *Glomerella tucumanensis* (Speg.) Arx and Muller), is a serious threat to sugarcane cultivation in India and many other sugarcane-growing countries. Management of the disease depends mainly on cultivation of disease-resistant cultivars. However, varietal replacement is a common feature in India because red rot resistance in sugarcane varieties is unstable due to the emergence of virulent pathotypes of *C. falcum*. Due to the complex polyploidy of sugarcane, heritability of red rot resistance in sugarcane is poorly understood (Viswanathan 2010).

Our early work revealed induction of pathogenesis related (PR) proteins and 3-deoxyanthocyanidin phytoalexins as a defense and induced response against *C. falcum* (Malathi *et al.* 2008; Ganesh Kumar *et al.* 2015; Viswanathan *et al.* 2003, 2005). Recently, a chitinase gene from sugarcane has been characterized as a class IV glycosyl hydrolase and its differential expression in sugarcane cultivars varying in red rot resistance was monitored through qRT-PCR (Rahul *et al.* 2015). Further, molecular techniques such as differential display (DD)-RT-PCR and suppression subtractive hybridization (SSH) were applied to study differential expression of the transcripts during the host-pathogen interaction. Upregulation of transcripts associated with jasmonic acid (JA), ethylene (ET), phosphoinositide (PI) and calcium (Ca^{2+}) signals, defense, reactive oxygen species (ROS) and the secretory pathway in red rot resistant cultivar were identified (Prathima *et al.* 2013; Rahul *et al.* 2016; Sathyabhama *et al.* 2015, 2016). In these studies, we used cane tissue inoculated with *C. falcum* or suspension cells treated with *C. falcum* elicitor to identify the differentially expressed transcripts. In the experimental trials, some cultivars of sugarcane exhibited differential responses to various *C. falcum* pathotypes over time (Viswanathan 2010). In our earlier transcriptomic studies, a set of sugarcane cultivars varying in red rot resistance was used to determine differential transcripts involved in *C. falcum* defense (Sathyabhama *et al.* 2015, 2016). The cultivars varying in red rot resistance may have variable genetic background and that may have a profound influence on identifying the genes/proteins of sugarcane involved in host defense.
In this study we chose a sugarcane host exhibiting compatible and incompatible interactions to two different \textit{C. falcatus} pathotypes to identify transcripts involved in defense responses. Two SSH libraries were constructed from sugarcane cv Co7805 exhibiting resistance response (RR) and susceptible response (SR) to \textit{C. falcatus} pathotypes \textit{C}87012 and \textit{C}94012, respectively. Illumina Hiseq 2000 sequencing platform was used to sequence subtracted transcripts from the two responses. About 14,060 high quality transcripts were obtained from the libraries. The transcripts are found to be involved in perception of PAMPs, signaling and activation of R genes in an incompatible host-pathogen interaction.

**MATERIALS AND METHODS**

**Plant material and pathogen culture**

The tropical sugarcane cultivar Co7805 was grown in the Institute field by following standard cultivation practices for a tropical sugarcane. Two pathotypes of \textit{C. falcatus}, \textit{C}87012 and \textit{C}94012, isolated from infected stalk tissues of sugarcane cvs Co87012 and Co94012, respectively, maintained in \textit{C. falcatus} culture collections, Plant Pathology lab of the Institute were used in this study. Consistent behaviour of the Co7805 as resistant and susceptible to the pathotypes \textit{C}87012 and \textit{C}94012, respectively, was confirmed earlier. The pathogen was inoculated in the third internode from the base of the cane by the standard plug method (Viswanathan 2010). Separate sets of sugarcane stalks were inoculated in triplicates with the two \textit{C. falcatus} pathotypes. Mock inoculated samples served as controls for the experimental samples. Stalk tissue samples were collected in triplicates with their respective controls 72 h post \textit{C. falcatus} inoculation.

**Suppression subtractive hybridization, Illumina library construction and sequencing**

Total RNA was extracted from all the samples with TRI reagent (Sigma-Aldrich, USA). The quality of RNA was checked in an agarose gel and quantified in a NanoDrop™ 1000 spectrophotometer (Thermo Scientific, USA). Forward and reverse subtractions for the cDNAs were done following the manufacturer’s instructions of PCR-Select™ cDNA subtraction kit (Clonetech, USA). Forward subtraction represents resistance response library (RRL), in which cv Co7805 challenged with \textit{C}87012 was used as a tester, the same cultivar challenged with \textit{C}94012 and a mock sample were used as the driver. In reverse subtraction, susceptible response library (SRL), cv Co7805 challenged with \textit{C}94012 was used as a tester, the same cultivar challenged with \textit{C}87012 and the mock sample were used as the driver. The cDNA pools of the subtracted two transcripts were sequenced by Illumina HiSeq 2000 paired end (PE) sequencing platform (Xcelris Genomics Pvt. Ltd, Ahmadabad, India). The transcript assembled contigs that belong to the metabolic pathways that are expressed in the interaction were identified through mapping the assembled transcripts to Kyoto Encyclopedia for Genes and Genomes (KEGG) eukaryotic database using KEGG Automatic Annotation Server (KAAS). All the transcripts were compared against KEGG-KAAS database using BLASTX with default threshold bit-score value of 60.

