Sugarcane yellow leaf virus impairs the transcriptomic defense responses of sugarcane to its new aphid vector *Melanaphis sorghi*

Ricardo José Gonzaga Pimenta[1], Alexandre Hild Aono[1], Danilo Augusto Sforça[2], Maria Natália Guindalini Melloni[3], Marcos Cesar Gonçalves[4], Luciana Rossini Pinto[3], Anete Pereira de Souza[1,2*]

[1]Institute of Biology, Universidade Estadual de Campinas (UNICAMP), Campinas, Brazil.

[2]Centre for Molecular Biology and Genetic Engineering (CBMEG), Universidade Estadual de Campinas (UNICAMP), Campinas, Brazil.

[3]Advanced Centre for Sugarcane Research and Development, Instituto Agronômico de Campinas (IAC/APTA), Ribeirão Preto, Brazil.

[3]Plant Protection Research Centre, Instituto Biológico (IB/APTA), São Paulo, Brazil.

***Corresponding author:** Anete Pereira de Souza. Address: Av. Cândido Rondon 400, Campinas, SP, Brazil, 13084-875. Email: anete@unicamp.br.

Abstract

Sugarcane (*Saccharum* spp.) is a preeminent sugar and bioenergy crop and has great economic importance in tropical countries. A major disease affecting this crop is yellow leaf disease, which is caused by sugarcane yellow leaf virus (SCYLV, *Polerovirus SCYLV*, *Solemoviridae*). The sugarcane aphid *Melanaphis sacchari* is considered the main vector of SCYLV, and the closely related sorghum aphid *Melanaphis sorghi*, which has recently emerged as a pest of great relevance in sorghum, has also been suspected as a vector. Genetic resistance is an important resource for preventing yield losses caused by SCYLV and its

vectors, but knowledge on the underlying molecular mechanisms is lacking. Therefore, the present work was aimed at investigating the transcriptomic responses of sugarcane to SCYLV and *M. sorghi*, which was reported to transmit this virus for the first time herein. Two sugarcane cultivars, one susceptible and one tolerant to SCYLV, were fed upon by aviruliferous and viruliferous aphids. The transcriptome of the plants was assessed by RNA-Seq using differential gene expression analyses and a gene coexpression network. The susceptible cultivar showed an incipient reaction to both *M. sorghi* and SCYLV, with very few differentially expressed genes (DEGs) identified in comparison with aphid-free plants. The response of the tolerant cultivar to aviruliferous *M. sorghi* involved pathways typically associated with defense against herbivory, which were also enriched in coexpression network modules in which DEGs were overrepresented. Some of these genes were hubs in their respective modules, indicating they are potential key regulators of the defense responses. However, these responses were diminished when viruliferous aphids were used, and other processes linked to infection with SCYLV were altered. These results indicated that SCYLV could affect sugarcane defense responses to its vector, similar to other viruses of the same genus. Some possible implications for the epidemiology and impact of SCYLV and *M. sorghi* are discussed.

Keywords: *Saccharum*, SCYLV, YLD, Sorghum aphid, RNA-Seq, Gene coexpression network

1. Background

Sugarcane (*Saccharum* spp.) is a major sugar and energy crop in tropical regions of the world, accounting for nearly 80% of the global production of sugar (ISO 2023). The yield of this crop is affected by several diseases, including yellow leaf disease (YLD), which is recognized mainly by the yellowing of abaxial leaf midribs of mature plants (Vega et al. 1997; ElSayed et al. 2015). This disease is caused by sugarcane yellow leaf virus (SCYLV, *Polerovirus SCYLV*), a positive-sense ssRNA virus belonging to the genus *Polerovirus* of the *Solemoviridae* family (Sõmera et al. 2021) present in all sugarcane-growing regions of the world (Komor et al. 2010). This virus is known to affect photosynthesis efficiency as well as the metabolism and transport of carbohydrates (Gonçalves et al. 2005; Lehrer et al. 2010; Viswanathan et al. 2014). This impairs growth and leads to substantial reductions in sugar productivity and tonnage, which reflect yield losses of up to 37% (Rassaby et al. 2003; Vasconcelos et al. 2009; Viswanathan et al. 2014).

SCYLV is transmitted in a persistent, circulative and nonpropagative manner by several aphid species, among which the sugarcane aphid *Melanaphis sacchari* (Zehntner, 1897) is considered the most relevant (Scagliusi and Lockhart 2000). *M. sacchari* is capable of reproducing by apomictic parthenogenesis and is widely distributed throughout the world (Singh et al. 2004). In sugarcane, economic losses caused by *M. sacchari* are usually not concerning, and the species is mostly regarded as an important vector of viruses (Hall 1987; White et al. 2001). In the last decade, a ‘superclone’ of a closely related species, the sorghum aphid *Melanaphis sorghi* (Theobald, 1904), has emerged as a major pest of sorghum (*Sorghum bicolor*) in the Americas (Nibouche et al. 2018, 2021). Although the transmission of SCYLV by *M. sorghi* has not been reported thus far, SCYLV has been detected in *S. bicolor* (Wei et al. 2016; ElSayed et al. 2018).

While the physiological consequences of infection by SCYLV and its impact on yield have been widely assessed in sugarcane, knowledge on the molecular mechanisms that underlie the response to this virus is still lacking. Genetic mapping studies targeting SCYLV resistance have been performed but are limited by the highly complex sugarcane genome, phenotyping constraints, and the quantitative nature of this trait (Debibakas et al. 2014; Pimenta et al. 2021). The ineffectiveness of the transmission of SCYLV by mechanical methods (Scagliusi and Lockhart 2000) has further restricted research on this virus, as it implies that data on the resistance to SCYLV and to its vectors are frequently intertwined. For instance, a marker associated with resistance to SCLYV identified by association mapping was found to be linked to a lipoxygenase gene associated with aphid resistance (Debibakas et al. 2014). Sugarcane resistance to aphids, which could help prevent infection by SCYLV, has been assessed at the phenotypic level (e.g., Akbar et al. 2011; Fartek et al. 2014; Bertasello et al. 2021), but its genetic basis has not been investigated in this crop.

With the decreasing cost of high-throughput next-generation sequencing technologies, RNA sequencing (RNA-Seq) has become a common method for investigating the mechanisms underlying biological phenomena. Research on sugarcane has contributed to the understanding of plant responses to pests and several pathogens, including viruses (Ling et al. 2018; Akbar et al. 2020). This type of analysis has not yet been undertaken for the feeding of sucking insects and has been performed for SCYLV infection only in a susceptible cultivar infested with aviruliferous and viruliferous *M. sacchari* (Shabbir et al. 2022). Assessing the responses of genotypes with contrasting responses to SCYLV, as well as those of aphid-free plants, can help elucidate the molecular mechanisms underlying resistance to both aphids and viruses, as well as their interactions.

In this study, the ability of *M. sorghi* to transmit SCYLV in sugarcane was demonstrated, with transmission rates reaching 100% after 48 hours of feeding.

Subsequently, the transcriptomic responses of sugarcane to SCYLV and *M. sorghi* were investigated using RNA-Seq. Two sugarcane hybrid cultivars, one susceptible and one tolerant to SCYLV, were exposed to either aviruliferous or viruliferous aphids. Comparing the transcriptomes of these plants with those of aphid-free controls and building a weighted gene coexpression network provided insights into the early molecular responses of sugarcane to SCYLV and *M. sorghi.*

2. Results

2.1. Viral isolate and aphid species identification

The complete genome of the SCYLV isolate infecting the source of inoculum, spanning 5,903 nucleotides, was obtained by Sanger sequencing and pairwise-aligned with publicly available genomes of 101 SCYLV isolates. The genome of this new isolate, named BRA-SP1, was found to be most closely related to that of BR-06-Leaf, which originated from a sugarcane sample collected in 2018 in Brazil. Forty-one polymorphisms were detected between the genomes of BRA-SP1 and BR-06-Leaf, which presented 99.3% similarity. Recombination analyses indicated that thirteen isolates included in the alignment were potential recombinants (Table S1), and these were thus excluded from the subsequent phylogenetic analysis to avoid confounding effects. A maximum likelihood tree was constructed with the remaining 89 aligned sequences using a general time reversible model with gamma-distributed invariant sites. The results indicated the existence of three well-supported major SCYLV clades, including all the isolates from Mauritius and Reunion (clade REU), a few isolates from Asia, the Americas and Papua New Guinea (clade CUB), and isolates from diverse locations (clade BRA). Isolate BRA-SP1 was positioned in clade BRA, close to other samples originating from Brazil (Figure 1).

A colony of aviruliferous *Melanaphis* sp. aphids collected from a sugarcane field was established, and a fragment of the *elongation factor 1-alpha* (*EF1-α*) gene harboring five polymorphic sites was sequenced from a pool of five insects. The nucleotides observed at positions 296, 478, 637, 663, and 735 were Y, T, A, G, and C, respectively. These results match the *EF1-α* haplotype "H2c" described by Nibouche et al. (2021), and thus the colony was identified as *M. sorghi*.

