High-throughput sequencing - based detection and characterization of sugarcane viruses in Ethiopia

Mereme Abide[1*] • Dawit Kidanemariam[2] • Misrak Kebede[1] • Dennis Knierim[3] • Paolo Margaria[3] • Wulf Menzel[3] and Adane Abraham[1, 4*]

[1] Department of Biotechnology, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia

[2] Department of Plant and Soil Sciences, Forestry and Agricultural Biotechnology Institute, University of Pretoria, South Africa

[3]Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures GmbH, Inhoffenstrasse 7B, Braunschweig, Germany

[4] Department of Biological Sciences and Biotechnology, Botswana International University of Science and Technology, Private Bag 16, Palapye, Botswana

*Corresponding authors: mereme.abide@aastu.edu.et and abrahama@biust.ac.bw

Abstract

A virus-discovery investigation using high-throughput sequencing in sugarcane samples from Ethiopia revealed the presence of sugarcane mild mosaic virus (SCMMV, genus *Ampelovirus*) and sugarcane umbra-like virus (SULV, genus *Umbravirus*). The genome sequences of two isolates of SCMMV and one SULV were determined. The sequences of the two SCMMV isolates were 13005 nucleotides in length and showed ~73.5% nucleotide identity along the genome and ~90.2, 96.8, and 90.4% amino acid sequence identity among each other in the RdRp, CP, and HSP70h, respectively. Isolate SCMMV ET2 showed a close relationship to group A isolates from Colombia, the USA, and the Philippines, with amino acid sequence identity of the predicted virus proteins in the range of 94–98.9%. Conversely, SCMMV ET1 shared a closer relationship with group B isolates from Colombia, Ivory Coast, and Argentina, exhibiting a 93–99% amino acid sequence identity. The complete genome sequence of SULV, comprising 3041 nucleotide, exhibited the highest identity with its counterpart from South Africa (MN868593). These findings contribute to expanding our understanding of the viral diversity within the sugarcane crop in Ethiopia.

Keywords: Diversity, SCMMV, SULV, virus

Statements and Declarations:

The authors declare that they have no conflict of interest.

Acknowledgment: The authors would like to thank Addis Ababa Science and Technology University for financial support and the Molecular Biology Laboratory of the Biotechnology Department for lab facilities. The contribution of Viola Weichelt (Leibniz Institute DSMZ) in library preparation is deeply acknowledged.

Funding: This study was financially supported by the Addis Ababa Science and Technology University. Illumina high-throughput sequencing was performed as part of the EVA-Global project which has received funding from the European Union's Horizon 2020 research and innovation program (grant agreement no. 871029)

1 Introduction

Sugarcane (*Saccharum officinarum* L.) is one of the most important crops globally, primarily cultivated for its high sugar content and used as a valuable source of sugar and bioenergy in the form of ethanol (Formann et al., 2020; Meena et al., 2022). Despite its importance, sugarcane cultivation is constantly threatened by various viral diseases, leading to substantial yield losses and economic damage. Viral diseases in sugarcane are caused by a diverse group of viruses with either DNA or RNA genomes (Addy et al., 2017; Wang et al., 2020; Asinari et al., 2021). One of the most economically significant viruses is the sugarcane mosaic virus (SCMV), which belongs to the *Potyviridae* family. SCMV causes mosaic symptoms on leaves, affecting plant growth and resulting in reduced stalk size and yield losses (Zhang et al., 2015; Lu et al., 2021). Another common virus is the sugarcane yellow leaf virus (SCYLV), a member of the *Solemoviridae* family. SCYLV induces yellowing of leaves, stunting, and reduced sucrose content in sugarcane plants (Singh et al., 2011; Madugula et al., 2017; Nithya et al., 2023).

Other notable sugarcane viruses include the sugarcane bacilliform virus (SCBV) from the *Caulimoviridae* family, which causes a range of symptoms including stunting, leaf curling, and reduced yield. Additional viruses such as sugarcane mild mosaic virus (SCMMV), sugarcane streak mosaic virus (SCSMV), sugarcane fiji disease virus (FDV), sugarcane chlorotic streak virus (SCSV), maize yellow mosaic virus (MaYMV), maize striate mosaic virus (MSMV), sugarcane white streak virus (SWSV), sugarcane striate virus (SStrV), and an sugarcane umbra-like virus (SULV) have also been reported in different countries (Zhang et al., 2018; Sun et al., 2019; Batista et al., 2021; Lin et al., 2021;Lu et al., 2021; Mollov et al., 2021; Nithya et al., 2021). These viruses are primarily transmitted through vegetative propagation, aphids, mealybugs, and other insect vectors (Sastry, 2013; Shabbir et al., 2022; Abide et al., 2023). Once infected, sugarcane plants may exhibit a range of symptoms, including mosaic patterns on leaves, stunting, yellowing, necrosis, and reduced stalk growth. The severity of symptoms varies based on the virus species/strain(s), sugarcane variety, environmental conditions, and the stage of plant growth.

The impact of sugarcane viruses on crop production is significant, with yield losses ranging from 10% to 80% (Holkar et al., 2020; Lu et al., 2021). In addition to reduced yield, infected plants often have poor stalk quality, and reducing their overall vigor, making them more susceptible to other pathogens (Putra et al., 2014; Viswanathan et al., 2014; Boukari et al., 2019). Diagnostic tools, such as ELISA, PCR/RT-PCR, and high-throughput sequencing (HTS), are essential for the screening of plant material and the early detection and identification of viral infections. Efforts are being made to develop resistant varieties, improve detection and diagnostic methods, and implement integrated pest management strategies in the field to mitigate the impact of these viral diseases (Adams et al., 2018; Kumar et al., 2019; Malapi-Wight et al., 2021; Strachan et al., 2022; Wang et al., 2022; Viswanathan, 2023).