**Validation of gene expression through qRT-PCR**

We performed qRT-PCR assays for 10 transcripts involved in recognition, signaling and defense mapped to KEGG-KAAS database to validate their gene expression in sugarcane after \textit{C. falcatus} inoculation (Step One Plus™ Real-Time PCR Systems, Life Technologies, USA). The transcripts included in the qRT-PCR assays are listed in Table 2. Tissue samples of the sugarcane cv Co7805 stalks inoculated with \textit{C. falcatus} pathotypes \textit{C}87012 and \textit{C}94012 were collected at 12, 36 and 72 h post \textit{C. falcatus} inoculation. Total RNA was extracted as mentioned before and cDNA was synthesized from 500 µg of total RNA. The first strand cDNA was diluted further to 200 ng and used as templates for qRT-PCR. SYBR green PCR mastermix (Life Technologies, USA) was used for the assays with 25S rRNA as the internal control. The samples were run in triplicates and relative quantitation (RQ) values determined by the method of Livak and Schmitgen (2001).

**RESULTS**

**Transcriptome sequencing, de novo assembly and KEGG-KAAS functional annotation**

The number of raw reads generated by Illumina HiSeq 2000 were 54,699,263 and 52,509,239 for RRL and SRL, respectively and after filtering the same had 41,025,151 and 42,001,812 reads. Quality trimming, adaptor sequence removal and size selection of transcript reads resulted in a total of 10,038 and 4,022 high quality reads for RRL and SRL, respectively (Table 1). After analyses the two set of the transcripts were mapped in KEGG-KAAS database to 12 categories
pertaining to carbohydrate metabolism, energy metabolism, lipid metabolism, nucleotide metabolism, amino acid metabolism, glycogen metabolism and biosynthesis, metabolism of cofactors and vitamins, metabolism of terpenoids and polyketides, biosynthesis of other secondary metabolites, genetic information processing, environment information processing and plant pathogen interaction. Of the total transcripts mapped, 42% were present in both the interactions and 47% of the transcripts were upregulated in RRL and 11% of the transcripts were upregulated in SRL.

**Table 1. Summary of sequencing, de novo assembly and annotation of sugarcane differential host interaction transcriptome.**

<table>
<thead>
<tr>
<th>Description (numbers)</th>
<th>RR</th>
<th>SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw reads</td>
<td>54,699,263</td>
<td>52,509,239</td>
</tr>
<tr>
<td>Filtered reads</td>
<td>41,025,151</td>
<td>42,001,812</td>
</tr>
<tr>
<td>Assembled transcripts</td>
<td>10,038</td>
<td>4,022</td>
</tr>
<tr>
<td>Transcripts with BLAST hits</td>
<td>7,849</td>
<td>2,899</td>
</tr>
<tr>
<td>Transcripts without BLAST hits (novel/hypothetical genes)</td>
<td>2,189</td>
<td>1,124</td>
</tr>
</tbody>
</table>