2.2. Inoculation experiment

Plants of the sugarcane cultivars SP71-6163 (SCYLV-susceptible) and IACSP95-5000 (SCYLV-tolerant) were exposed to *M. sorghi* aphids for varying periods to assess viral transmission. Reverse transcription followed by quantitative PCR (RT□qPCR) confirmed that SCYLV was not present in any plants exposed to aviruliferous aphids or plants not exposed to aphids, which were used as controls. The number and percentage of plants in which SCYLV was detected at different inoculation access periods (IAPs) are summarized in Table S2. The virus was detected in SP71-6163 after as little as four hours of exposure to *M. sorghi*, showing that this aphid is capable of vectoring SCYLV in sugarcane under controlled conditions. The virus was detected in 100% of the plants of both cultivars after only 48 hours of exposure to *M. sorghi*; thus, this IAP was chosen for subsequent RNA-Seq analysis. At this time point, the mean SCYLV relative quantification of SP71-6163 (0.0013) was significantly ($p = 0.03$) greater than that of IACSP95-5000 (0.0003).

2.3. Transcriptome reference assembly and quantification

The sequencing of the 18 RNA-Seq libraries yielded more than a billion paired-end reads; over 95% of these (961.2 million) could be attributed to the libraries, resulting in an average

of 53.4 million reads per sample. After trimming, all reads presented an average Phred score ≥ 24, and 99.7% of them presented a Phred score ≥ 30 (Table S3).

In conjunction with these data, trimmed RNA-Seq data from leaf samples of twelve diverse *Saccharum* genotypes (Correr et al. 2020) were used to construct a new reference sugarcane transcriptome. Filtered reads were aligned to the allele-resolved genomes of the two main ancestral *Saccharum* species: *Saccharum spontaneum* and *Saccharum officinarum*. The numbers of assembled transcripts for the *S. spontaneum* haplotypes were (A) 33,969, (B) 33,913, (C) 33,576 and (D) 33,672. For *S. officinarum*, the quantities were as follows: (A) 34,993, (B) 34,081, (C) 33,832, (D) 32,513, (E) 31,720, (F) 30,280, (G) 29,124 and (H) 25,926. Reduction of redundancies in the two references using CD-HIT resulted in 95,533 and 141,626 transcripts for *S. spontaneum* and *S. officinarum*, respectively. These two sets were then combined, yielding a final reference containing 206,426 transcripts with an N50 of 2,097 bases. Searches using BUSCO showed that this reference included 99.3 and 98.4% of the conserved orthologs of eukaryotes and green plants, respectively (Table S4). A total of 43,291 of these transcripts, hereafter referred to as “genes”, could be annotated with Trinotate.

On average, 89.4% of the reads from each RNA-Seq library generated in the present study were mapped to this reference (Table S3). The first two gene filtering steps based on count per million (CPM) values resulted in the retention of 85,685 expressed genes. A multidimensional scaling analysis performed with these data is shown in Figure S1, where the proximity between replicates of each genotype and biological condition is observed.

2.4. Differential gene expression analyses

After performing t tests, sets of genes to be excluded in each condition were obtained and combined according to the comparisons to be carried out in the differential gene expression

(DGE) analyses. The quantities of genes excluded in each condition and in the union of sets, as well as the number of genes retained for each comparison, can be found in Table S5. DGE analyses were performed for nine different comparisons. The number of differentially expressed genes (DEGs) identified in each group can be seen in Table 1, which distinguishes upregulated ($\log_2$(fold change) ($\log_2$(FC)) > 1) and downregulated ($\log_2$(FC) < 1) genes, and in Figure 2, which shows intersections between DEG sets. Additionally, Table S6 presents a complete list of DEGs and associated information. The expression patterns of eight DEGs were validated through RT□qPCR using *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) and Cluster_58111 as endogenous controls. Overall, strong and significant linear correlations between the CPM and ΔCq values were observed, with R^2 values ranging from 0.56 to 0.92 (Figure S2).

Table 1. Number of up- and downregulated differentially expressed genes in each comparison performed.

Comparison	Differentially expressed genes		
	Upregulated	**Downregulated**	**Total**
SP71-6163 NA vs. SP71-6163 AA	63	7	70
SP71-6163 NA vs. SP71-6163 VA	23	9	32
SP71-6163 AA vs. SP71-6163 VA	2	2	4
IACSP95-5000 NA vs. IACSP95-5000 AA	589	64	653
IACSP95-5000 NA vs. IACSP95-5000 VA	27	138	165
IACSP95-5000 AA vs. IACSP95-5000 VA	33	328	361
SP71-6163 NA vs. IACSP95-5000 NA	7	19	26
SP71-6163 AA vs. IACSP95-5000 AA	163	50	213
SP71-6163 VA vs. IACSP95-5000 VA	21	156	177

NA: no aphids (control); AA: aviruliferous aphids; VA: viruliferous aphids.

The infestation by aviruliferous and viruliferous aphids led mostly to the upregulation of genes in the SCYLV-susceptible cultivar SP71-6163. Sixty-three genes were upregulated and 7 were downregulated in response to aviruliferous aphids, while 23 were upregulated and 9 were downregulated in response to viruliferous aphids. In the tolerant cultivar IACSP95-

5000, infestation by aviruliferous aphids led mainly to gene upregulation, with 589 upregulated and 64 downregulated genes. However, infection by viruliferous aphids led to the downregulation of several genes, with 27 upregulated and 138 downregulated genes. These results were also clearly observed in the comparisons between the two cultivars, in which a low number of DEGs (26) were found in the control condition. Additionally, in both cultivars, the number of DEGs in the comparison between controls and plants with aviruliferous aphids (70 in SP71-6163 and 653 in IACSP95-5000) was much greater than that in the comparisons between controls and plants with viruliferous aphids (32 in SP71-6163 and 165 in IACSP95-5000).

A more detailed investigation of DEG annotations and the results of Gene Ontology (GO) enrichment analyses further demonstrated the differences in the responses of each cultivar to both *M. sorghi* and SCYLV. REVIGO TreeMaps showing the biological processes predominating in each DEG set are shown in Figures 3, S3 and S4. In the response of SP71-6163 to aviruliferous aphids (Figure S3), terms associated with ‘ent-kaurene oxidation to kaurenoic acid’ and ‘negative regulation of transcription by RNA polymerase II’ were enriched. IACSP95-5000 seemed to activate more defense-specific pathways in response to *M. sorghi*. In this cultivar, two processes with a clear association with plant resistance to aphids were highly represented: the metabolism of flavonoids and the regulation of Rac protein signal transduction (Figure S4). The representative genes for these processes are Cluster_17139, encoding a C-type lectin domain containing 16A (CLEC16A) protein, and Cluster_45858, encoding afadin- and alpha-actinin-binding protein B, both of which were upregulated in response to *M. sorghi*. When the two cultivars were compared, processes associated with ‘innate immune response’, ‘regulation of histone H3-K9 methylation’ and ‘signal peptide processing’ were found to be enriched (Figure 3). The genes Cluster_10338 and Cluster_5506, annotated as a phytosulfokine receptor 2 and an E3 ubiquitin-protein

ligase BRE1-like 2, respectively, were representative of the ‘innate immune response’ process and were upregulated in IACSP95-5000 infested with aviruliferous aphids.

Differences were also clear when analyzing the response of the two cultivars to viruliferous aphids. In SP71-6163, ‘methionine biosynthetic process’ and ‘methylation’ categories were enriched (Figure S3). In IACSP95-5000, processes associated with the biogenesis of ribosomes constituted a high proportion of enriched GO terms, along with others such as the biosynthesis of amylopectin (Figure S4). Ribosome biogenesis was apparently altered in IACSP95-5000 in response to SCYLV, with the downregulation of a 25S rRNA (cytosine-C(5))-methyltransferase gene (Cluster 21838), and amylopectin synthesis was also affected, with differential expression of a soluble starch synthase (Cluster_1012). These two processes were also clearly represented when the two cultivars were compared, together with ‘photosynthetic acclimation’ and ‘vacuolar transmembrane transport’ (Figure 3).

In the comparison of IACSP95-5000 infested with aviruliferous and viruliferous aphids, several processes that were present in other comparisons also stood out. GO terms associated with ‘flavonoid transport from endoplasmic reticulum to plant-type vacuole’, for instance, were present in the comparisons between plants with viruliferous aphids and both the control and plants with aviruliferous aphids, with several DEGs identified in within-genotype comparisons also present in these comparisons. However, genes linked to these terms, such as Cluster_17139, were upregulated in response to *M. sorghi* alone and downregulated in response to SCYLV-viruliferous aphids. Notably, genes associated with the regulation of Rac protein signal transduction were upregulated in response to aviruliferous aphids but were not differentially expressed between the control plants and plants infested with viruliferous aphids. Similarly, GO terms associated with the biogenesis of ribosomes and amylopectin biosynthesis were identified via comparisons of plants exposed to

viruliferous aphids with both the non-inoculated controls and plants exposed to aviruliferous aphids.

2.5. Gene coexpression network analyses

After additional filtering steps, transcript per million (TPM)-normalized data from 85,685 genes were used to construct a gene coexpression network via WGCNA. The pairwise Pearson correlation coefficients were increased to a soft threshold with $\beta = 7$, corresponding to an R^2 of 0.798 and a mean connectivity of 536 (Figure S5). Using the unweighted pair group method with arithmetic mean (UPGMA) approach, 47 functional modules were defined in this network, with sizes ranging from 125 to 14,816 genes and an average size of 2,112 genes (Table S8 and Figure S6). Additionally, the genes with the highest connectivity in each module were identified as hubs. Three genes identified as hubs in their modules were also identified as DEGs. These genes were Cluster_21034 in the darkorange module, annotated as an L-type lectin domain-containing receptor kinase; Cluster_140592 in the ivory module, annotated as histone H4; and Cluster_39399 in the violet module, annotated as autophagy-related protein 16-1.