Previous studies have described the occurrence of SCBV in sugarcane in Ethiopia (Haregu et al., 2021; Abide et al., 2023) along with its management by the use of virus-free plants generated through tissue culture (Abide et al., 2022). The present study aimed to investigate virus infections in sugarcane samples from Ethiopia using high-throughput sequencing. The nearly complete sequences of two SCMMVs and the complete genome of SULV were reconstructed and are reported for the first time in Ethiopia.

2 Materials and methods

2.1 Samples collection

Fifteen sugarcane stem cuttings showing symptoms resembling viral diseases were collected from the Wonji-Shewa research farm during a field survey. The cuttings were planted in a sterilized soil mixture with a 1:1:1 ratio of soil, sand, and farmyard manure in plastic pots and maintained at the Addis Ababa Science and Technology University (AASTU) greenhouse. After one year of growth in greenhouse conditions, two symptomatic samples were selected, and leaf samples were preserved using silica gel and sent to Leibniz-Institute DSMZ (Braunschweig, Germany) for Illumina high-throughput sequencing.

2.2 RNA extractions, HTS, RT-PCR, and genome assembly

Total RNA was extracted from the two samples using a Spectrum® Plant Total RNA kit (Sigma), according to the manufacturer´s instructions. Two independent libraries were prepared using the Illumina DNA Prep Kit, following exactly to the manufacturer´s recommendations, and sequenced on a NextSeq2000 instrument as paired-end reads (2X150 bp). Due to capacity reasons, the individual RNA extracts from sugarcane were mixed at a ratio of 1:1 with yam samples from an independent study. The bioinformatic analysis was conducted in Geneious Prime version 2023.2.1 software (Biomatters, Auckland, New Zealand) using a custom workflow developed at the Plant Virus Department-DSMZ. In brief, the raw reads were paired, trimmed, normalized, and mapped against host sequences [chromosome sequences from NCBI BioProject PRJNA744175 and PRJNA695139; chloroplast sequences (NC_006084, NC_035224, NC_024170, NC_039707, NC_039836, NC_039851, NC_052854), mitochondrion sequence (NC_031164)]. Un-mapped reads were then used for *de novo* assembly and reconstruction of contigs, which were subsequently screened by BLASTn and BLASTp against a custom reference database for virus discovery, taxonomic assignment, and genome reconstruction.

The nearly complete genome sequences of the two SCMMV isolates were determined from Illumina reads. As the two SCMMV isolates were present in mixed infection, the assembly and genomes reconstruction were conducted using stringent parameters in Geneious Prime (number of allowed maximum mismatches per read was set to 0 in the final sequence assembly and curation), to avoid the possible generation of *in silico* hybrids. The complete genome sequence of SULV was assembled from both the Illumina reads and Sanger sequences obtained from the sequencing of RACE amplification products. For this purpose, the RNA was polyadenylated using a poly(A) polymerase (Poly(A) Polymerase, Yeast, Thermo Fisher Scientific), and the cDNA was synthesized using a poly(T) primer and a WarmStart RTx Reverse Transcriptase (NEB), according to the manufacturer´s instructions. Two specific primers at a distance of about 500 nucleotide (nt) from the expected 3'-end (Umbra-F1 5'-CACAAGAACCTACAACGGCTT-3' and Umbra-F2 5'-ACCTGTTCAAGA AACTGCGAA C-3'), and two specific primers binding about 300 nt after the expected 5'-end (Umbra-R1 5'-CCCAACCAACTTCCGCATAGT-3' and Umbra-R2 5'-TTTGCACT TCCGAGTACCCAC-3') were designed using Geneious Prime software using the "Design New Primers" tool based on the assembled sequences. The following PCR reactions were conducted with the Q5 Hot Start High-Fidelity 2X Master Mix (NEB) kit, following the manufacturer´s instructions. The generated PCR fragments were directly sequenced by Sanger sequencing at Microsynth Seqlab GmbH (Göttingen, Germany). To verify the assignment of the identified genomes to the individual samples, specific primers were derived based on the newly determined genome sequences (Table S2). The RT-PCR was carried out using the SuperScript One-Step RT-PCR kit with Platinum Taq (Invitrogen) according to the manufacturer´s instructions. For confirmation, the obtained PCR fragments were also directly sequenced by Sanger sequencing.

2.3 Sequence analysis

The ClustalW algorithm, integrated into the MEGA11 software, was used to conduct multiple sequence alignments of each sequence, along with related viral sequences retrieved from GenBank. The aligned sequences were then used to construct a phylogenetic tree using MEGA-11 software with the maximum likelihood method and 1000 bootstrap replications, employing the Kimura-2 parameter model (Tamura et al., 2021). The nearly complete genome and the most conserved sequence regions (heat shock protein 70 (HSP70), RNA dependent RNA polymerase (RdRp), and capsid protein (CP)) of the SCMMV ET1 and ET2 isolates, along with 17 SCMMV isolates reported from different countries, were used for this analysis. Additionally, thirty-two viruses belonging to the family *Tombusviridae*, consisting of umbra-like viruses and umbra viruses retrieved from GenBank, were used for the phylogenetic analysis of SULV. The aligned genome sequences of the two SCMMVs and a SULV were used to detect the possible occurrence of recombination events using Recombination Detection Program version 4.101 (RDP) software. Pairwise sequence comparison was performed using SDT version 1.2. SnapGene viewer version 1.5.3 software was used to determine the ORFs and predict the molecular weights of the encoded proteins.

3 Results

3.1 Virome investigation by HTS

The bioinformatic analysis allowed the assembly of contigs assigned to SCMMV in sample 1 (further analyzed in detail below) and SULV in sample 2 (Table S1). The presence of identified genomes in the symptomatic samples was ascertained by RT-PCR (Figure S2) using specific primers (Table S2) followed by direct sequencing. Contigs yielding hits to badnaviruses in the BLAST analysis were also found, but were not further considered in this work. The reconstruction of virus genomes resulted in the nearly complete genome sequence of two SCMMV isolates from sample 1 and the complete genome sequence of SULV in sample 2. The assembled sequences of SCMMV and SULV were 13005 nt and 3041 nt, respectively, and have been deposited in GenBank (OR906130, OR906131, and OR906133).

