SEQUENCE ANALYSIS OF COAT PROTEIN GENE AND 3` NON-CODING REGION OF SWEET POTATO FEATHERY MOTTLE VIRUS (SPFMV) ISOLATED FROM EGYPT

A. ASHOUB AND MERVAT M. M. EL FAR

Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), 9 Gamaa St. 12619, Giza, Egypt

Before the advances of data obtained by sequencing of potyviruses, species and strains of viruses were distinguished on the bases of host range, symptoms on indicator plants and serological relationships. Currently, availability of coat protein and the full genome sequences established the criteria that can be used in particular to distinguish sequences representing closely-related virus species from those strains of the same species (Shukla and Ward, 1988; Adams *et al*., 2005). Additionally, comparisons showed that nucleotide sequences of the 3' untranslated region of the potyviral genome can serve as an aid to identify and classify potyviruses (Frenkel *et al*., 1989). Increased knowledge of conserved viral sequences allowed the design of oligonucleotide primers for their use in reverse transcriptional polymerase chain reaction (RT-PCR) amplification and rapid identification of uncharacterized potyviruses (Colint *et al*., 1994; Langeveld *et al*., 1991; Pappu *et al*., 1993).

Sweet potato feathery mottle virus (SPFMV) belongs to potyviruses, the largest group of plant positive sense single stranded RNA viruses (Ward and Shukla, 1991). It has long been recognized to infect sweet potatoes worldwide (Moyer and Salazar, 1989). The virus is transmitted mechanically and by several genera of aphids including the cotton aphid (*Aphis gossypii*) and the green peach aphid (*Myzus persicae*) in a non-persistent manner (Brunt *et al*., 1996). SPFMV is a major component of sweet potato virus disease that is drastically affecting the yield of sweet potatoes worldwide (Gibson *et al*., 1998; Gibson and Aritua, 2000). Several isolates and strains of SPFMV have been characterized in different parts of the world, the most important ones being the ordinary strain “O” (Usugi *et al*., 1991), the russet crack strain “RC” (Moyer and Salazar, 1989), the common strain “C” (Cali and Moyer, 1981) and severe strain “S” (Mori *et al*., 1995). The first full-length sequence of SPFMV was published earlier (Sakai *et al*., 1997) and showed that it is the largest genome among potyviruses with a MW of 10,820 nucleotides excluding the poly (A) tail, which is incapcidated into a 35 kDa coat protein subunits. Several attempts were carried out to identify SPFMV on the basis of RT-PCR and sequence analysis

(Colinet *et al*., 1998; Ryu *et al*., 1998; Hanada *et al*., 2000; Sakai, 2000). Recently, several partial sequences of SPFMV coat protein gene and 3' non coding region from different geographic areas are made available on the databases (NCIB database).

In this study, we present the sequence of coat protein gene and the 3' non-coding region of SPFMV isolated from Egypt with its homology to the already available sequences on the database.

MATERIALS AND METHODS

Sweet potato (*Ipomoea batatas* var. Abees) carrying the Egyptian isolate of SPFMV (Ashoub *et al*., 2007) was used to isolate RNA and amplifying the coat protein gene and the 3'end of the virus as following:

Total RNA of SPFMV-sweet potato infected plants was extracted following the method described by Ashoub *et al*. (2006) with further purification using Promega SV-40 total RNA extraction kit (Promega, USA) according to the manufacturer`s recommendations.

One µg RNA was denatured at 70°C for 5 min. in the presence of 100 pmol primer Oligo-dT (5'-TTTTTTTTTTTTTTTTTTYI-3'), where Y is A, G or C residue and I is an insoin residue. Samples were chilled on ice for 2 min. RNA was reverse-transcribed for 1 h. at 42°C in the presence of 1 × reverse transcriptase buffer, 20 mM dNTPS, and 200 units M-MuLV reverse transcriptase (Promega, USA). After incubation, PCR was carried out using 5 µl cDNA in 1 × PCR buffer, 50 pmol primer Oligo-dT, 10 pmol primer P-SPFMV-CP3`-F (5`-CTTCAGTGACGTTGCTGAAGC-3`), 1.5 mM $MgCl_2$, 10 mM dNTPs, and 2 units *Taq* DNA polymerase (Fermentas, USA). Samples were subjected to 30 cycles of PCR with 15 sec. of denaturing at 94°C, 15 sec. of annealing at 50°C, and 1 min. of extension at 72°C. The series of cycles was preceded by 3 min. initial denaturing at 94°C. Five µl of the PCR products were analysed on a 1% agarose gel in TAE buffer, visualized by ethidium bromide staining and photographed (Ausubel *et al.*, 1995). PCR product was purified using Wizard SV gel and PCR clean-up system (Promega, USA) following the manufacturer`s instructions. After purification, PCR was introduced into pGEM-T Easy vector system (Promega, USA) following the manufacturer instructions to produce pSPFMV-CP3'. Vector was transformed into *E. coli* JM109. Positives were selected and plasmid DNA was isolated using Wizard plus SV minipreps DNA purification System (Promega, USA) following the manufacturer`s instructions.