Some DEG sets were overrepresented in eight network modules, with some highly significant associations (Table S9). Notably, the three modules whose hub genes were also differentially expressed—darkorange, ivory, and violet—were among them. The most significant associations involved the modules darkorange, plum1 and violet, associated with comparisons including IACSP95-5000 infested with aviruliferous aphids, and magenta, associated with comparisons that included IACSP95-5000 infested with viruliferous aphids. These modules were enriched for various biological processes, as depicted in Figure S7. In the darkorange module, ‘regulation of DNA stability’ and ‘lipoxygenase pathway’ were the most prevalent categories. In the plum1 module, ‘regulation of glucose metabolic process’

and ‘flavonoid transport from endoplasmic reticulum to plant-type vacuole’ were predominant. In the violet module, ‘regulation of muscle cell differentiation’ and ‘protein urmylation’ were the most prevalent categories. In the magenta module, ‘shade avoidance’ and ‘copper ion transport’ prevailed. Other interesting processes enriched in other modules were ‘regulation of peroxidase activity’ and ‘regulation of ncRNA transcription’ in the ivory module and ‘regulation by virus of viral protein levels in host cell’ in the midnightblue module.

3. Discussion

Herein, *M. sorghi* was shown to be a vector of SCYLV in sugarcane. Until recently, this aphid species was considered restricted to Africa, but this view changed after the invasion of sorghum by the multilocus lineage (MLL)-F in the United States in 2013 (Bowling et al. 2016). Initially identified as *M. sacchari* (Nibouche et al. 2018), MLL-F was later identified by Nibouche et al. (2021) as a lineage of *M. sorghi.* These authors also presented evidence that despite having a preference for sorghum, this species also infests sugarcane. Many previous studies that focus on *Melanaphis* spp. do not mention species identification methodologies and often reference *M. sacchari*, but others have supported the hypothesis that *M. sorghi* is responsible for recent aphid outbreaks in sorghum in the Americas (Harris-Shultz et al. 2022; Balbi et al. 2023). Given the threat this aphid has posed in recent years, the evidence that it is capable of transmitting SCYLV is worrisome, especially given the potential detrimental effects of SCYLV on the host response to *M. sorghi.*

The DGE analysis performed here indicated that SCYLV-susceptible SP71-6163 has an incipient response to both SCYLV and *M. sorghi*, with very few genes activated or repressed in response to this virus and its vector. Its response to aviruliferous aphids led to the upregulation of Cluster_3551, encoding a leucine repeat-rich (LRR) receptor-like kinase

that triggers plant innate immunity (Wang et al. 2021). However, the functions of most DEGs identified in comparisons involving only SP71-6163 were overall not consistent with what is commonly described for plant transcriptomic responses to sucking insects, which usually include secondary metabolic processes, jasmonic and salicylic acid biosynthesis, and signal transduction (Zogli et al. 2020), neither to viruses, such as RNA silencing and other pathways related to plant host defense and stress responses (Zanardo et al. 2019).

However, some of these processes were clearly and differentially represented in the responses of SCYLV-tolerant IACSP95-5000. In addition to the upregulation of the LRR kinase-encoding gene Cluster_3551, plants of this cultivar exposed to aviruliferous *M. sorghi* exhibited upregulation of genes associated with the biosynthesis of flavonoids, a class of chemically diverse secondary metabolites that are highly relevant in the defense of plants against insects (Treutter 2006). Genes involved in flavonoid biosynthesis have been previously shown to be upregulated in sorghum infested with *M. sacchari* (Kiani and Szczepaniec 2018; Tetreault et al. 2019). Here, genes encoding a CLEC16A protein (Cluster_17139), which has domains associated with intracellular membrane trafficking and the accumulation of flavonoids in plants (Ichino et al. 2014), and a flavonoid 3'-monooxygenase (Cluster_78445), which is involved in the biosynthesis of these compounds, were upregulated in response to *M. sorghi* in IACSP95-5000. Similarly, the signaling of Rac proteins, a class of Rho GTPases, may be enhanced in IACSP95-5000 through the upregulation of afadin- and alpha-actinin-binding proteins observed in plants exposed to *M. sorghi*. These proteins are also linked to the production of reactive oxygen species and thus to plant responses to biotic stresses (Torres 2010). Indeed, a Rac protein has been demonstrated to improve the defense responses of cotton plants against the aphid *Myzus persicae* (Yang et al. 2018).

IACSP95-5000 also responded very differently to infection by SCYLV. Notably, infection with this virus seemed to lead to the regulation of genes linked to the biogenesis of ribosomes and the biosynthesis of amylopectin. A representative upregulated gene associated with the biogenesis of ribosomes is an rRNA methyltransferase, which participates in rRNA methylation, a process that optimizes ribosomal function (Zou et al. 2020). In general, RNA viruses are known to induce heterogeneous ribosome production (Wang et al. 2022), and differential expression of ribosomal protein genes has been reported in studies analyzing plant responses to SCYLV and other plant viruses (Shabbir et al. 2022; Zhang et al. 2022).

The synthesis of amylopectin may also be affected by the modulation of the expression of starch synthase, an enzyme that plays a key role in the synthesis of this polysaccharide (Irshad et al. 2021) and has even stronger links with infection by SCYLV. Amylopectin is the major component of starch, composing ca. 80% of this polymeric carbohydrate in sugarcane (Figueira et al. 2011; Alves et al. 2014). The most notorious effects of infection by SCYLV in sugarcane relate to the metabolism and transport of carbohydrates, one of which is the increased content of starch in the leaves, especially in tolerant or resistant varieties (Lehrer et al. 2007, 2010; Yan et al. 2008). Thus, the differential regulation of genes affecting the synthesis of the main constituents of starch in response to SCYLV could be associated with its accumulation during infection.

One of the most notable results observed here was the fewer DEGs identified in both sugarcane cultivars exposed to viruliferous aphids than in those exposed to aviruliferous aphids. This is an indication that infection by SCYLV has a negative effect on the plant's response to its vector, an occurrence that has been observed in other members of the *Polerovirus* genus, including potato leafroll virus (PLRV) and turnip yellows virus (TuYV) (Patton et al. 2020; Krieger et al. 2023). Krieger et al. (2023) recently demonstrated that TuYV has a suppressive effect on the defense mechanisms of *Arabidopsis thaliana* against

M. persicae to increase its transmission efficiency. It is possible that a similar phenomenon occurs in the sugarcane-*M. sorghi*-SCYLV system, in which the virus seemingly suppresses the activation of pathways leading to the biosynthesis of compounds that would otherwise be upregulated in response to the aphid alone.

Additionally, a gene coexpression network revealed broader biological processes associated with sugarcane responses to *M. sorghi* and SCYLV. The darkorange module had an overrepresentation of DEGs from comparisons involving IACSP-95-5000 infested with aviruliferous aphids, among which was the top hub of this module. This gene encodes an L-type lectin domain-containing receptor kinase, a class of protein that has been shown to act as an immune receptor in plants and activate downstream defense pathways (Wang and Bouwmeester 2017). The recognition of attack by *M. sorghi* by such a kinase could lead to the activation of defense-related processes such as the lipoxygenase pathway, which was enriched in the darkorange module. Lipoxygenases are a family of enzymes with diverse biological roles, including the response to biotic and abiotic stresses. Some lipoxygenases are differentially expressed in sorghum in response to *M. sacchari* feeding (Shrestha et al. 2021), so their involvement in the response of sugarcane to *M. sorghi* is biologically plausible.

A similar scenario was observed for the violet module, which also had an overrepresentation of DEGs associated with IACSP-95-5000 infested with aviruliferous aphids, including the top hub of the module. This gene encodes an autophagy-related protein, a class of proteins that modulates cell death in plants under biotic stresses, including aphid feeding (Hao et al. 2023). Consistent with this finding, the violet module was enriched with several biological processes linked with autophagy, such as cell killing and protein catabolism, which may be enhanced in IACSP95-5000 in response to *M. sorghi*. Crucially, the hubs of the violet and darkorange modules were not differentially expressed in response

to viruliferous aphids, reinforcing the hypothesis that SCYLV disrupted the associated immune responses.

The plum1 module, which also showed high overrepresentation of DEGs identified in comparisons involving IACSP95-5000 infested with aviruliferous aphids, is also worthy of further discussion. In addition to the terms associated with flavonoid transport, whose role in response to aphids has been discussed, this module was enriched with terms linked with the metabolism of glucose. In sorghum, *M. sacchari* feeding is known to affect the content of saccharides and increase the content of glucose, which has also been correlated with increased damage caused by this aphid (Knoll et al. 2021; Cardona et al. 2023). Finally, the midnightblue module, which was associated with comparisons involving IACSP95-5000 infested with viruliferous aphids, was enriched with 'regulation by virus of viral protein levels in host cell' terms, which can be linked to the plant response to SCYLV. These results revealed further potential effects of *M. sorghi* feeding and infection with SCYLV that were not captured in the DGE analysis alone, demonstrating the ability of gene coexpression network methodologies to expand the scope of transcriptomic analyses.