In all the categories, several transcripts were mapped in common, i.e. those transcripts were present in both the responses. However, in carbohydrate metabolism, hexokinase, probable phosphoglycerate mutase, and pyruvate dehydrogenase E1 component subunit beta were present in RRL and absent in SRL. In energy metabolism, an important enzyme succinate dehydrogenase (ubiquinone) flavoprotein subunit was upregulated only in RRL. A few transcripts upregulated in RRL involved in terpenoid biosynthesis include phosphomevalonate kinase, diphosphomevalonate decarboxylase, famesyl diphasate synthase and geranylglycerol diphasate synthase. In the MAPK signaling pathway, extracellular signal regulated kinase was present in both the responses, and protein phosphatase 3, regulatory subunit and MAPK kinase 1 were found only in the RRL. In the calcium signaling pathway, phosphatidylinositol phospholipase C, delta and protein phosphatase 3 regulatory subunit were upregulated in RRL and calmodulin was upregulated in SRL. In plant hormone signal transduction, the transcripts involved in auxin, abscisic acid, brassinosteroid and JA signaling were present in both the interactions. The transcripts were auxin-responsive protein IAA, auxin response factor, abscisic acid receptor PYR/PYL family, protein phosphatase 2C, serine/threonine-protein kinase SRK2, ABA responsive element binding factor, BR-signaling kinase, coronatine-insensitive protein 1 and jasmonate ZIM domain-containing protein. Two-component system components such as Arabidopsis histidine kinase 2/3/4 (cytokinin receptor), histidine-containing phosphor transfer protein and two-component response regulator ARR-B family were upregulated only in RRL. The ethylene signaling components, ethylene receptor, ethylene-insensitive protein 2, ethylene-insensitive protein 3, EIN3-binding F-box protein, jasmonic acid-amino synthetase, transcription factor TGA and pathogenesis-related protein 1 were also upregulated only in RRL. In RRL, E2f transcription factor 1/3, glycogen synthase kinase 3 beta, cyclin-dependent kinase 2, anaphase-promoting complex subunit 6, 10 and 11, cohesion complex subunit, cell cycle arrest protein, etc were upregulated only in RRL.

The Ub-activating enzyme E1 and isoforms of Ub-conjugating enzymes E2 were present in both the responses. The E3 Ub-protein ligases TRIP12, NEDD4, HUWE1 and SIAH1, anaphase-promoting complex (APC) subunits were upregulated only in RR. S-phase kinase associated protein1 (Skp1), cullins 1, 3 and 4 were present in both the interactions. The well-known RING-box protein 1 was found only in RR. F-box and LRR protein1 (Skp2) and a DNA damage-binding protein 1 was found only in SR. The secondary metabolite transcripts, namely phenylalanine ammonia lyase (PAL), 4-coumarate CoA ligase (4-CouCoAL), coniferyl alcohol dehydrogenase (ConADH), cinnamyl alcohol dehydrogenase (CinAlcDH), phosphomevalonate kinase (PMK) and phytoene synthase (PhyS), were present only in RR.

**Gene expression through qRT-PCR**

The gene expression of the transcripts CEBiP, MAPKKK1, MAPKK1, DRPRPM1, DRPRPS5 and CBPCML showed a gradual increase in their expression at 12 h and 36 h and a decline at 72 h post *C. falcatum* inoculation in RR. Whereas, in SR, CEBiP, MAPKKK1 and CBPCML exhibited an inconsistency in their expression at all the time intervals. MAPKK1 showed gradual decrease from 12 h to 72 h. DRPRPM1 showed a similar response as RR but the transcript level was less than 2 fold. The transcripts CNGC and CDPK showed a gradual increase in RR whereas a gradual decrease was seen in SR from 12 to 72 h post *C. falcatum* inoculation. Transcript accumulation was less than 2 fold in SR for both the genes. The transcript SOD Cu Zn showed an increase in expression from 12 to 36 h and a decline at 72 h in the SR whereas in
RR, there was a drastic reduction in its expression after 12 h post *C. falcum* inoculation. BRSK showed an inconsistent expression in both the responses. At 12 h, BRSK showed more than 5 fold expression in RR. Overall, RR revealed higher expression of differential transcripts upon *C. falcum* inoculation whereas in SR, except for few transcripts, the gene expression was not prominent (Table 2).

**Table 2.** Validation of selected genes expression in sugarcane cv Co7805 after challenge with two pathotypes by qRT-PCR.

<table>
<thead>
<tr>
<th>Sl No</th>
<th>Transcript</th>
<th>Host reaction</th>
<th>RQ values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Control 12 h 36 h 72 h</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>CEBiP</td>
<td>RR</td>
<td>1.0   7.0 11.3 6.6</td>
</tr>
<tr>
<td>2</td>
<td>MAPKK1</td>
<td>RR</td>
<td>1.0   1.9 4.6 3.5</td>
</tr>
<tr>
<td>3</td>
<td>MAPKK1</td>
<td>RR</td>
<td>1.0   1.6 3.2 2.5</td>
</tr>
<tr>
<td>4</td>
<td>BRSK</td>
<td>RR</td>
<td>1.0   9.2 1.4 1.1</td>
</tr>
<tr>
<td>5</td>
<td>DRPRPM1</td>
<td>RR</td>
<td>1.0   1.6 2.5 1.5</td>
</tr>
<tr>
<td>6</td>
<td>DRPRPS5</td>
<td>RR</td>
<td>1.0   34.3 44.0 34.3</td>
</tr>
<tr>
<td>7</td>
<td>CNGC</td>
<td>RR</td>
<td>1.0   1.8 2.0 2.6</td>
</tr>
<tr>
<td>8</td>
<td>CDPK</td>
<td>RR</td>
<td>1.0   1.2 2.6 3.0</td>
</tr>
<tr>
<td>9</td>
<td>CBPCML</td>
<td>RR</td>
<td>1.0   1.0 0.4 0.3</td>
</tr>
<tr>
<td>10</td>
<td>SOD Cu Zn</td>
<td>RR</td>
<td>1.0   4.3 0.5 0.9</td>
</tr>
</tbody>
</table>