3.2 Virus genome organization

The finalized complete genomes of SCMMV consist of seven open reading frames (ORF1a, ORF1a/b, ORF2, ORF3, ORF4, ORF5, and ORF6). ORF1a encodes for a polyprotein that is 1986 amino acid (aa) long and has a molecular weight of 224.5 kDa. It has two conserved domains, namely, viral methyltransferase (pfam01660) and viral helicase (pfam01143), at 492–799 and 1687–1953 aa, respectively. ORF1a/b encodes a protein of 2510 aa with a predicted molecular weight of 284 kDa, assigned to the RdRp superfamily. ORF2 encodes a 54 aa protein with a molecular weight of 5.8 kDa. ORF3 codes for a protein of 533 aa with a molecular weight of 58.6 kDa, assigned to HSP70. ORF4 codes for a protein of 547 aa with a molecular weight of 62.8 kDa, assigned to heat shock protein 90 (HSP90). ORF 5 encodes the CP and consists of 287 aa for SCMMV ET1 and 282 aa for SCMMV ET2. The molecular weights of these proteins are estimated to be 30.3 kDa and 29.8 kDa, respectively. The 3′ terminal ORF6 encodes a protein of 218 aa with a molecular weight of 23.8kDa that has similarity with p24 protein.

The SULV genome ends were determined by 5′ and 3′ RACE amplification, resulting in a 3041-nt full-length sequence with four open reading frames (ORF1-4). ORF1 encodes a protein of 200 aa with a molecular weight of 22.3 kDa. The slippery heptameric sequence (GGGTTTT) was identified at position 589 – 595 at the end of ORF1, where the -1 ribosomal frameshift is predicted. ORF1-ORF2 encodes a protein of 727 aa with a molecular weight of 82.7 kDa and the conserved catalytic core domain of RdRp. ORF3 encodes a protein of 117 with a molecular weight of 13 kDa. The final ORF (ORF4) encodes a protein of 206 aa with a molecular weight of 23.1 kDa.

3.3 Phylogenetic and genome recombination analyses

The phylogenetic analysis revealed that the two SCMMV isolates in sample 1 were grouped into two groups (A and B). The first cluster included SCMMV ET2 from the current study, as well as SCMMV isolates from Colombia, the USA, and the Philippines that were previously classified as group A (Mollov et al., 2023). The SCMMV ET1 isolate clustered with group B SCMMV isolates from Colombia, Argentina, and Ivory Coast (Figure 1). A similar grouping pattern was also observed when the tree was constructed using the HPS70, RdRp, and CP conserved sequence regions (data not shown).

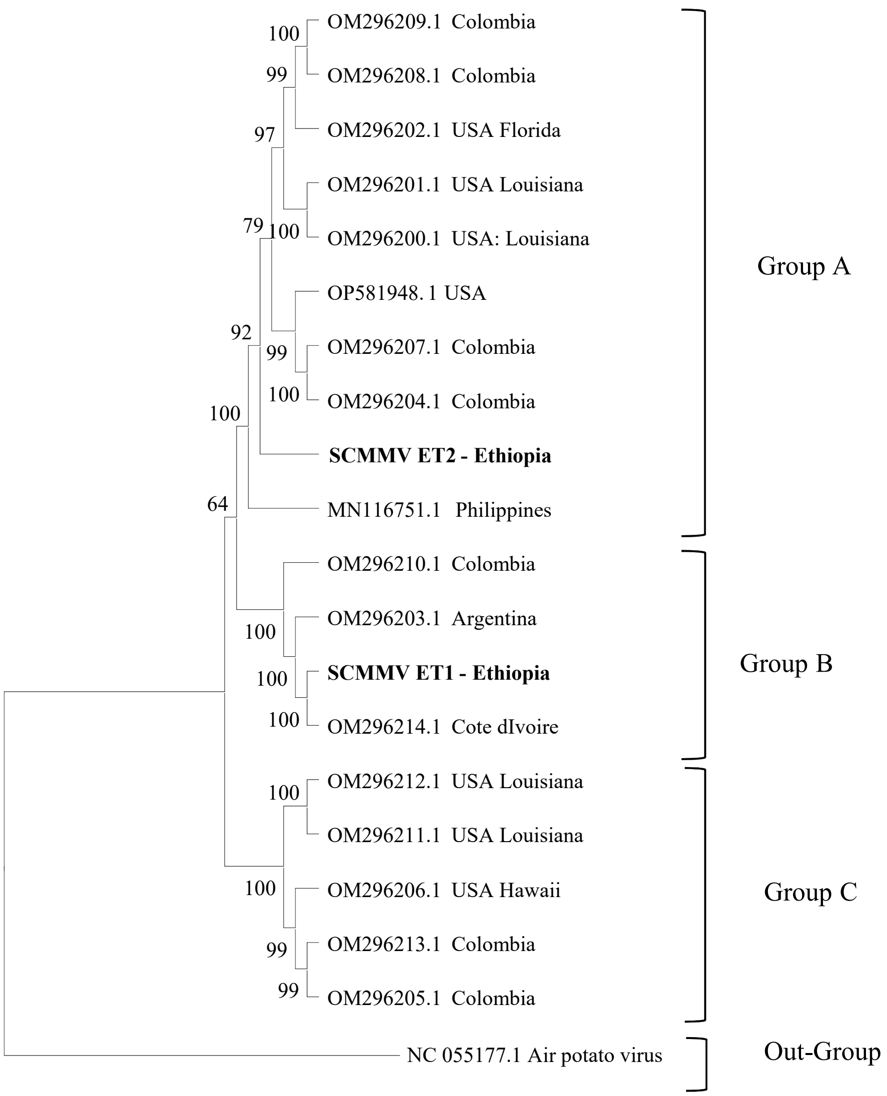

Figure 1. Phylogenetic tree of SCMMV generated by the Maximum Likelihood method in MEGA-11 based on the nt genome sequences of 17 selected SCMMV isolates from GenBank. Air potato virus was used as an out-group. Isolates from this study are highlighted in bold.