The potential of SCYLV to affect host response to *M. sorghi* highlights the need for further investigations on the incidence of SCYLV in both this aphid and in the crops it infests, which could exacerbate the direct damage the aphid inflicts. Research on the percentage of *Melanaphis* spp. aphids carrying SCYLV has never been carried out on any crop. Studies analyzing aphids vectoring other poleroviruses have revealed incidences of up to 100% (Singh et al. 1997; Thango 2019). SCYLV is often highly prevalent in commercial sugarcane fields (Lehrer et al. 2008; Amata et al. 2016; Filloux et al. 2018) and could serve as an inoculum source for sorghum fields, where the incidence of this virus has thus far been shown to be relatively low (Boukari et al. 2021; Viswanathan et al. 2023). Recent research has shown that *M. sacchari* feeding on different sugarcane cultivars infected with SCYLV

present differences in their behavior and biology (Bertasello et al. 2021). Thus, it would also be valuable to evaluate the performance and behavior of *Melanaphis* spp. on SCYLV-infected and virus-free plants to assess how the biology of these aphids is affected by the virus, which is consistent with the findings reported herein.

4. Conclusions

This study revealed that *M. sorghi* is capable of vectoring SCYLV, posing an additional threat to the crops currently impacted by this aphid. Moreover, these findings indicate that infection by SCYLV has a negative effect on plant defense responses to its vector, leading to shifts in the expression of genes involved in pathways modulated in response to feeding by *M. sorghi*. Contrasting responses were observed in a comparison between SCYLV-susceptible and SCYLV-tolerant sugarcane genotypes, indicating that susceptibility may result from a combination of ineffective responses to both SCYLV and *M. sorghi*. This highlights the need for further investigations on the impacts of SCYLV on its vectors and hosts and corresponding implications for virus epidemiology.

5. Methods

5.1. Plant material and virus isolate

Two Brazilian hybrid sugarcane cultivars were used in the experiment. SP71-6163 is efficiently infected by SCYLV, develops severe symptoms of YLD and is strongly affected by the virus in terms of yield (Vega et al. 1997; Rodrigues 2015); therefore, it was chosen as an SCYLV-susceptible genotype. Despite being a highly suitable host to *M. sacchari* and exhibiting SCYLV infection rates similar to those of SP71-6163 (Rodrigues 2015; Bertasello et al. 2021), IACSP95-5000 has not been reported to exhibit YLD symptoms or high yield

losses (M.G.A. Landell, personal communication); for these reasons, this cultivar was selected as an SCYLV-tolerant genotype. Virus-free plants of the two cultivars were generated by meristem micropropagation at the Instituto Agronômico de Campinas Advanced Centre for Sugarcane Research and Development's biofactory.

Plants clonally propagated from an SCYLV-infected SP71-6163 plant from Ribeirão Preto, Brazil (Pimenta et al. 2021), were used as a source of inoculum. Total RNA was extracted from the leaf of one plant following a protocol adapted from Kistner and Matamoros (2005). RNA was quantified on a NanoDrop 2000 spectrophotometer and subjected to electrophoresis on a 1% denaturing agarose gel stained with ethidium bromide for an integrity check. RT was performed with the ImProm-II Reverse Transcription System Kit (Promega, Madison, USA). Phusion High-Fidelity PCR Master Mix (New England BioLabs, Ipswich, USA) was used to amplify overlapping fragments covering the complete SCYLV genome. Primers were designed using Primer3Plus v. 3.3.0 (Untergasser et al. 2007); details are provided in Table S9. PCR products were purified with polyethylene glycol and sequenced on a 3500xL Genetic Analyzer (Applied Biosystems, Foster City, USA).

Geneious v. 7.1.3 software (Kearse et al. 2012) was used for Sanger sequencing data handling. Low-quality bases and primer sequences were manually trimmed, and reads were used to produce a de novo assembly of a contiguous complete genomic sequence. The complete genomes of 101 SCYLV isolates from 17 different countries, as well as the PLRV reference genome (Table S10), were retrieved from GenBank, pairwise-aligned using ClustalW v. 2.1 (Thompson et al. 2003) and trimmed to the length of the shortest sequences. Pairwise Hamming distances (Hamming 1950) were calculated on UGENE v. 48.1 (Okonechnikov et al. 2012). The presence of recombinants was determined using the Recombination Detection Program v. 5.33 (RDP5) (Martin et al. 2021) using seven different methods with default parameters: RDP, GENECONV, BOOTSCAN, MAXCHI,

CHIMAERA, SISCAN and 3SEQ. Only recombination events detected by at least five methods and showing significant support with a Bonferroni-corrected *p* value cutoff of 0.05 were considered valid (Pagán and Holmes 2010). Recombinant isolates were removed for subsequent analyses, and MEGA11 v. 11.0.11 (Tamura et al. 2021) was used to find the best-fit DNA/protein model based on the Bayesian information criterion (BIC) score. The model with the lowest BIC score was used to construct a maximum likelihood tree rooted by the *Polerovirus* type species *Potato leafroll virus* (PLRV) genome via the same software. Bootstrap procedures with 1,000 replicates were used to generate standard error estimates, and the resulting tree was visualized with FigTree v. 1.4.4 218 software (Rambault 2018).

5.2. Aphid rearing and identification

Melanaphis sp. aphids were obtained from a sugarcane field in Ribeirão Preto, Brazil, and reared on detached sugarcane leaf fragments in test tubes with 1% agar kept in a plant growth chamber at 30°C under 12-hour light/dark cycles. A colony of aviruliferous aphids was obtained by transferring individual adult aphids onto virus-free sugarcane leaves and transferring newborn nymphs to separate virus-free leaves. To ensure the effectiveness of the process, nymph transfer was repeated for several aphid generations, and the colonies obtained were tested for the presence of SCYLV by RT□PCR. For this purpose, RNA was extracted from pools of five aphids using TRIzol (Invitrogen, Carlsbad, USA), quantified on a NanoDrop spectrophotometer and converted into cDNA using the ImProm-II Reverse Transcription System Kit. cDNA was used for PCR diagnosis as described by Gonçalves et al. (2002). Prior to inoculation experiments, a viruliferous colony was obtained by transferring aphids to SCYLV-infected SP71-6163 plants kept in aphid-proof cages. Aphids were allowed an acquisition access period on infected plants of at least 48 hours for virus acquisition.

While *M. sacchari* is the most common *Melanaphis* species that infests sugarcane, *M. sorghi* has been recently reported in Brazil (Nibouche et al. 2021; Harris-Shultz et al. 2022). Because the two species are very similar morphologically, a molecular diagnostic marker was used to determine the species of the established colony. DNA was extracted from a pool of five aphids using a protocol adapted from Sunnucks and Hales (1996). The primers MelF2 (GTGCTTATTGTCGCTGCTGG) and MelR2 (TTGTCTCCGGGAACAGCTTC) were designed to amplify a 683-bp fragment of the *EF1-α* gene harboring five polymorphic sites in *M. sacchari* and *M. sorghi*, one of which is sufficient to distinguish the two species (Nibouche et al. 2021). PCR was performed using the Phusion High-Fidelity PCR Master Mix and an annealing temperature of 60°C. The PCR products were purified using polyethylene glycol and sequenced on a 3500xL Genetic Analyzer. Geneious was used to manually trim low-quality bases and primer sequences, perform pairwise alignment of forward and reverse reads, and analyze the nucleotides present at the polymorphic sites.

5.3. Inoculation experiment and sample processing

The inoculation experiment was performed in a plant growth chamber kept at 30°C under 12-hour light/dark cycles; three biological replicates consisting of 11-month-old plants were used for each experimental condition. Clip cages were positioned on the median region of the top visible dewlap leaf of each plant, in which five viruliferous or aviruliferous adult aphids were placed; for aphid-free control plants, no aphids were used. After inoculation access periods (IAPs) of 4, 24 or 48 hours, aphids were removed, and whole leaves were collected, flash-frozen in liquid nitrogen and stored at -80°C until processing.

Total RNA was extracted from the ~10 cm leaf section around the clip cage placement site, quality was checked, and RT was performed as previously described. cDNA was used to confirm infection and estimate the SCYLV titer by a qPCR-based relative

quantification method detailed by Pimenta et al. (2021). The normality and variance of the viral quantification data were verified by Shapiro Wilk and F tests, respectively, and differences in quantification between genotypes were assessed by an unpaired t test in R software (R Core Team 2011).

RNA from selected samples was used to prepare RNA-Seq libraries using the TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, USA). After their fragment size profiles were verified on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, USA), the libraries were quantified on a Qubit 3.0 fluorometer (Thermo Scientific, Wilmington, USA), pooled and requantified. A 2×150 bp paired-end sequencing library was prepared with the NovaSeq 6000 SP Reagent Kit v1.5 (Illumina, San Diego, USA) and sequenced in two lanes of a NovaSeq 6000 instrument (Illumina, San Diego, USA).

5.4. Reference transcriptome assembly and annotation

In addition to the data generated for the samples from the present study, previously available RNA-Seq data were included in the analyses to contribute to the assembly of a representative transcriptomic reference and to the construction of a robust gene coexpression network. These consisted of data from leaf samples collected from twelve diverse *Saccharum* genotypes generated by Correr et al. (2020). The quality of all RNA-Seq data was assessed using FastQC (Andrews 2010). Trimmomatic v.0.39 (Bolger et al. 2014) was used for read trimming, using a 5-bp sliding window with a minimum average Phred quality score of 20 and removing reads shorter than 75 bp.

