MOLECULAR PHYLOGENETICS AND EVOLUTION
Vol. 10, No. 3, December, pp. 323–332, 1998
ARTICLE NO. FY980535

Molecular Cloning, Sequencing, and Phylogenetic Relationships of a New Potyvirus: Sugarcane Streak Mosaic Virus, and a Reevaluation of the Classification of the Potyviridae[1]

Jeffrey S. Hall,* Byron Adams,* Thomas J. Parsons,*[,2] Roy French,† Leslie C. Lane,* and Stanley G. Jensen†[,3]

*Department of Plant Pathology and †USDA-ARS, University of Nebraska–Lincoln, Lincoln, Nebraska 68583-0722

Received August 5, 1997; revised April 23, 1998

The nucleic acid of a serologically distinct potyvirus, originally isolated out of sugar cane from Pakistan, was reverse transcribed and the 3′ terminal 2000 bp was PCR amplified, cloned, and sequenced. Phylogenetic comparisons of viruses representing each genus of the Potyviridae show that the Pakistani isolate is most closely related to the rymoviruses wheat streak mosaic virus (WSMV) and brome streak mosaic virus. We therefore propose that this new virus species be named sugar cane streak mosaic virus to reflect its similarity to WSMV. The phylogenetic data also show that the genus *Rymovirus* contains at least two unique evolutionary lineages. Thus the current taxonomy, based on transmission vector, is paraphyletic. We present an analysis of the taxonomic relationships among members of the family and propose a classification that both resolves the paraphyly and more accurately represents the evolutionary history of the Potyviridae.
© 1998 Academic Press

INTRODUCTION

The Potyviridae is the largest and most economically important family of plant viruses. With over 200 species described, this family accounts for nearly 25% of the known plant viruses and causes diseases in almost all commercial crops. The virions are flexuous rods, roughly 750 nm in length and have a genome of positive-sense RNA, approximately 10 kb long with a viral encoded protein (VPg) covalently attached to the 5′ terminus. There is a single long open reading frame (ORF) encoding a polyprotein that is posttranslationally processed into the individual gene products by viral proteases. These viruses also characteristically induce the formation of nuclear and cytoplasmic inclusion bodies in the host's cells (reviewed by Shukla *et al.,* 1994).

Because of the importance of this family, potyviruses have been studied extensively. Classification historically has been based on the type of vector, host range, symptom expression, capsid protein serology, and protease digestion patterns, and also the morphology and serology of the inclusion bodies (Shukla *et al.,* 1994). With advances in technology, the amino acid and nucleic acid sequences, particularly of the coat protein, have become standards for comparing these viruses (Ward *et al.,* 1992). Nevertheless, the phylogenetic relationships and classification of the Potyviridae continue to generate much discussion and debate.

Current classification places the Potyviridae in the picornavirus-like supergroup of plus-stranded, RNA viruses, with the families Picornaviridae, Nepoviridae, and Comoviridae. The family is composed of four genera *Potyvirus, Rymovirus, Bymovirus,* and *Ipomovirus* (Shukla *et al.,* 1994; Colinet *et al.,* 1996) that are defined by the type of vectors that transmit the viruses, namely aphids, eriophyid mites, fungi, and whiteflies, respectively.

Recent evidence suggests that the aphid-transmitted macluraviruses comprise another genus (Badge *et al.,* 1997) and that the genus *Rymovirus* is actually two distinct groups of viruses (Salm *et al.,* 1996b). The increasing availability of potyvirus sequence data presents new opportunities for comparisons that may reveal insights about the phylogeny and evolution of this important family of viruses. An accurate phylogenetic framework is an essential component of accurate nomenclature and the comparative understanding of the etiology, epidemiology, and evolution and speciation of these viruses.

[1] This study was a cooperative investigation of the USDA-ARS and the University of Nebraska Agricultural Experiment Station and was published as Paper No. 11966, Journal Series, Nebraska Agricultural Experiment Station.

[2] Current address: Armed Forces DNA Identification Laboratory, Armed Forces Institute of Pathology, 1413 Research Blvd., Rockville, MD 20850.

[3] To whom correspondence should be addressed at 406 Plant Science Hall, University of Nebraska–Lincoln, Lincoln, NE 68583-0722. Fax: (402) 472-2853. E-mail: sjensen@unlinfo.unl.edu.

1055-7903/98 $25.00
Copyright © 1998 by Academic Press
All rights of reproduction in any form reserved.