h – hours post *C. falcum* inoculation, RR – Resistant response, SR- Susceptible response.  
CEBiP - chitin elicitor binding protein, MAPKKK1 - mitogen activated protein kinase kinase kinase 1,  
MAPKK1 - mitogen activated protein kinase kinase 1, BRSK - brassinosteroid signalling kinase,  
DRPRPM1 - disease resistant protein RPM1, DRPRPS5 - disease resistant protein RPS5,  
CNGC - cyclic nucleotide gated channel, CDPK - calcium dependent protein kinase,  
CBPCML - calcium binding protein CML, SOD Cu Zn - superoxide dismutase Cu Zn.

**DISCUSSION**

Co-evolution of *C. falcum* with new sugarcane varieties has occurred in the past 100 years in India and many sugarcane varieties exhibited differential interaction to *C. falcum* pathotypes (Viswanathan 2010). In our study, we attempted to reduce variation in host genetic background by exploiting the differential behaviour of a sugarcane cultivar to two *C. falcum* pathotypes and two separate SSH cDNA populations were obtained. Additionally, the NGS platform was applied to sequence the transcripts induced variably during the two interactions. We were successful in the attempt and obtained ~14,060 high quality transcripts induced during RR and SR. Critical analyses of these transcripts generated new information on the variable transcript accumulation during RR and SR against *C. falcum*. Earlier studies involving different molecular techniques revealed limited information on specific induction of transcripts and proteins during sugarcane and *C. falcum* interaction (Muthumeena et al. 2013; Rahul et al. 2016; Sathyabhama et al. 2015, 2016; Viswanathan 2012). These studies revealed a strong induction of WRKY transcription factors from 6 to 48 h post *C. falcum* inoculation. In the present study, transcripts homologous to CEBiP and STPK were found in RRL and SRL. The kinases involved in MAPK cascade, MAPKKK1 and MAPKK1 were found only in RR and absent in SRL. MAPK signal cascade directly targets the nucleus in the presence of WRKY transcription factor and activates defense gene expression. In case of compatible interaction, there was induction at 12h and the expression was unstable thereafter. Even though CEBiP and STK were found in both RR and SR, the downstream signaling cascade, the MAPK cascade is not activated in SR. The perception of pathogenic determinants must be a common phenomenon in both the responses. However, further defense responses involve a cascade of signaling events. Elevated levels of JA need ET and both act synergistically to trigger defense gene expression and defense related secondary metabolite biosynthesis (Bilgin et al. 2010). Linoleate 9S lipoxigenase, the
enzyme involved in JA biosynthesis was found in both the interactions. Previously, the isoform of lipoygenase was studied in compatible and incompatible interactions between sugarcane and C. falcum. Throughout the study, the temporal expression showed unstable expression and was found to be enhanced in the incompatible interaction (Sathyabhama et al. 2015).

Ubiquitin proteasome system (UPS) is involved in defense responses against a wide range of pathogens (Dielen et al. 2010). In both the defense responses, transcripts homologous to several 26S proteasome regulatory subunits N and T, 20S proteasome subunits alpha and beta and proteasome maturation protein were present. The proteasome inhibitor subunit 1 was present only in SR. During host-pathogen interactions, activation of defense genes is the primary response and in this study, the defense genes namely DRPRPM1 and DRPRPS5 were upregulated only in RR and absent in SR. Overall, our results very clearly indicated differential expression of transcripts involved in defense, signaling, carbohydrate and energy metabolism and other functional groups in sugarcane between compatible and incompatible interactions in response to C. falcum infection. Moreover, transcript induction is high in RR as compared to SR probably due to the complex network of activating defense responses.

CONCLUSIONS

We have identified that the signaling responses ROS, Ca2+, BR, JA and ABA were present in both RR and SR responses in sugarcane. In addition, MAPK, ET, PI signals and JA amino conjugation were present only in RR. Perception of PAMPs may be a common mechanism, but in RR downstream signaling and activation of R genes was rapid and this resulted in activation of transcripts involved in defense and secondary metabolites. Our study has given new insights that the defense responses in sugarcane vary with respect to C. falcum pathotypes. Further studies on characterizing PAMP/effector molecules of the pathogen could throw more light on the molecular interaction between the host and the pathogen.