Filtered reads were aligned to the genomes of *S. officinarum* (GenBank GCA_020631735.1) and *S. spontaneum* (GenBank GCA_022457205.1) using STAR v. 2.7.3 (Dobin et al. 2013). These genomes are allele-resolved, and each haplotype was treated as a different reference. Thus, twelve different alignments were performed for each sample,

representing the eight haplotypes of *S. officinarum* and the four of *S. spontaneum*. After comparative mapping, the resulting files were sorted based on their genomic positions using SAMtools v. 1.12 (Li et al. 2009) and used to define twelve distinct transcriptomes via Stringtie v. 2.1.6 (Pertea et al. 2015). CD-HIT v. (Fu et al. 2012) was then used to reduce redundancies in the reference. This software was first used for the transcriptomes of each species to combine genes from individual haplotypes and subsequently to combine the transcriptomes generated for the two species. This final reference was evaluated with BUSCO v. 5.0.0 (Simão et al. 2015) using datasets of conserved orthologs from Eukaryota and Viridiplantae. It was then annotated with Trinotate v. 4.0.1 (Bryant et al. 2017) by performing a homology search of sequences on UniProt release 2024_1 (UniProt Consortium 2019).

5.5. Differential gene expression analyses

Salmon v.1.1.0 software (Patro et al. 2017) was used for transcript quantification via selective alignment of the sequencing reads and correction for fragment-level GC biases. The tximport R package v. 1.26.1 (Soneson et al. 2015) was used to import quantification data into R software. The edgeR package v 3.40.2 (Robinson et al. 2010) was used to normalize quantifications by CPM. For the differential gene expression analysis, in which only samples generated in the present work were used, specific filters were used. First, to remove genes with low expression across treatments, only genes with a sum of CPMs ≥ 3 in all samples of a treatment were retained.

Next, genes with inconsistent expression across replicates of each treatment were filtered out. To achieve this goal, paired t tests were performed to identify genes with significant ($p < 0.05$) differences in CPM values across the six lanes encompassing each treatment, which were discarded. Subsequently, the gene sets for each biological condition

were combined according to the nine comparisons performed: (1) SP71-6163 Control vs. SP71-6163 Aviruliferous Aphids; (2) SP71-6163 Control vs. SP71-6163 Viruliferous Aphids; (3) SP71-6163 Aviruliferous Aphids vs. SP71-6163 Viruliferous Aphids; (4) IACSP95-5000 Control vs. IACSP95-5000 Aviruliferous Aphids; (5) IACSP95-5000 Control vs. IACSP95-5000 Viruliferous Aphids; (6) IACSP95-5000 Aviruliferous Aphids vs. IACSP95-5000 Viruliferous Aphids; (7) SP71-6163 Control vs. IACSP95-5000 Control; (8) SP71-6163 Aviruliferous Aphids vs. IACSP95-5000 Aviruliferous Aphids; and (9) SP71-6163 Viruliferous Aphids vs. IACSP95-5000 Viruliferous Aphids.

DGE analyses were performed with the R package DESeq2 v. 1.38.3 (Love et al. 2014) using a false discovery rate (FDR)-corrected *p* value of 0.05 and a $\log_2$(FC) value of 1 as thresholds for the delimitation of DEGs. The UpSetR R package (Conway et al. 2017) was used to investigate overlaps between sets of DEGs and generate an UpSet plot. GO enrichment analysis was performed on each DEG set with the topGO R package v. 2.50.0 (Alexa et al. 2006) using Fisher's test with $\alpha = 0.05$. REVIGO v. 1.8.1 (Supek et al. 2011) was used for the visualization and analysis of the enriched GO categories.

5.6. Gene coexpression network

A gene coexpression network was constructed using data generated in the present study as well as those from Correr et al. (2020). Quantification data were filtered to retain genes with a CPM ≥ 1 in all samples from at least one treatment. The edgeR R package was used to normalize quantifications by TPM, and the WGCNA R package v. 1.72.5 (Langfelder and Horvath 2008) was used to construct the gene coexpression network. Pairwise Pearson correlations of TPM values considering a power function fitting scale-free independence were used. A soft threshold value corresponding to an R^2 value of approximately 0.8 was used to approximate a scale-free topology model. Functional modules in the network were

defined using UPGMA based on a topological overlap matrix and dynamic dendrogram pruning based only on the dendrogram. Hubs, defined as the genes with the highest connectivity in each module, were identified with the dedicated WGCNA function.

Fisher's exact test with an FDR correction with $\alpha = 0.05$, performed in R, was used to verify whether DEG sets identified in the DGE analyses were overrepresented in any of the modules of the network. GO enrichment analysis was performed on the genes included in these modules with topGO using Fisher's test with $\alpha = 0.05$, and enriched GO terms were visualized with REVIGO.

5.7. RT□qPCR validation

RNA-Seq libraries were validated by RT□qPCR. Primers to amplify eight DEGs and one gene stably and highly expressed across all samples were designed using Primer3Plus with parameters optimized for qPCR. Additionally, primers designed to amplify one common sugarcane control gene, *GAPDH* (Iskandar et al. 2004), were synthesized. The sequences of all the qPCR primers can be found in Table S11. cDNA was synthesized using the QuantiTect Reverse Transcription Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. qPCRs were performed using the iTaq Universal SYBR Green Supermix (Bio-Rad, Philadelphia, USA) on a Bio-Rad CFX384 Touch detection system (Bio-Rad, Philadelphia, USA) according to the manufacturer's cycling recommendations. The mean Cq values of the two endogenous controls were used to correct the DEG Cq values. All reactions were performed using three technical replicates. Linear models correlating the CPM and ΔCq values were fitted and plotted in R using the ggplot2 (Wickham 2016) and ggpubr (Kassambara 2023) packages.

List of abbreviations

IAP: Inoculation access period

BIC: Bayesian information criterion

CLEC16A: C-type lectin domain containing 16A

CPM: Counts per million

DGE: Differential gene expression

DEG: Differentially expressed gene

EF1-α: Elongation factor 1-alpha

LRR: Leucine repeat-rich

GAPDH: Glyceraldehyde-3-phosphate dehydrogenase

GO: Gene Ontology

MLL: Multilocus lineage

PLRV: Potato leafroll virus

RT□qPCR: Reverse transcription followed by quantitative polymerase chain reaction

SCYLV: Sugarcane yellow leaf virus

TPM: Transcripts per million

TuYV: Turnip yellows virus

UPGMA: Unweighted pair group method with arithmetic mean

YLD: Yellow leaf disease

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The RNA sequencing data generated by the present work have been deposited in SRA under the BioProject ID PRJNA1125051. The complete genome of the BRA-SP1 isolate and the *M. sorghi EF1-α* partial sequence have been deposited in GenBank under accession numbers PP796534 and PP796535, respectively. Table S6 is available on FigShare under the DOI 10.6084/m9.figshare.27236814.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grants from the São Paulo Research Foundation (FAPESP), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant 424050/2016-1), and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Computational Biology Program). RJGP received MSc and PhD fellowships from FAPESP (grants 2018/18588-8 and 2019/21682-9). AHA received a PhD fellowship from FAPESP (grant 2019/03232-6).

Authors' contributions

MCG, LRP and APS conceived the project and designed the experiments. MNGM generated the virus-free plants. RJGP and DAS performed further experimental procedures. RJGP and AHA analyzed the data. RJGP interpreted the results and wrote the manuscript. All the authors have read and approved the manuscript.

Acknowledgements

We thank Silvana Creste for the use of the IAC biofactory for the production of virus-free sugarcane plants, Maicon Volpin for assistance in aphid rearing and greenhouse work and Aline C. L. Moraes for assistance in constructing and sequencing the RNA-Seq library.

References

Abu Ahmad Y, Rassaby L, Royer M, Borg Z, Braithwaite KS, Mirkov TE, et al. Yellow leaf of sugarcane is caused by at least three different genotypes of sugarcane yellow leaf virus, one of which predominates on the Island of Réunion. Arch Virol. 2006; 151:1355-71.

Akbar S, Yao W, Yu K, Qin L, Ruan M, Powell CA, et al. Photosynthetic characterization and expression profiles of sugarcane infected by Sugarcane mosaic virus (SCMV). Photosynth Res. 2020; 150:279-94.

Akbar W, Showler AT, Beuzelin JM, Reagan TE, Gravois KA. Evaluation of aphid resistance among sugarcane cultivars in Louisiana. Ann Entomol Soc Am. 2011; 104:699-704.

Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics. 2006; 22:1600-7.

Alves FV, Polesi LF, Aguiar CL, Sarmento SBS. Structural and physicochemical characteristics of starch from sugar cane and sweet sorghum stalks. Carbohydr Polym. 2014; 111:592-7.

Amata RL, Fernandez E, Filloux D, Martin DP, Rott P, Roumagnac P. Prevalence of sugarcane yellow leaf virus in sugarcane-producing regions in Kenya revealed by reverse-transcription loop-mediated isothermal amplification method. Plant Dis. 2016; 100:260-8.

Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. www.bioinformatics.babraham.ac.uk/projects/fastqc.

Balbi EI, Decker-Franco C, Szwarc D, Casuso VM, Saluso A, Arneodo JD. Unveiling the occurrence of *Melanaphis sorghi* in Argentina following a major aphid outbreak. Phytoparasitica. 2023; 51:13-8.

Bertasello LET, Carmo-Sousa M, Maluta NKP, Pinto LR, Lopes JRS, Gonçalves MC. Effect of sugarcane cultivars infected with sugarcane yellow leaf virus (ScYLV) on feeding behavior and biological performance of *Melanaphis sacchari* (Hemiptera: Aphididae). Plants. 2021; 10:2122.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30:2114-20.

Boukari W, Mollov D, Wei C, Tang L, Grinstead S, Tahir MN, et al. Screening for sugarcane yellow leaf virus in sorghum in Florida revealed its occurrence in mixed infections with sugarcane mosaic virus and a new marafivirus. Crop Prot. 2021; 139:105373.