To amplify the rest of coat protein gene, the same described above approach was carried out except that primer PSPFMV-CP5`-R (5'-AAGAGGTTATGTATATTTCTAGTAA-3') was used to generate the cDNA in combination with SPFMV-CP5`-F (5'-

GAGTACAACCTAGCGCTGACGATC TCAGC-3') to generate the PCR product. Additionally, the concentration of PSPFMV-CP-R and SPFMV-CP-F was 10 pmol each for the PCR reaction and the annealing temperature was 60°C. After ligation as mentioned above, the generated vector was named pSPFMV-CP5`.

Vectors were subjected to sequence analysis using T7 and SP6 promoter primers (Promega, USA), sequence reactions were carried out using the 310 sequencer (Applied bio systems) and analyzed using DNA star program.

RESULTS AND DISCUSSION

In this study, the coat protein gene and the 3`non coding region were reverse transcribed, amplified as PCR products, cloned and subjected to sequence analysis in two steps. In the first step, reverse transcription reaction was carried out based on the Oligo-d-T primer since all potyviruses have a polyadenilation 3' terminal. In PCR amplification reaction, a forward primer was designed from a conserved region for all SPFMV published in the database of NCIB (Ashoub *et al*., 2007). The amplified PCR product (Fig. 1-A) has a molecular weight size of 550 bp. When pSPFMV-CP3' was subjected to sequencing reaction, it read 544 nucleotides, excluding the polyadinilation tail and representing 222 nucleotides of 3' non coding region of the virus in addition to 322 nucleotides of the coat protein gene C-terminus. Comparing SPFMV sequence obtained from the Egyptian isolate with sequences published on the database, it revealed that the similarity was 98% to the South Africa strain C (accession number AY459601). For the second step to amplify the rest of the coat protein gene, primer SPFMV-CP5' -F was synthesized based on the sequence of the South Africa strain C located 104 nucleotides upstream of the coat protein gene start codon combined with primer PSPFMV-CP5`-R designed on the bases of the available sequence of SPFMV Egyptian isolate obtained from the first step. The amplified PCR product (Fig. 1-B) has a molecular weight of 1024 bp. The sequence of pSPFMV-CP5', excluding the 104 nucleotides upstream of the CP gene, in combination with pSPFMV-CP3' represent the complete coat protein gene sequence of the SPFMV Egyptian isolate in addition to the 3' non-coding region with a molecular size of 1164 nucleotides (Fig. 2-A). The comparison analysis with other published SPFMV coat protein gene and 3´non coding region sequences indicated the highest similarity (98%) of the Egyptian isolate with South Africa strain C. Strains used in the comparison, their accession number, geographical location and degree of similarity to the Egyptian isolate are indicated in Table (1). Results of the phylogenetic tree showed the similarity between the Egyptian isolate and the other isolates used in the comparison, which indicated that the Egyptian isolate falls within the group of strain C of SPFMV (Fig. 3). Deduced amino acid analysis of the coat protein gene of SPFMV from Egypt indicated that the protein codes for

313 amino acids with a calculated molecular weight of 35 kDa (Fig. 2-B) which is in agreement with Sakai *et al*. (1997). It has 98.4% similarity with the South Africa strain C when both were compared based on their deduced amino acids. Moreover, the 3`non coding region of the Egyptian isolate was 98.2% similar to the South Africa isolates, while the coat protein gene was 98% similar to the South Africa isolate C.

The availability of sequence information for the coat protein gene is a necessity for downstream applications like molecular and serological diagnostic tools and improving sweet potato crop by introducing resistance *via* established gene transformation strategies.