The phylogenetic tree for SULV was constructed using the RdRp region from the current study along with the RdRp sequence of thirty-two viruses belonging to the family *Tombusviridae*. The analysis revealed that the SULV from the current study clustered together with sugarcane umbra-like virus (MN868593), Ethiopia maize-associated virus (MN715238), johnsongrass umbra-like virus (OM937760), fig umbra-like virus (MW480893), opuntia umbra-like virus (MH579715.1), and citrus yellow vein-associated virus (JX101610.1) (Figure 2). The RDP4 analysis was also conducted to assess the potential presence of recombination events in the three sequences (2 SCMMV and 1 SULV). However, no recombination events were identified by the RDP4 analysis for any of the sequences.

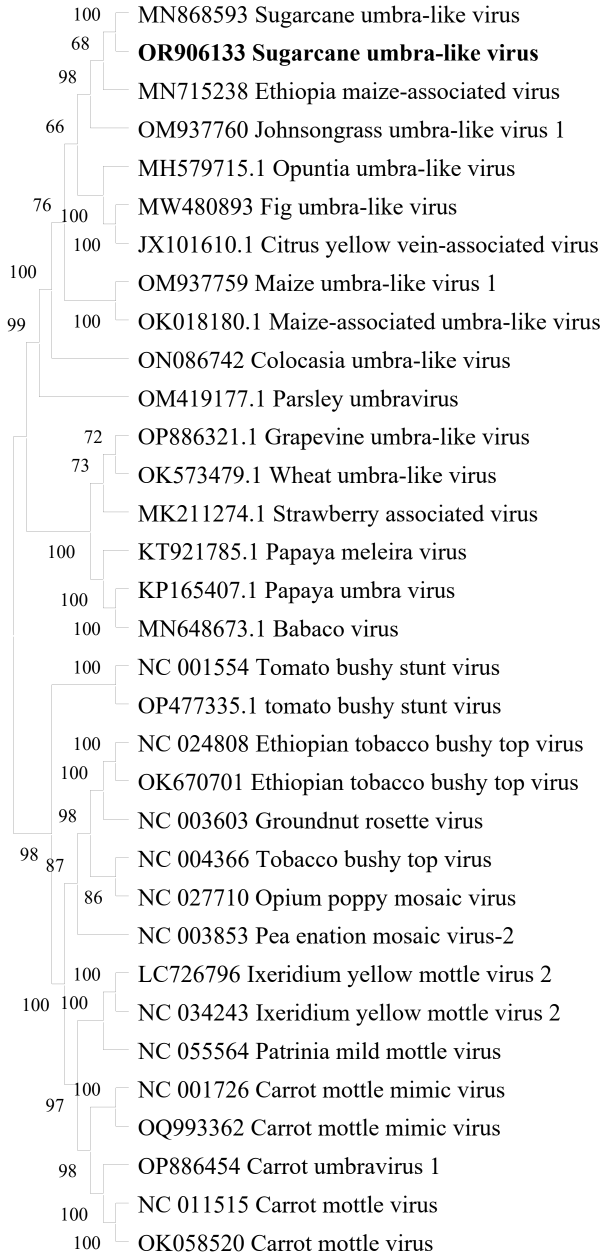

Figure 2. Phylogenetic tree of SULV generated by the Maximum Likelihood method in MEGA-11 based on aa RdRp sequences (from the ORF1/ORF2 frameshift to the stop codon) of SULV with umbra-like viruses and umbraviruses retrieved from GenBank. The sequence from this study is highlighted in bold.

3.4 Pairwise sequence analysis

Pairwise sequence analysis was conducted to compare the sequence identity of the SCMMV and SULV genomes with reference sequences. The analysis revealed that the two SCMMV isolates from this study exhibited a nt sequence identity of ~73.5% along the reconstructed genome sequence. These isolates demonstrated nt sequence identities ranging from 78% to 98.8%, with SCMMV isolates previously categorized into group A, B, and C (Mollov et al., 2023).

SCMMV ET2 showed the highest aa identity (94–98.9%) with group A SCMMV isolates identified from Colombia, the USA, and the Philippines. This isolate also showed 82% aa sequence identity with group B and C SCMMV isolates identified from Cote d'Ivoire, Argentina, the USA, and Colombia (Figure 3). The SCMMV ET1 isolate in this study exhibited the highest aa sequence identity (93–99%) with group B SCMMV isolates from Colombia, Ivory Coast, and Argentina. Conversely, this isolate shared a 78-82% identity with group A and C SCMMV identified from the USA and Colombia.

Pairwise identity analyses were further conducted for each SCMMV protein (polyprotein, RdRp, P6, HSP70, SHP90, CP, and P24). SCMMV ET1 displayed protein sequence identities ranging from 73–100% with SCMMVs from the three groups. ORF1a of SCMMV ET1 showed the highest sequence identity (89.6-98%) with group B SCMMV isolates, while shared 73% -75% protein sequence identity with group A and C isolates. Protein sequence identity of SCMMV ET1 ORF1b (RdRp) ranged from 89 to 99.8%, with the highest similarity to group B SCMMV isolates. The HSP70 protein (ORF3) shared 90–99.2% sequence identity with the three SCMMV groups. The protein sequence identities of HSP90 (ORF4) and the coat protein (ORF5) were 88–99.6% and 95–99.6%, respectively.