Bowling RD, Brewer MJ, Kerns DL, Gordy J, Seiter N, Elliott NE, et al. Sugarcane aphid (Hemiptera: Aphididae): a new pest on sorghum in North America. J Integr Pest Manag. 2016; 7:12.

Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, Payzin-Dogru D, et al. A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. Cell Rep. 2017; 18:762-76.

Cardona JB, Grover S, Bowman MJ, Busta L, Kundu P, Koch KG, et al. Sugars and cuticular waxes impact sugarcane aphid (*Melanaphis sacchari*) colonization on different developmental stages of sorghum. Plant Sci. 2023; 330:111646.

Chinnaraja C, Viswanathan R, Karuppaiah R, Bagyalakshmi K, Malathi P, Parameswari B. Complete genome characterization of sugarcane yellow leaf virus from India: evidence for RNA recombination. Eur J Plant Pathol. 2013; 135:335-49.

Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics. 2017; 33:2938-40.

Correr FH, Hosaka GK, Gómez SGP, Cia MC, Vitorello CBM, Camargo LEA, et al. Time-series expression profiling of sugarcane leaves infected with *Puccinia kuehnii* reveals an ineffective defense system leading to susceptibility. Plant Cell Rep. 2020; 39:873-89.

Debibakas S, Rocher S, Garsmeur O, Toubi L, Roques D, D'Hont A, et al. Prospecting sugarcane resistance to sugarcane yellow leaf virus by genome-wide association. Theor Appl Genet. 2014; 127:1719-32.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-Seq aligner. Bioinformatics. 2013; 29:15-21.

ElSayed AI, Boulila M, Odero DC, Komor E. Phylogenetic and recombination analysis of sorghum isolates of sugarcane yellow leaf virus. Plant Pathol. 2018; 67:221-32.

ElSayed AI, Komor E, Boulila M, Viswanathan R, Odero DC. Biology and management of sugarcane yellow leaf virus: an historical overview. Arch Virol. 2015; 160:2921-34.

ElSayed AI, Weig AR, Komor E. Molecular characterization of Hawaiian sugarcane yellow leaf virus genotypes and their phylogenetic relationship to strains from other sugarcane-growing countries. Eur J Plant Pathol. 2011; 129:399-412.

Fartek B, Nibouche S, Atiama□Nurbel T, Reynaud B, Costet L. Genotypic variability of sugarcane resistance to the aphid *Melanaphis sacchari*, vector of the sugarcane yellow leaf virus. Plant Breed. 2014; 133:771-6.

Figueira JDA, Carvalho PH, Sato HH. Sugarcane starch: quantitative determination and characterization. Food Sci Tech Brazil. 2011; 31:806-15.

Filloux D, Fernandez E, Comstock JC, Mollov D, Roumagnac P, Rott P. Viral metagenomic-based screening of sugarcane from florida reveals occurrence of six sugarcane-infecting viruses and high prevalence of sugarcane yellow leaf virus. Plant Dis. 2018; 102:2317-23.

Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012; 28:3150-2.

Gao S-J, Lin Y-H, Pan Y-B, Damaj MB, Wang Q-N, Mirkov TE, et al. Molecular characterization and phylogenetic analysis of sugarcane yellow leaf virus isolates from China. Virus Genes. 2012; 45:340-9.

Gaur R, Rao GP, Lehrer A. Detection of sugarcane yellow leaf luteovirus of India. Nat Preced. 2009; doi:10.1038/npre.2009.3388.1.

Gonçalves MC, Vega J, Oliveira JG, Gomes MMA. Sugarcane yellow leaf virus infection leads to alterations in photosynthetic efficiency and carbohydrate accumulation in sugarcane leaves. Fitopatol Bras. 2005; 30:10-6.

Gonçalves MC, Klerks MM, Verbeek M, Vega J, van den Heuvel JFJM. The use of molecular beacons combined with NASBA for the sensitive detection of sugarcane yellow leaf virus. Eur J Plant Pathol. 2002; 108:401-7.

Hall DG. The sugarcane aphid, *Melanaphis sacchari* (Zehntner), in Florida sugarcane. J Am Soc Sugar Cane Technol. 1987; 7:26-9.

Hamming RW. Error detecting and error correcting codes. Bell Syst Tech J. 1950; 29:147-60.

Harris-Shultz K, Armstrong JS, Carvalho G, Jr., Segundo JP, Ni X. *Melanaphis sorghi* (Hemiptera: Aphididae) clonal diversity in the United States and Brazil. Insects. 2022; 13:416.

Hao LK, Dangol A, Shavit R, Pitt WJ, Nalam V, Brotman Y, et al. Self-eating while being eaten: Elucidating the relationship between aphid feeding and the plant autophagy machinery in *Arabidopsis* leaves. bioRxiv. 2023.

Ichino T, Fuji K, Ueda H, Takahashi H, Koumoto Y, Takagi J, et al. GFS 9/TT 9 contributes to intracellular membrane trafficking and flavonoid accumulation in *Arabidopsis thaliana*. Plant J. 2014; 80:410-23.

Irshad A, Guo H, Rehman SU, Wang X, Wang C, Raza A, et al. Soluble starch synthase enzymes in cereals: an updated review. Agronomy. 2021; 11:1983.

Iskandar HM, Simpson RS, Casu RE, Bonnett GD, Maclean DJ, Manners JM. Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression in sugarcane. Plant Mol Biol Report. 2004; 22:325-37.

ISO. International sugar organization. London: ISO; 2023.

Kassambara A. ggpubr: 'ggplot2' based publication ready plots. R package version 0.6.0. 2023.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012; 28:1647-9.

Keese P, Martin RR, Kawchuk LM, Waterhouse PM, Gerlach WL. Nucleotide sequences of an Australian and a Canadian isolate of potato leafroll luteovirus and their relationships with two European isolates. J Gen Virol. 1990; 71:719-24.

Kiani M, Szczepaniec A. Effects of sugarcane aphid herbivory on transcriptional responses of resistant and susceptible sorghum. BMC Genom. 2018; 19:774.

Kistner C, Matamoros M. RNA isolation using phase extraction and LiCl precipitation. In: Márquez AJ, editor. *Lotus japonicus* handbook. Dordrecht: Springer; 2005. p. 123-4.

Knoll JE, Uchimiya M, Harris-Shultz K. Juice chemical properties of 24 sorghum cultivars under varying levels of sugarcane aphid (*Melanaphis sacchari*) infestation. Arthropod Plant Interact. 2021; 15:707-19.

Komor E, ElSayed A, Lehrer AT. Sugarcane yellow leaf virus introduction and spread in Hawaiian sugarcane industry: retrospective epidemiological study of an unnoticed, mostly asymptomatic plant disease. Eur J Plant Pathol. 2010; 127:207-17.

Krieger C, Halter D, Baltenweck R, Cognat V, Boissinot S, Maia-Grondard A, et al. An aphid-transmitted virus reduces the host plant response to its vector to promote its transmission. Phytopathology. 2023; 113:1745-60.

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinform. 2008; 9:559.

Lehrer A, Yan S-L, Fontaniella B, ElSayed A, Komor E. Carbohydrate composition of sugarcane cultivars that are resistant or susceptible to sugarcane yellow leaf virus. J Gen Plant Pathol. 2010; 76:62-8.

Lehrer AT, Kusalwong A, Komor E. High incidence of sugarcane yellow leaf virus (SCYLV) in sugar plantations and germplasm collections in Thailand. Australas Plant Dis Notes. 2008; 3:89-92.

Lehrer AT, Moore PH, Komor E. Impact of sugarcane yellow leaf virus (ScYLV) on the carbohydrate status of sugarcane: comparison of virus-free plants with symptomatic and asymptomatic virus-infected plants. Physiol Mol Plant Pathol. 2007; 70:180-8.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009; 25:2078-9.

Lin Y-H, Gao S-J, Damaj MB, Fu H-Y, Chen R-K, Mirkov TE. Genome characterization of sugarcane yellow leaf virus from China reveals a novel recombinant genotype. Arch Virol. 2014; 159:1421-9.

Ling H, Huang N, Wu Q, Su Y, Peng Q, Ahmed W, et al. Transcriptional insights into the sugarcane-sorghum mosaic virus interaction. Trop Plant Biol. 2018; 11:163-76.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. 2014; 15:550.

Lu J-J, He E-Q, Bao W-Q, Chen J-S, Sun S-R, Gao S-J. Comparative genomics reveals insights into genetic variability and molecular evolution among sugarcane yellow leaf virus populations. Sci rep. 2021; 11:7149.

Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T, et al. RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evol. 2021; 7:veaa087.

Moonan F, Molina J, Mirkov TE. Sugarcane yellow leaf virus: an emerging virus that has evolved by recombination between luteoviral and poleroviral ancestors. Virology. 2000; 269:156-71.

Nibouche S, Costet L, Holt JR, Jacobson A, Pekarcik A, Sadeyen J, et al. Invasion of sorghum in the Americas by a new sugarcane aphid (*Melanaphis sacchari*) superclone. PLoS One. 2018; 13:e0196124.

Nibouche S, Costet L, Medina RF, Holt JR, Sadeyen J, Zoogones A-S, et al. Morphometric and molecular discrimination of the sugarcane aphid, *Melanaphis sacchari*, (Zehntner, 1897) and the sorghum aphid *Melanaphis sorghi* (Theobald, 1904). PLoS One. 2021; 16:e0241881.

Okonechnikov K, Golosova O, Fursov M. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics. 2012; 28:1166-7.