SUMMARY

The coat protein gene and the 3' non-coding region of sweet potato feathery mottle virus, SPFMV, isolated from Egypt was subjected to reverse transcriptional polymerase chain reaction, RT-PCR, cloned and applied to sequence analysis. The 1164 nucleotides sequenced representing the full-length coat protein gene were found to code for 330 amino acids with MW of 35 kDa. In addition, 222 nucleotides of 3' non-coading region excluding the poly-A tail were sequenced. Data comparison to the published sequences revealed that the nucleic acid sequence of the Egyptian SPFMV isolate has 98% homology with the common strain, C, isolated from South Africa and 98.4% homology of the deduced amino acids. The phylogenic analysis indicated that the Egyptian isolate occurs within the C strains of identified SPFMV.

REFERENCES

Adams, M. J., J. F. Antoniw and C. M. Fauquet (2005). Molecular criteria for genus and species discrimination within the family *Potyvirida.* Archives of Virology, 150: 459-479.

Ashoub, A., M. Knoblauch, W. S. Peters and A. J. E. van Bel (2006). A simple extraction method for RNA isolation from plants. Egypt. J. Genet. & Cytol., 35: 187-194.

Ashoub, A., M. M. El Far, D. Prüfer and T. Nasr El-Din (2008). Comparison of methods to detect sweet potato feathery mottle virus (SPFMV) in sweet potatoes. Egyptian J. Genet. & Cytol. (Accepted)

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. D. Seidman, J. A. Smith and K. Struhl (Edt.) (1995). Current Protocols in Molecular Biology. (New York: John Wiley and Sons).

Brunt, A. A., K. Crabtree, M. J. Dallwitz, A. J. Gibbs, L. Watson and E. J. Zurcher (Eds.) (1996). Plant Viruses Online: Descriptions and Lists from the VIDE Database. Version: 20th August 1996.' URL http://biology.anu.edu.au/Groups/MES/vide/

Cali, B. B. and J. W. Moyer (1981). Purification, serology, and particle morphology of two russet crack strains of sweetpotato feathery mottle virus. Phytopatology, 71: 302-305.

Colinet, D., J. Kummert, P. Lepoivre and J. Semal (1994). Identification of distinct potyviruses in mildly infected sweetpotato by the polymerase chain reaction with degenerate primers. Phytopathology, 84: 65-69.

Colinet, D., M. Nguyen, J. Kummert, P. Lepoivre and F. Z. Xia (1998). Differentiation among potyviruses infecting sweet potato based on genus and virus-specific reverse transcription polymerase chain reaction. Plant Diseases, 82: 223-229.

Frankel, M. J., C. W. Ward and D. D. Shukla (1989). The Use of 3' Noncoding Nucleotide Sequences in the Taxonomy of Potyviruses: Application to Watermelon Mosaic Virus 2 and Soybean Mosaic Virus-N. Journal of General Virology, 70: 2775-2783.

Gibson, R. W. and V. Aritua (2000). Sweetpotato virus disease in Africa. African Potato Conference Proceeding, 5: 373-377.

Gibson, R. W., I. Mpembe, T. Alicai, E. E. Carey, R. O. M. Mwanga, S. E. Seal and H. J. Vetten (1998). Symptoms, aetiology and serological analysis of sweet potato virus disease in Uganda. Plant Pathology, 47: 95-102.

Hanada, K., J. Sakai, M. Nishiguchi, T. Usugi and M. Onuki (2000). Molecular characterization of sweetpotato viruses occurring in Japan. p 10-13. In: Nakazawa, Y., and K. Ishiguro (Eds). International Workshop on Sweetpotato Cultivar Decline Study, Proceedings, Kyushu National Agricultural Experiment Station, Miyakonjo, Japan.

Langeveld, S. A., J. M. Dore, J. Memelink, A. F. L. M. Derks, C. I. M. van der Vlugt, C. J. Asjes and J. F. Bol (1991). Identification of potyviruses using the polymerase chain reaction with degenerate primers. Journal of General Virology, 72: 1531-1541.

Mori, M., J. Sakai, T. Kimura, T. Usugi, T. Hayashi, K. Hanada and M. Nishiguchi (1995). Nucleotide sequence analysis of two nuclear inclusion body and coat protein genes of a sweet potato feathery mottle virus severe strain (SPFMV-S) genomic RNA. Archives of Virology, 140: 1473-1482.

Moyer, J. W. and L. F. Salazar (1989). Virus and virus-like diseases of

sweet potato. Plant Disease, 73: 451- 455.