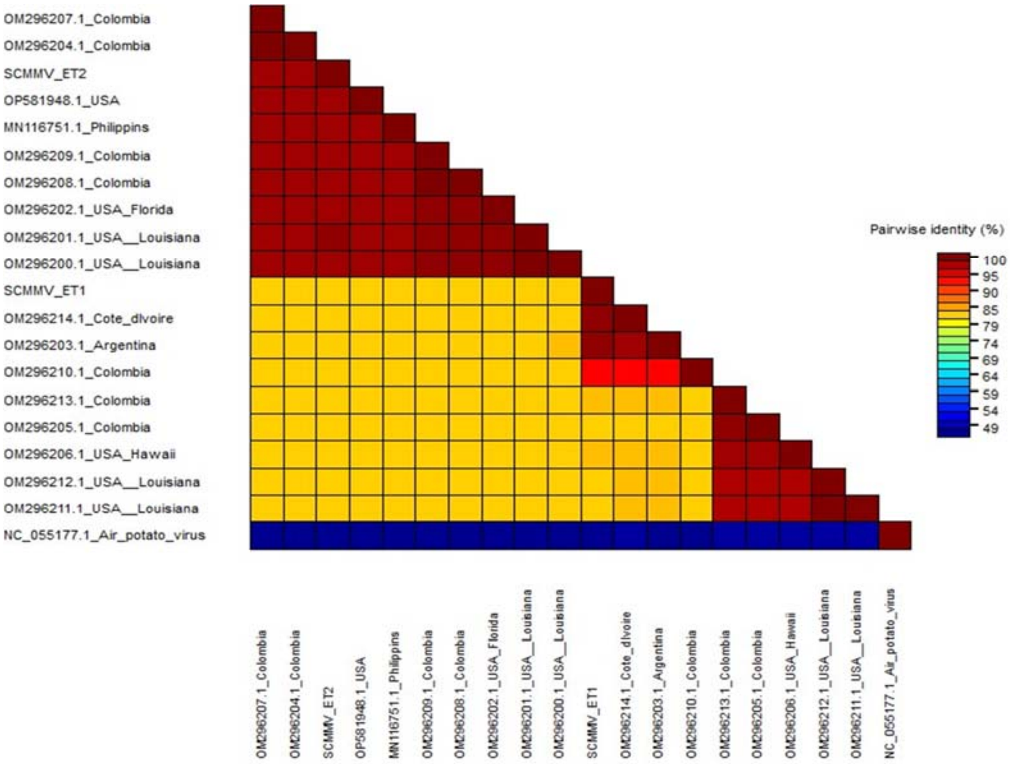

Figure 3. Pairwise nucleotide identity matrix based on the near complete genome sequence of SCMMV isolates from Ethiopia together with isolates from NCBI.

On the other hand, the SCMMV ET2 ORFs (ORF1a–ORF6) showed the highest protein sequence identity (97.7–100%) with group A SCMMV isolates and 73–97.8% sequence identities with group B and C SCMMV isolates. The comparison of the seven ORFs in SCMMV ET2 with group A SCMMV isolates exhibited sequence identities ranged from 97–98.2% for the ORF1a protein, 99-99.6% for the RdRp, 98–100% for the P6, 99% for the HSP70, 98.7–99.7% for the HSP90, 98.6%–100% for the CP, and 98-100% for P24 protein. Besides, ORF1a protein, RdRp, P6, HSP70, HSP90, CP, and P24 shared 73–97.5% with group B and C SCMMV isolates. Out of the seven ORFs, ORF 5 (CP) is the most conserved protein. It showed the highest sequence identity among the three SCMMV groups in both SCMMV ET1 and SCMMV ET2 isolates.

Pairwise sequence comparison showed that the SULV isolate investigated in this study exhibited sequence identities ranging from ~38% to ~86% with unclassified umbra viruses (umbra-like viruses) and umbra viruses. The highest aa sequence identity was observed with sugarcane umbra-like virus (85.8%) and Ethiopia maize-associated virus (74.1%).

4 Discussion

Recent investigations into viruses in sugarcane in Ethiopia have shown the widespread occurrence of SCBV (Haregu et al., 2022; Abide et al., 2023). In this study, HTS revealed the presence of two viruses in sugarcane that had not been previously reported in Ethiopia. In addition to sequences assigned to the previously identified SCBV, the nearly complete genome of two SCMMV (genus *Ampelovirus*) isolates and the complete genome of an unclassified sugarcane umbra-like virus (family *Tombusviridae*) have been reconstructed. The two isolates of SCMMV were present in a mixed infection within a single plant. Due to their genome-level diversity and the parameters used in the bioinformatics analysis for genome reconstruction, the two sequences could be confidently assembled as distinct isolates. Studying whole genome sequences offers a more profound insight for detailed analysis of genetic structure and molecular epidemiology.

The initial report of SCMMV dates back to 1992 (Lockhart et al. 1992); however, it did not receive recognition from the *Closteroviridae* ICTV study group. Currently this virus, known to be associated with mosaic disease in sugarcane, is classified within the *Ampelovirus* genus (Rott et al., 2018; He et al., 2022). The virus has been found to have a fairly wide geographical distribution, including Africa, Australia, the Caribbean, North America, South America, and Southeast Asia (Fernandez et al., 2020; He et al., 2022; Mollov et al., 2023). The current study reports its first occurrence in Ethiopia. Although SCMMV is known to occur in four African countries namely Malawi, Mauritius, South Africa, and Ivory Coast, information on its genomic sequence is reported only for one isolate from the latter (Mollov et al., 2023). The nearly complete genome of two isolates of SCMMV obtained in this work from Ethiopia, representing East Africa, will expand our knowledge of genome structure and diversity in the continent. Previous research has shown that SCMMV is primarily transmitted by the pink sugarcane mealybug, *Saccharicoccus sacchari* (Lockhart et al., 1992), which is confirmed to be the vector for SCBV, in the country (Abide et al., 2023). This insect, which commonly colonizes sugarcane crops in the country, is thus likely to play an important role in the field spread of both SCBV and SCMMV, possibly also resulting in a mixed infection. The presence of SCMMV in Ethiopia raises concerns about its potential impact on sugarcane production in the country. The detection of this virus could have significant implications for the agricultural sector. Further research is needed to understand the spread and impact of SCMMV in Ethiopia and to develop effective management strategies to mitigate its effects on sugarcane production.