ACKNOWLEDGMENTS

The research work was supported by ICAR, New Delhi under ‘Outreach Project on ALCOCERA’.

REFERENCES


Analyse transcriptomique de l'interaction hôte-pathogène entre la canne à sucre et Colletotrichum falcatum par hybridation suppressive soustractive et séquençage Illumina

Résumé. Une analyse à haut débit de l'interaction entre la canne à sucre et Colletotrichum falcatum, pathogène hémibiotrophe causant la morve rouge, a été effectuée en comparant les transcriptomes des interactions compatibles et incompatibles chez Co7805, cultivar qui présente des réactions différentielles aux pathotypes CF4012 et CF87012. L'hybridation suppressive soustractive (SSH) et le séquençage à haut débit Illumina 2000 ont été utilisés pour identifier les transcripts différentiels dans la bibliothèque de réponse de la résistance (RRL) et dans la bibliothèque de réponse de la sensibilité (SRL). Après avoir traité et filtré les lectures de séquences brutes fournies par le séquençage à haut débit Illumina 2000 des produits soustraits, nous avons obtenu 10 038 et 4 022 transcripts de haute qualité dans respectivement la RRL et la SRL. La cartographie des transcripts dans la banque de données KEGG-KASS a permis d'identifier dans les deux réponses la présence d'un récepteur CEBiP et les signaux ROS, Ca²⁺, BR, JA et ABA. Cependant, les signaux MAPK, ET, PI et l’amo-no-conjugaison JA n’ont été trouvés que dans l’interaction incompatible, et l’expression de 10 transcripts impliquées dans ces voies a été validée par qPCR. Notre étude conclut que l’observation de PAMP (motifs moléculaires associés aux pathogènes) intervient dans les deux systèmes, mais que la signalisation en aval par l’intermédiaire de MAPK, ET, PI et l’amo-no-conjugaison JA ainsi que l’activation des gènes R n’interviennent que dans l’interaction incompatible. Il s’agit de la première analyse transcriptomique détaillée d’interactions compatibles et incompatibles chez la canne à sucre avec deux pathotypes différents de Colletotrichum falcatum en utilisant la SSH et une plateforme de séquençage de nouvelle génération.

Mots-clés: Canne à sucre, Colletotrichum falcatum, séquençage Illumina, réponses de la résistance et de la sensibilité, KEGG-KASS, PRR, voies de signalisation, qRT-PCR

Análisis transcriptómico de la interacción huésped-patógeno entre caña de azúcar y Colletotrichum falcatum mediante hibridación substractiva por supresión y secuenciamiento Illumina

Resumen. El análisis de la interacción entre la caña de azúcar y Colletotrichum falcatum, un patógeno hemibiotrofo causante de la pudrición roja, se realizó en la variedad Co7805 que presenta reacción diferencial a los patotipos CF94012 y CF87012, mediante la comparación de transcriptomas de interacciones compatibles e incompatibles, respectivamente. Se utilizó la hibridación substractiva por supresión (SSH), en combinación con la secuenciación de alto rendimiento Illumina 2000, para identificar transcripciones diferenciales en la biblioteca de respuesta resistente (RRL) y la biblioteca de respuesta susceptible (SRL). Después de procesar y filtrar las lecturas crudas de los productos sustraídos y secuenciados por illumina 2000, obtuvimos 10.038 y 4.022 transcripciones de alta calidad, desde la RRL y SRL, respectivamente. Basándose en el mapa transcriptómico de la base de datos KEGG-KASS, la presencia del receptor CEBiP y las señales de ROS, Ca²⁺, BR, AJ y ABA, se encontraron en ambas respuestas. Sin embargo, las señales MAPK, ET, IP y la conjugación AJ amino, sólo se encontraron en la interacción incompatible, mientras que la expresión de 10 transcripciones que participan en estas vías se validaron usando qRT-PCR. Nuestro estudio concluyó que la percepción de PAMP se presenta en ambos sistemas, pero la cascada de señalización corriente abajo a través de MAPK, ET, PI, la conjugación AJ amino y la activación de los genes R se produce sólo en la interacción incompatible. Este es el primer análisis transcriptomico detallado de la interacción compatible e incompatible de la caña de azúcar con dos patotipos diferentes de Colletotrichum falcatum a través de SSH y la plataforma de secuenciación de nueva generación (NGS).

Palabras clave: Caña de azúcar, Colletotrichum falcatum, secuenciamiento Illumina, respuesta de resistencia y susceptibilidad, KEGG-KASS, PRR, vías de señalización, qRT-PCR