Pappu, S. S., R. Brand, H. R. Pappu, E. P. Rybicki, K. H. Gough, M. J. Frenkel and C. L. Niblett (1993). A polymerase chain reaction method adapted for selective amplification and cloning of 3´ sequences of potyviral genomes: Application to dasheen mosaic virus. Journal of Virological Methods, 41: 9-20.

Ryu, K. H., S. J. Kim and W. M. Park (1998). Nucleotide sequence analysis of the coat protein genes of two Korean isolates of sweet potato feathery mottle potyvirus. Archives of Virology, 143: 557-562.

Sakai, J. (2000). Identification of four SPMV strains and SPVG by RT-PCR and RFLP methods. p. 85-89. In: Nakazawa, Y. and K. Ishiguro (Eds). International Workshop on Sweetpotato Cultivar Decline Study, Proceedings, Kyushu National Agricultural Experiment Station, Miyakonjo, Japan.

Sakai, J., M. Mori, M. Morshita, K. Tanaka, K. Hanada, T. Usugi and M. Nishiguchi (1997). Complete nucleotide sequence and genome organization of sweet potato feathery mottle virus (S strain) genomic RNA: the large coading region of the P1 gene. Archives of Virology, 142: 1553-1562.

Shukla, O. D. and W. C. Ward (1988). Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. Journal of General Virology, 69: 2703-2710.

Usugi T., M. Nakano, A. Akira and T. Hayashi (1991). Three filamentous viruses from sweet potato in Japan. Annals of the Phytopathological Society of Japan, 57: 512-521.

Ward, C. W. and D. D. Shukla (1991). Taxonomy of potyviruses: current problems and some solutions. Intervirology, 32: 269-296.

Table (1): Accession numbers of SPFMV strains found in the NCBI database and their percentage of similarity in comparison to the Egyptian isolate.

Accession Number	SPFMV Isolates and Strians	Similarity to SPFMV from Egypt
AY459601	South Africa strain C	98%
AY459594	Kenya: Kisumu strain C	95.7%
AY459596	Uganda: Namulonge strain C	94.5%
AY459591	Kenya: Kakamega strain C	93.8%
S43451	USA: Strain C	88.5%
AY459592	Kenya: Kakamega strain EA	74.1%
AY459600	Spain, Canary Islands strain EA	74.1%
AF015541	Korean strain 2	74%
AY459599	Portugal strain EA	74%
AY459593	Kenya: Kisii strain EA	73.9%
AY459597	Madagascar strain EA	73.8%
D16664	ordinary strain, SPFMV-O	73.7%
AY459598	Tanzania strain RC	73.7%
D38543	Strain Severe	73.5%
AY459602	China strain RC	73.3%
AF015540	Korean strain 1	73.2%
AY459595	Uganda: Arua strain O	73.2%
S43450	Strain RC	73.2%

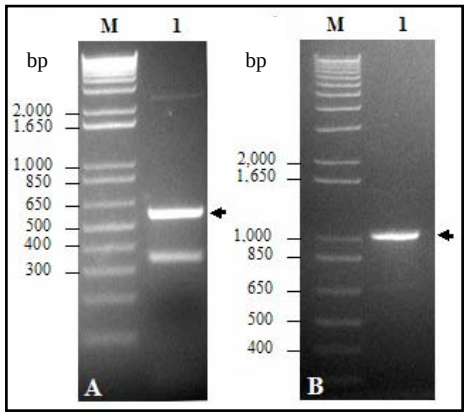

Fig. (1): RT-PCR amplifications of the SPFMV coat protein gene and 3`non-coading region. Amplification of the 3`part using Oligo-dT as reverse primer and P-SPFMV-CP3`-F as forward primer (A), and amplification of the 5`part using PSPFMV-CP5`-R as reverse primer and P-SPFMV-CP5`-F as forward primer (B). PCR products are indicated by an arrow in comparison to the molecular weight marker, M.