SCMMV ET2 possesses a 282 aa coat protein, while SCMMV ET1 has a predicted longer coat protein (287 aa). The aa sequence of the CP showed a high degree of conservation among all groups of SCMMV, which may indicate the presence of a strong evolutionary pressure to retain these sequences. Phylogenetic analysis of the complete and partial genome sequences of SCMMV ET1 and SCMMV ET2 isolates revealed the distinct clustering of these isolates with group B and A SCMMVs, respectively as described by Mollov et al. (2023). The co-existence of these two isolates with a low degree of identity in a single sample suggests the possibility of multiple strains of the virus within the same environment. These distinct clades may have arisen due to different selection pressures. It is also likely that ET1 and ET2 isolates have been introduced to Ethiopia on multiple occasions, perhaps from different countries with sugarcane cuttings which are the major means of trade and germplasm exchange across the world.

A virus with characteristics similar to umbraviruses has been discovered in various plant species, including sugarcane, edible fig (Ficus carica), maize, johnsongrass, papaya, babaco, and strawberry (Quito-Avila et al., 2015; Cornejo-Franco et al., 2021; Daugrois et al., 2021; Tahir et al., 2021; Wang, 2021; Koloniuk et al., 2022; Quito-Avila et al., 2022). Previously, only maize-associated viruses have been reported in Ethiopia (Adams et al., 2017). The current study reports the first identification of this virus infecting sugarcane in Ethiopia. Based on complete-genome BLAST searches, the isolate from this study showed 85.8% and 74.1% sequence identities to a sugarcane umbra-like virus found in South Africa (MN868593.1) (Tahir et al., 2021) and Ethiopia maize-associated virus (MN715238) detected in maize affected by maize lethal necrosis disease (Adams et al., 2017), respectively.

Phylogenetic tree constructed from RdRp gene (Fig. 2) which is generally used for evolutionary classification of members of the family *Tombusviridae* placed our SULV isolate in a clade that consists of all umbra-like viruses described from various crops in recent years.

Although this clade shares the most recent common ancestor with members of genus *Umbravirus*, studies showed that the members of this group have different genome features, such as mostly smaller genome size, varying number of genes and encoded proteins, and regulatory elements such as -1 frameshift positions. In addition, unlike umbraviruses, no helper viruses are identified for most umbra-like viruses raising questions as to whether they use coat proteins of other viruses for encapsulation or vector transmission. Considering these features distinguishing them from umbraviruses and their evolutionary relationship, Tahir et al. (2021) proposed the classification of umbra-like viruses in separate genus in the family *Tombusviridae.* A more recent study (Ying et al. 2024) indicates that umbra-like viruses are unique among plant viruses in not encoding movement protein and relying on host movement protein for systemic spread within a plant. The results of our phylogenetic analysis add more evidence in support of the proposal for the creation of a new genus consisting of umbra-like viruses such as SULV in the family. The *Tombusviridae* working group of ICTV should also come up with species demarcation criteria within the proposed new genus, as the one currently used for umbraviruses may not work for this distinct group of viruses.

In conclusion, the present study provides the first report of the occurrence of SCMMV and SULV on sugarcane in Ethiopia, their complete genome sequences, and insight into their evolutionary relationship to known viruses. The study also offers valuable information on the diversity of viruses within the sugarcane crop in Ethiopia. Further research is needed on the prevalence, diversity and impact of these viruses on sugarcane production, as well as their mode of transmission, including dependence on helper viruses if any. In addition, studies on the possible cross-infection of these viruses, particularly with economically important crop hosts such as maize and sorghum, will provide basic information that can be used for devising suitable disease management options.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Mereme Abide, Adane Abraham, Dawit Kidanemariam, and Misrak Kebede. Illumina high-throughput sequencing was performed at Leibniz Institute-DSMZ. The first draft of the manuscript was written by Mereme Abide, and all the authors commented on the previous version of the manuscript. All authors read and approved the final manuscript.

Data availability

The sequences mentioned in this paper can be accessed through GenBank accession numbers OR906130, OR906131, and OR906133.

Reference

Abide, M., Kidanemariam, D., Kebede, M. &Abraham, A. (2023). Sugarcane bacilliform viruses in Ethiopia: genetic diversity and transmission by pink sugarcane mealybug. *Australasian Plant Pathology*, 52, 613-624

Abide, M., Kidanemariam, D., Yimam, T., Worku, Y., Kebede, M. & Abraham, A. (2022). Protocol optimization for elimination of sugarcane bacilliform virus and rapid propagation of virus-free sugarcane using meristem tip culture. *Tropical Plant Pathology*, *47*(5), 693-697.

Abraham, A.D., Menzel, W., Lesemann, D. –E, Varrelmann, M. and Vetten, H.J. 2006. Chickpea chlorotic stunt virus: a new polerovirus infecting cool-season food legumes in Ethiopia. Phytopathology 96:437-44

Adams, I.P., Braidwood, L.A., Stomeo, F., Phiri, N., Uwumukiza, B., Feyissa, B., Mahuku, G., Wangi, A., Smith, J., Mumford, R. and Boonham, N. (2017). Characterising maize viruses associated with maize lethal necrosis symptoms in Sub Saharan Africa. *BioRxiv*, 161489.

Adams, I.P., Fox, A., Boonham, N., Massart, S. and De Jonghe, K. (2018). The impact of high throughput sequencing on plant health diagnostics. *European Journal of Plant Pathology*, *152*,909-919.