Pagán I, Holmes EC. Long-term evolution of the *Luteoviridae*: time scale and mode of virus speciation. J Virol. 2010; 84:6177-87.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017; 14:417-9.

Patton MF, Bak A, Sayre JM, Heck ML, Casteel CL. A polerovirus, potato leafroll virus, alters plant–vector interactions using three viral proteins. Plant Cell Environ. 2020; 43:387-99.

Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-Seq reads. Nat Biotechnol. 2015; 33:290-5.

Pimenta RJG, Aono AH, Burbano RCV, Coutinho AE, da Silva CC, Dos Anjos IA, et al. Genome-wide approaches for the identification of markers and genes associated with sugarcane yellow leaf virus resistance. Sci rep. 2021; 11:15730.

R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011.

Rambault A. FigTree v1.4.4, a graphical viewer of phylogenetic trees. 2018. tree.bio.ed.ac.uk/software/figtree.

Rassaby L, Girard JC, Letourmy P, Chaume J, Irey MS, Lockhart BEL, et al. Impact of sugarcane yellow leaf virus on sugarcane yield and juice quality in Réunion Island. Eur J Plant Pathol. 2003; 109:459-66.

Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26:139-40.

Rodrigues MP. Evaluation of sugarcane cultivars resistance to sugarcane yellow leaf virus (SCYLV) and to its aphid vector *Melanaphis sacchari*. Master's thesis. Jaboticabal, Brazil: São Paulo State University; 2015.

Rott P, Grinstead S, Dallot S, Foster ZSL, Daugrois JH, Fernandez E, et al. Genetic diversity, evolution, and diagnosis of sugarcane yellow leaf virus from 19 sugarcane-producing locations worldwide. Plant Dis. 2023; 107:3437-47.

Scagliusi SM, Lockhart BEL. Transmission, characterization, and serology of a luteovirus associated with yellow leaf syndrome of sugarcane. Phytopathology. 2000; 90:120-4.

Shabbir R, Zhaoli L, Yueyu X, Zihao S, Pinghua C. Transcriptome analysis of sugarcane response to sugarcane yellow leaf virus infection transmitted by the vector *Melanaphis sacchari*. Front Plant Sci. 2022; 13:921674.

Shrestha K, Pant S, Huang Y. Genome-wide identification and classification of Lipoxygenase gene family and their roles in sorghum-aphid interaction. Plant Mol Biol. 2021; 105:527-41.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015; 31:3210-2.

Singh BU, Padmaja PG, Seetharama N. Biology and management of the sugarcane aphid, *Melanaphis sacchari* (Zehntner) (Homoptera: Aphididae), in sorghum: a review. Crop Prot. 2004; 23:739-55.

Singh RP, Kurz J, Boiteau G, Moore LM. Potato leafroll virus detection by RT-PCR in field-collected aphids. Am Potato J. 1997; 74:305-13.

Smith GR, Borg Z, Lockhart BEL, Braithwaite KS, Gibbs MJ. Sugarcane yellow leaf virus: a novel member of the *Luteoviridae* that probably arose by inter-species recombination. Microbiology. 2000; 81:1865-9.

Sõmera M, Fargette D, Hébrard E, Sarmiento C, ICTV Report Consortium. ICTV virus taxonomy profile: *Solemoviridae* 2021. J Gen Virol. 2021; 102:001707.

Soneson C, Love MI, Robinson MD. Differential analyses for RNA-Seq: transcript-level estimates improve gene-level inferences. F1000Res. 2015; 4:1521.

Sunnucks P, Hales DF. Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Mol Biol Evol. 1996; 13:510-24.

Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011; 6:e21800.

Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021; 38:3022-7.

Tetreault HM, Grover S, Scully ED, Gries T, Palmer NA, Sarath G, et al. Global responses of resistant and susceptible sorghum (*Sorghum bicolor*) to sugarcane aphid (*Melanaphis sacchari*). Front Plant Sci. 2019; 10:426693.

Thango S. An investigation of the identity, transmission efficiency and control of TuYV-vectoring aphid species in South Africa. Master of Science, MSc. South Africa: University of Johannesburg; 2019.

Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. Curr Protoc Bioinform. 2003; 2:2-3.

Torres MA. ROS in biotic interactions. Physiol Plant. 2010; 138:414-29.

Treutter D. Significance of flavonoids in plant resistance: a review. Environ Chem Lett. 2006; 4:147-57.

UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019; 47:D506-15.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 2007; 35:W71-4.

Vasconcelos ACM, Gonçalves MC, Pinto LR, Landell MG, Perecin D. Effects of sugarcane yellow leaf virus on sugarcane yield and root system development. Functional Plant Sci Biotechnol. 2009; 3:31-5.

Vega J, Scagliusi SMM, Ulian EC. Sugarcane yellow leaf disease in brazil: evidence of association with a luteovirus. Plant Dis. 1997; 81:21-6.

Viswanathan R, Chinnaraja C, Malathi P, Gomathi R, Rakkiyappan P, Neelamathi D, et al. Impact of sugarcane yellow leaf virus (ScYLV) infection on physiological efficiency and growth parameters of sugarcane under tropical climatic conditions in India. Acta Physiol Plant. 2014; 36:1805-22.

Viswanathan R, Nithya K, Vishnuvardhan J, Balasaravanan S, Kaverinathan K, Visalatchi D. Sorghum (*Sorghum bicolor*) a new host to sugarcane yellow leaf and mosaic viruses in India. Indian Phytopathol. 2023; 76:867-77.

Wang MQ, Zhou GH. A near-complete genome sequence of a distinct isolate of sugarcane yellow leaf virus from China, representing a sixth new genotype. Virus Genes. 2010; 41:268-72.

Wang X, Zhang N, Zhang L, He Y, Cai C, Zhou J, et al. Perception of the pathogen□induced peptide RGF7 by the receptor□like kinases RGI4 and RGI5 triggers innate immunity in *Arabidopsis thaliana*. New Phytol. 2021; 230:1110-25.

Wang X, Zhu J, Zhang D, Liu G. Ribosomal control in RNA virus-infected cells. Front Microbiol. 2022; 13:1026887.

Wang Y, Bouwmeester K. L-type lectin receptor kinases: new forces in plant immunity. PLoS Pathog. 2017; 13:e1006433.

Wei C, Hincapie M, Larsen N, Nuessly G, Rott P. First report of sugarcane yellow leaf virus infecting grain sorghum (*Sorghum bicolor*) in the United States. Plant Dis. 2016; 100:1798.

White WH, Reagan TE, Hall DG. *Melanaphis sacchari* (Homoptera: Aphididae), a sugarcane pest new to Louisiana. Fla Entomol. 2001; 84:435.

Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.

Yan SL, Lehrer AT, Hajirezaei MR, Springer A, Komor E. Modulation of carbohydrate metabolism and chloroplast structure in sugarcane leaves which were infected by sugarcane yellow leaf virus (SCYLV). Physiol Mol Plant Pathol. 2008; 73:78-87.

Yang Z, Liu J, Luo L, Ye S, Yang Y, Zhang G, et al. The cotton *GhRac6* gene encoding a plant ROP/RAC protein improves the plant defense response to aphid feeding. Plant Mol Biol Report. 2018; 36:888-96.

Zanardo LG, de Souza GB, Alves MS. Transcriptomics of plant–virus interactions: a review. Theor Exp Plant Physiol. 2019; 31:103-25.

Zhang J, Wang T, Jia Z, Jia X, Liu Y, Xuan J, et al. Transcriptome analysis reveals a comprehensive virus resistance response mechanism in pecan infected by a novel badnavirus pecan virus. Int J Mol Sci. 2022; 23:13576.

Zogli P, Pingault L, Grover S, Louis J. Ento(o)mics: the intersection of ‘omic’ approaches to decipher plant defense against sap-sucking insect pests. Curr Opin Plant Biol. 2020; 56:153-61.

Zou M, Mu Y, Chai X, Ouyang M, Yu L-J, Zhang L, et al. The critical function of the plastid rRNA methyltransferase, CMAL, in ribosome biogenesis and plant development. Nucleic Acids Res. 2020; 48:3195-210.