TCTGGTAACCCCCCTGAATTTAAAGATGCAGGTGCGAACCCACCAGCACCAAAGCCAAAAG
GGCCATATGTAGCACCAGAAATCACAGAAGTTACTGATCCCGAGGATCCCAAACAAGCTGC
GCTCCGCGAAGCTAGACAGAAACAACCTGCTGTCACACCCGAATCATACGGTAGAGATACA
GGTGAGAAACCTATGCGCTCTGTTTCACCACAAAGGGTGAAAGACAAGGATGTCAATGTTG
GTACGACAGGTACATTTTTAGTACCACGAGTTAAGCTTCATACTAGTAAAATGCGCCAACC
GAGAGTCAATGGAGTCTCCGTAGTAAACTTACAACACCTTGCAACCTATGAACCTGAGCAA
CATAACATTGGGAATACACGCTCAACTCAGGAACAGTTTCAAGCATGGTACGAAGGTGTCA
AGGGTGATTATGGTGTTGATGACGCTGGAATGGCAATCTTGTTGAATGGATTGATGGTATG
GTGCATAGAGAATGGAACATCTCCAAATATAAATGGCGTTTGGACGATGATGGACGGGGAT
GAACAAGTGACGTACCCCATAAAGCCATTACTGGATCATGCAGTGCCTACTTTTAGGCAAA
TCATGACACACTTCAGTGACGTTGCTGAAGCGTATATAGAGATGAGGAATCGCACCAAGGC
ATATATGCCTAGGTATGGATTACAACGTAATTTGACTGATATGAGTCTTGCGCGATATGCA
TTTGATTTCTATGAGCTGCACTCAACAACACCTACACGTGCTAAAGAAGCACACATGCAGA
TGAAAGCAGCTGCACTTAAGAATGCGCATAATCGGTTGTTTGGTTTGGATGGAAACGTCTC
CACGCAAGAAGAGGATACGGAGAGGCACACTGCGACTGATGTTACTAGAAATATACATAAC
CTCTTAGGAATGAGGGGTGTGCAC**TAG**GTGAAACTTTGCACTGTATTTATTTACTTATGTG
GTTTTTAGTATGCCTTTATTTAAATTCGTGTTCTTCAGTCCCGACAGAAATGGTTGGGTGT
ATCGACAAAGTGGGCTTTTAGCCTGGTCCATACACTTGAGAAGTTTCTGGTCTATTACGTA
TCATAAGGGACTCTTAAAAGTGAGGAGTACCTCGTAAGAAAAGCCTTTTTGGTTCGTGATC
GAGCA(N)

Fig. (2A): SPFMV Egyptian isolate coat protein and 3`non coding region nucleotide sequence. The TAG stop codon of the coat protein gene is underlined and the (N) indicates an unidentified number of the adenosine residue at the 3`poly-A tail.

SGNPPEFKDAGANPPAPKPKGPYVAPEITEVTDPEDPKQAALREARQKQPAVTPESYGRDT
GEKPMRSVSPQRVKDKDVNVGTTGTFLVPRVKLHTSKMRQPRVNGVSVVNLQHLATYEPEQ
HNIGNTRSTQEQFQAWYEGVKGDYGVDDAGMAILLNGLMVWCIENGTSPNINGVWTMMDGD
EQVTYPIKPLLDHAVPTFRQIMTHFSDVAEAYIEMRNRTKAYMPRYGLQRNLTDMSLARYA
FDFYELHSTTPTRAKEAHMQMKAAALKNAHNRLFGLDGNVSTQEEDTERHTATDVTRNIHN
LLGMRGVH*

Fig. (2B): Translation of SPFMV coat protein gene. The asterisk indicates the termination codon of the coat protein.

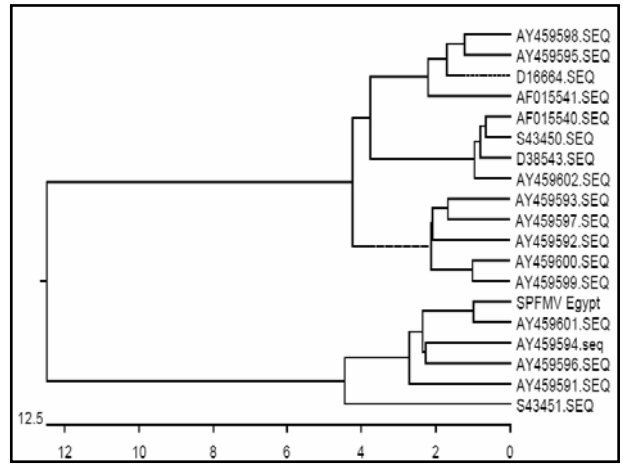

Fig. (3): Phylogenetic tree of SPFMV isolates used in this study and their relation to the Egyptian isolate.