Addy, H.S., Nurmalasari, Wahyudi, A.H.S., Sholeh, A., Anugrah, C., Iriyanto, F.E.S., Darmanto, W. and Sugiharto, B. (2017). Detection and response of sugarcane against the infection of Sugarcane mosaic virus (SCMV) in Indonesia. *Agronomy*, *7*(3), 50.

Asinari, F., Perez Gomez, S.G., Easdale, C., Rago, A.M., Balzarini, M. and Cafrune, E.E. (2021). Impact of yellow leaf virus on sugarcane commercial fields. *Tropical Plant Pathology*, *46*(1), 37-43.

Batista, J.G., Boiteux, L.S., Melo, F.F.S., Miranda, B.E.C., Ribeiro, S.G., Fonseca, M.E.N. and Pereira–Carvalho, R.C. (2021). First report of maize striate mosaic virus (Geminiviridae) in sugarcane (Saccharum officinarum). *Australasian Plant Disease Notes*, *16*, 1-3.

Boukari, W., Kaye, C., Wei, C., Hincapie, M., LaBorde, C., Irey, M. and Rott, P. (2019). Field infection of virus-free sugarcane by Sugarcane yellow leaf virus and effect of yellow leaf on sugarcane grown on organic and on mineral soils in Florida. *Plant disease*, *103*(9), 2367-2373.

Cornejo-Franco, J.F., Flores, F., Mollov, D. and Quito-Avila, D.F.M. (2021). An umbra-related virus found in babaco (Vasconcellea× heilbornii). *Archives of virology*, *166*(8), 2321-2324.

Daugrois, J.H., Filloux, D., Julian, C., Claude, L., Ferdinand, R., Fernandez, E., Fontes, H., Rott, P.C. and Roumagnac, P. (2021). Comparison of the virome of quarantined sugarcane varieties and the virome of grasses growing near the quarantine station. *Viruses*, *13*(5), 922.

Fernandez, E., Ferdinand, R., Filloux, D., Guinet, I., Julian, C., Rott, P., Roumagnac, P. and Daugrois, J.H. (2020). Improvements in virus detection at CIRAD's sugarcane quarantine using both viral metagenomics and PCR-based approaches. *Sugar Tech*, *22*, 498-503.

Formann, S., Hahn, A., Janke, L., Stinner, W., Sträuber, H., Logroño, W. and Nikolausz, M. (2020). Beyond sugar and ethanol production: value generation opportunities through sugarcane residues. *Frontiers in Energy Research*, *8*, 579577.

Haregu, S., Kidanemariam, D. and Abraham, A. (2022). Molecular identification and characterization of badnaviruses infecting sugarcane in Ethiopia. *Acta Virol*, 3-10.

He, E.Q., Bao, W.Q., Sun, S.R., Hu, C.Y., Chen, J.S., Bi, Z.W., Xie, Y., Lu, J.J. and Gao, S.J. (2022). Incidence and distribution of four viruses causing diverse mosaic diseases of sugarcane in China. *Agronomy*, *12*(2), 302.

Holkar, S.K., Balasubramaniam, P., Kumar, A., Kadirvel, N., Shingote, P.R., Chhabra, M.L., Kumar, S., Kumar, P., Viswanathan, R., Jain, R.K. and Pathak, A.D. (2020). Present status and future management strategies for Sugarcane yellow leaf virus: A major constraint to the global sugarcane production. *The Plant Pathology Journal*, *36*(6), 536.

Koloniuk, I., Přibylová, J., Čmejla, R., Valentová, L. and Fránová, J. (2022). Identification and characterization of a novel umbra-like virus, strawberry virus A, infecting strawberry plants. *Plants*, *11*(5), 643.

Kumar, A., Pal, S. and Chand, H. (2019). Insect pests of sugarcane and their management: an overview. *Advances in agricultural entomology*, *3*, 1-18.

Lin, Y., Ali, N., Hajimorad, M.R., Zhang, L., Qi, X., Zhou, L., Wen, R. and Chen, B., 2021. Incidence, Geographical Distribution, and Genetic Diversity of Sugarcane Striate Virus in Saccharum Species in China. *Plant Disease*, *105*(11), 3531-3537. 10.1094/PDIS-10-20-2307-RE

Lu, G., Wang, Z., Xu, F., Pan, Y.B., Grisham, M.P. and Xu, L. (2021). Sugarcane Mosaic Disease: Characteristics, Identification and Control. *Microorganisms* 2021, 9, 1984.

Lockhart, B.E., Autrey, J.C. & Comstock, J.C (1992) Partial purification and serology of sugarcane mild mosaic virus, a mealybugtransmitted closterolike virus. Phytopathology, 82, 691–695.

Madugula, S. and Gali, U.D. (2017). Virus indexing for Sugarcane Yellow Leaf Virus (SCYLV) in field varieties and in vitro regenerated plantlets of sugarcane. *Australasian Plant Pathology*, *46*, 433-439.

Malapi-Wight, M., Adhikari, B., Zhou, J., Hendrickson, L., Maroon-Lango, C.J., McFarland, C., Foster, J.A. and Hurtado-Gonzales, O.P.(2021). HTS-based diagnostics of sugarcane viruses: Seasonal variation and its implications for accurate detection. *Viruses*, *13*(8), 1627.

Meena, M.R., Appunu, C., Arun Kumar, R., Manimekalai, R., Vasantha, S., Krishnappa, G., Kumar, R., Pandey, S.K. and Hemaprabha, G. (2022). Recent advances in sugarcane genomics, physiology, and phenomics for superior agronomic traits. *Frontiers in Genetics*, *13*, 854936.

Mollov, D., Filloux, D., Grinstead, S., Bolus, S., Claude, L., Fernandez, E., Julian, C., Candresse, T., Lockhart, B., Roumagnac, P. and Rott, P. (2023). Sugarcane mild mosaic virus: The rediscovery of an identified but unrecognized virus. *Plant Pathology*, *72*(2), 312-321.