Figures

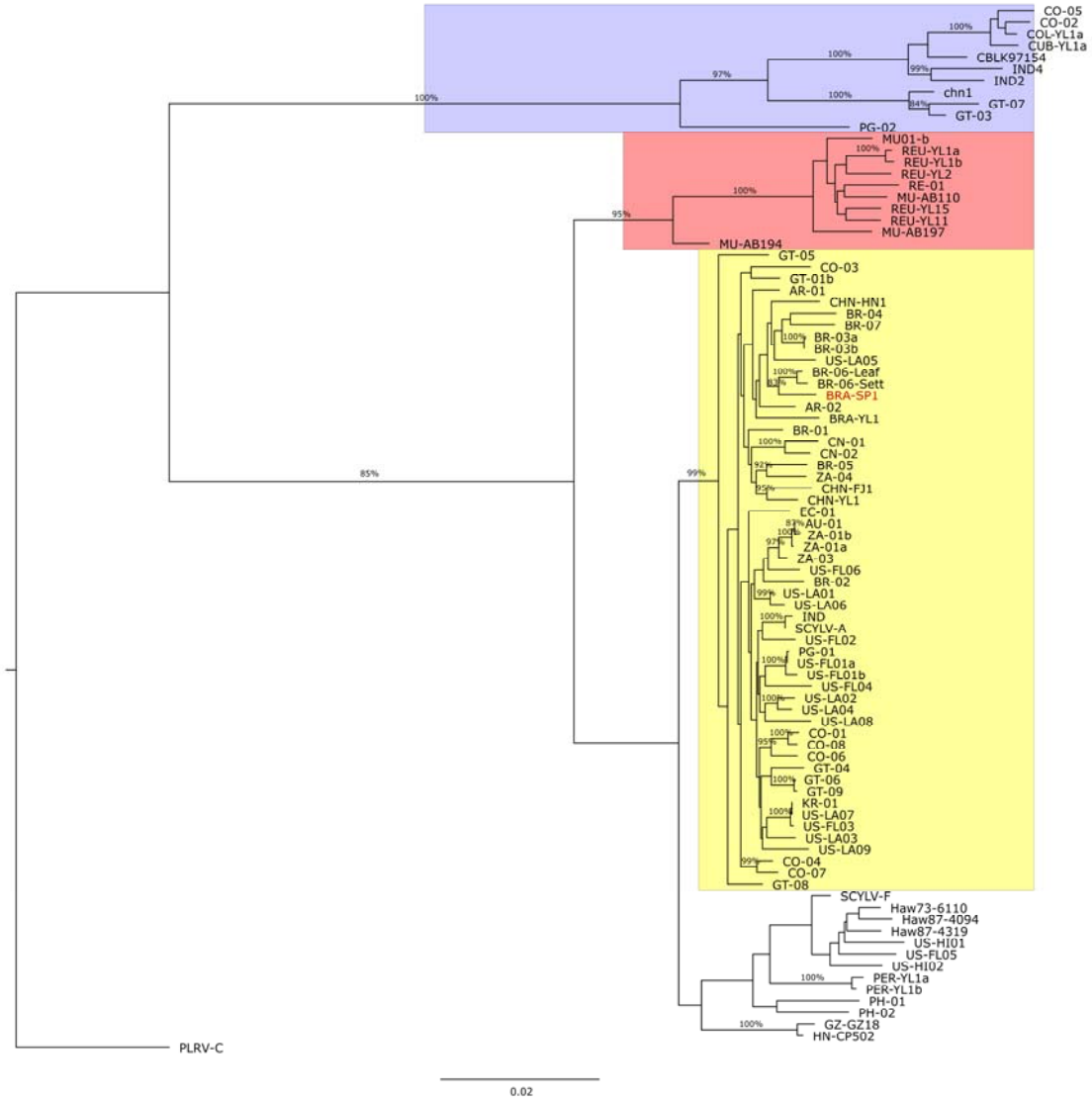

Figure 1. Maximum likelihood phylogenetic tree constructed using complete genome sequences of 89 sugarcane yellow leaf virus isolates. Numbers on the branches represent bootstrap values, for which a cutoff of 80% was used. The scale bar signifies a genetic distance of 0.02 nucleotide substitutions per site. Three major clades were identified: CUB (blue), REU (red) and BRA (yellow).

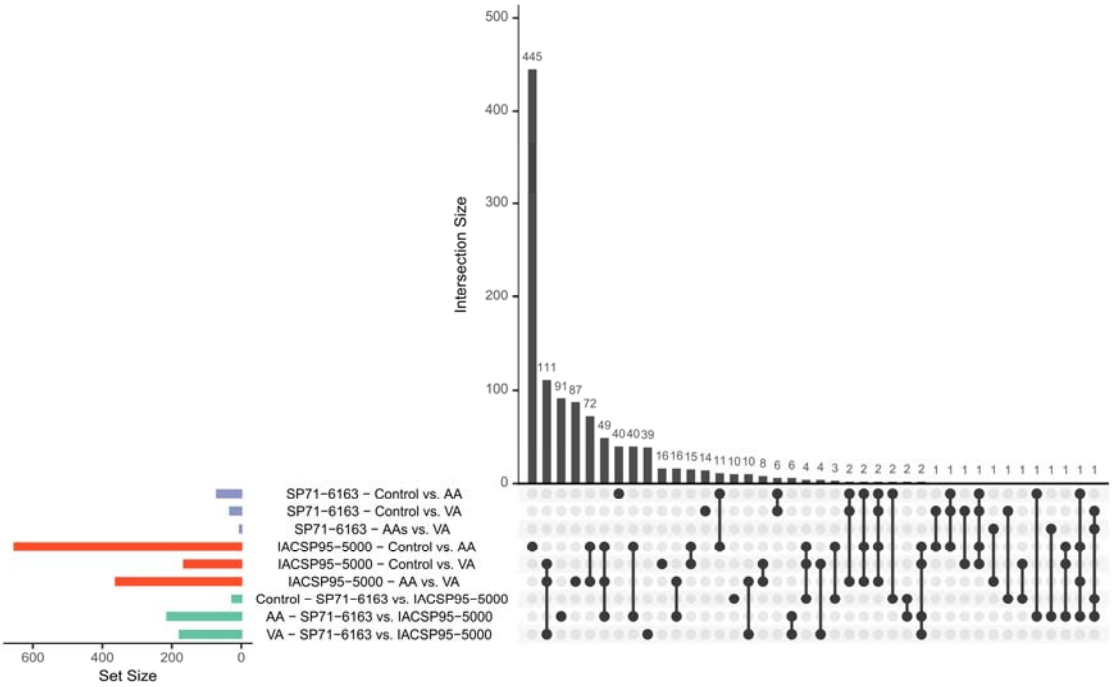

Figure 2. Upset plot displaying intersections between differentially expressed gene sets in SP71-6163 and IACSP95-5000.

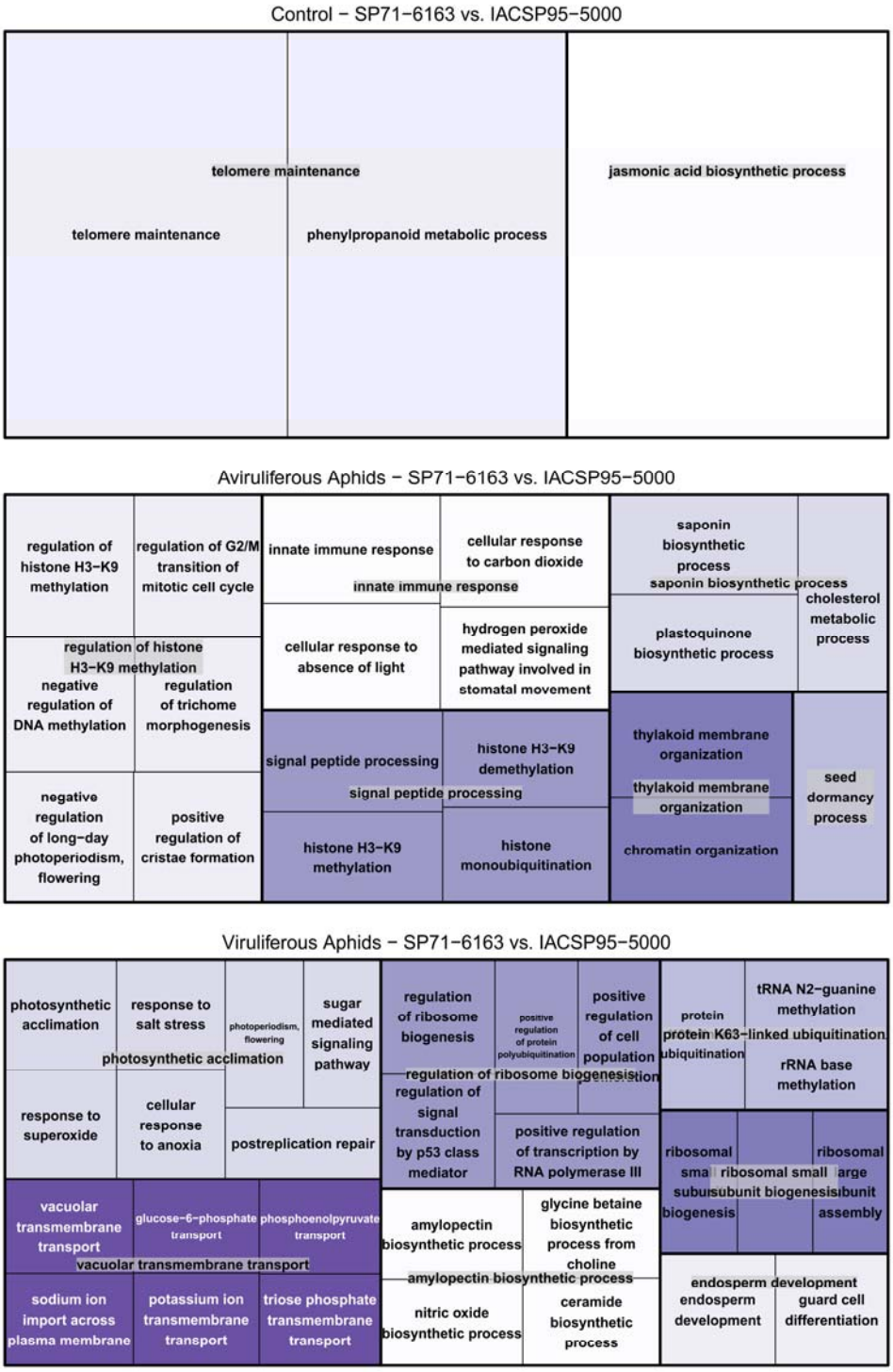

Figure 3. REVIGO representation of Gene Ontology categories enriched in the sets of differentially expressed genes from comparisons between SP71-6163 and IACSP95-5000 under the same treatments.