Nithya, K., Parameswari, B., Kumar, S., Annadurai, A., Nithyanantham, R., Mahadevaswamy, H.K. and Viswanathan, R., (2023). Prospecting true ScYLV resistance in Saccharum hybrid parental population in India by symptom phenotyping and viral titre quantification. *3 Biotech*, *13*(5), 125.

Nithya, K., VishnuVardhan, J., Balasaravanan, S., Vishalakshi, D., Kaverinathan, K. and Viswanathan, R. (2021). First report of Maize yellow mosaic virus (MaYMV) infecting sugarcane in India and its molecular characterization. *Australasian Plant Pathology*, *50*, 633-638.

Putra, L.K., Kristini, A., Achadian, E.M.A. and Damayanti, T.A. (2014). Sugarcane streak mosaic virus in Indonesia: distribution, characterisation, yield losses and management approaches. *Sugar Tech*, *16*, 392-399.

Quito-Avila, D.F., Alvarez, R.A., Ibarra, M.A. and Martin, R.R. (2015). Detection and partial genome sequence of a new umbra-like virus of papaya discovered in Ecuador. *European Journal of Plant Pathology*, *143*, 199-204.

Quito-Avila, D.F., Reyes-Proaño, E.G., Mendoza, A., Margaria, P., Menzel, W., Bera, S. and Simon, A.E. (2022). Two new umbravirus-like associated RNAs (ulaRNAs) discovered in maize and johnsongrass from Ecuador. *Archives of Virology*, *167*(10), 2093-2098.

Rott, P. (2018) Achieving sustainable cultivation of sugarcane. Vol. 2. Breeding, pests and diseases. Cambridge, UK: Burleigh Dodds Science Publishing.

Sastry, K.S. and Sastry, K.S. (2013). Plant virus transmission through vegetative propagules (asexual reproduction). *Seed-borne plant virus diseases*, 285-305.

Shabbir, R., Zhaoli, L., Yueyu, X., Zihao, S. and Pinghua, C. (2022). Transcriptome Analysis of Sugarcane Response to Sugarcane Yellow Leaf Virus Infection Transmitted by the Vector Melanaphis sacchari. *Frontiers in Plant Science*, *13*, 921674.

Singh, D., Rao, G.P., Snehi, S.K., Raj, S.K., Karuppaiah, R. and Viswanathan, R. (2011). Molecular detection and identification of thirteen isolates of Sugarcane yellow leaf virus associated with sugarcane yellow leaf disease in nine sugarcane growing states of India. *Australasian Plant Pathology*, *40*, 522-528.

Strachan, S., Bhuiyan, S.A., Thompson, N., Nguyen, N.T., Ford, R. and Shiddiky, M.J. (2022). Latent potential of current plant diagnostics for detection of sugarcane diseases. *Current Research in Biotechnology*. 4:475-492. doi: 10.1016/j.crbiot.2022.10.002

Sun, S.R., Chen, J.S., Yang, J., Huang, X.C., Huang, M.T. and Gao, S.J. (2019). First report of Maize yellow mosaic virus infecting sugarcane in China. *Plant Disease*, *103*(9), 2482.

Tahir, M.N., Bolus, S., Grinstead, S.C., McFarlane, S.A. and Mollov, D. (2021). A new virus of the family Tombusviridae infecting sugarcane. *Archives of Virology*, *166*(3), 961-965. https://doi.org/ 10.1007/ s00705-020-049 08-9

Tamura, K., Stecher, G. and Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022-3027. doi:10.1093/ molbev/msab120

Viswanathan, R. (2023) Emerging diseases of sugarcane and new disease management strategies. *Journal of Sugarcane Research* 12: 01–15. https://doi.org/10.37580/JSR .2022. 1.12 .1-15

Viswanathan, R., Chinnaraja, C., Malathi, P., Gomathi, R., Rakkiyappan, P., Neelamathi, D. and Ravichandran, V. (2014). Impact of Sugarcane yellow leaf virus (ScYLV) infection on physiological efficiency and growth parameters of sugarcane under tropical climatic conditions in India. *Acta physiologiae plantarum*, *36*, 1805-1822.

Wang, K.L., Deng, Q.Q., Chen, J.W. and Shen, W.K. (2020). Physiological and molecular mechanisms governing the effect of virus-free chewing cane seedlings on yield and quality. *Scientific Reports*, *10*(1), 10306. doi: 10.1038/s41598-020-67344-4

Wang, W., Wang, J., Feng, X., Shen, L., Feng, C., Zhao, T., Xiao, H., Li, S. and Zhang, S., 2022. Breeding of virus-resistant transgenic sugarcane by the integration of the Pac1 gene. *Frontiers in Sustainable Food Systems*, *6*, 925839. doi: 10.3389/fsufs.2022.925839

Wang, X., Olmedo-Velarde, A., Larrea-Sarmiento, A., Simon, A.E., Kong, A., Borth, W., Suzuki, J.Y., Wall, M.M., Hu, J. and Melzer, M. (2021). Genome characterization of fig umbra-like virus. *Virus Genes*, *57*, pp.566-570. https://doi.org/10.1007/s11262-021-01867-4.

Zhang, R.Y., Li, W.F., Huang, Y.K., Pu, C.H., Wang, X.Y., Shan, H.L., Cang, X.Y., Luo, Z.M. and Yin, J. (2018). Genetic diversity and population structure of Sugarcane streak mosaic virus in Yunnan province, China. *Tropical Plant Pathology*, *43*, 514-519.

Zhang, Y., Huang, Q., Yin, G., Jia, R., Lee, S., Xiong, G., Yu, N., Pennerman, K.K., Liu, Z. and Zhang, S. (2015). Genetic diversity of viruses associated with sugarcane mosaic disease of sugarcane inter-specific hybrids in China. *European journal of plant pathology*, *143*, 351-361.