An apparent potyvirus isolated from quarantined sugar cane imported from Pakistan was tentatively identified by Gillaspie *et al.* (1978, 1984) as sugar cane mosaic virus-strain F (SCMV-F) based on symptom expression in indicator hosts. However, the host range of the Pakistani virus was unlike that of other known potyviruses and the purified capsid protein gave a unique tryptic digest profile (Jensen and Hall, 1993a, b). Analysis by ELISA and Western blotting revealed no serological relationships between this virus and sugar cane mosaic virus-strain A (SCMV-A), maize dwarf mosaic virus-strain A (MDMV-A), johnsongrass mosaic virus (JgMV), and sorghum mosaic virus (SrMV), all viruses that infect monocotyledonous plants. There was also no serological cross reaction with two rymoviruses, *Hordeum* mosaic virus (HoMV) or *Agropyron* mosaic virus (AgMV) (Jensen and Hall, 1993a). These preliminary data suggested that the Pakistani virus was actually a novel virus of sugar cane.

The objectives of this study were to characterize the Pakistani virus and analyze its phylogenetic relationships within the family Potyviridae. In doing so, we have inferred relationships among the taxa that more accurately represent the evolutionary history of this important virus family.

MATERIALS AND METHODS

Virus Preparation

The virus was propagated and maintained in sorghum (*Sorghum bicolor* var. QL-3) by mechanical inoculation and purified by the procedure developed by Lane (1986) with the modification that the virus was isolated and purified with 0.1 M Tris (pH 9.0) buffer.

Western Blotting

Polyacrylamide gel electrophoresis, protein transfer, and membrane staining were performed using the published methods of Brakke *et al.* (1990).

RNA Extraction and cDNA Synthesis

RNA was isolated from the intact virus by digestion with Proteinase K (100 μg/ml) and 1% SDS in 50 mM glycine buffer (pH 9.5) containing 50 mM NaCl and 5 mM EDTA and followed by phenol/chloroform extraction and subsequent ethanol precipitation. The purified viral RNA was used as a template for cDNA synthesis using 10 μM of primer RCF1 (5′-AGCTGGATCC(T_{14})-3′ [*Hin*DIII sticky end + *Bam*HI site + oligo d(T)]) and 0.5 units/μl of AMV reverse transcriptase (Boehringer Mannheim) in 50 mM Tris–HCl (pH 8.0), 5 mM 2-mercaptoethanol, 10 mM $MgCl_2$, 70 mM KCl, and 0.4 mM each dNTP.

Polymerase Chain Reaction and Cloning

The cDNA was amplified by polymerase chain reaction (PCR) in 50-μl reactions using 50 pM each of primers RCF1 and Poty4 (5′-GCGGGATCCGTNTGYGTNGAYGAYTTYAAYAA-3′) (Zerbini *et al.,* 1995; Salm *et al.,* 1996a), 2.5 units of *Taq* DNA polymerase (Cetus) in 1× reaction buffer (supplied with the enzyme), and 0.2 mM each dNTP and 2 μl of cDNA. Poty 4 is a degenerate primer that corresponds to a conserved sequence in the potyvirus NIb gene and has been used to amplify several potyvirus sequences. A PCR program of 94°C (45 s), 43°C (45 s), 72°C (1 min) for 35 cycles with a final 72°C extension for 5 min was performed in a Perkin–Elmer 9600 Thermal Cycler.

The 2-kb PCR product was cleaved into ≥300-bp

TABLE 1

The Current Classification of the Family Potyviridae with Associated Vectors and Virus Species Used for the Phylogenetic Analysis

Genus	Vector	Virus	GenBank Accession No.
Bymovirus	Fungi	Barley yellow mosaic virus (BYMV)	X69757
		Wheat spindle streak mosaic virus (WSSMV)	X73883
Macluravirus[a]	Aphids	*Maclura* mosaic virus (MacMV)	U58771
		Narcissus latent virus (NLV)	U58770
Ipomovirus[a]	Whiteflies	Sweet potato mild mottle virus (SPMMV)	Z48058
Rymovirus	Eriophyid mites	Ryegrass mosaic virus (RgMV)	U27383
		Hordeum mosaic virus (HoMV)	U30615
		Agropyron mosaic virus (AgMV)	U30616
		Brome streak mosaic virus (BrSMV)	Z48506
		Wheat streak mosaic virus (WSMV)	Niblet *et al.,* 1991; R. French-unpublished data
Potyvirus	Aphids	Potato virus-strain Y (PVY)	D00441
		Plum pox virus (PPV)	M21847
		Maize dwarf mosaic virus-strain A (MDMV-A)	U07216
		Sorghum mosaic virus-strain H (SrMV-H)	U07219

[a] The genera *Macluravirus* and *Ipomovirus* are proposed members of the family along with the three established genera, *Potyvirus, Rymovirus,* and *Bymovirus.*

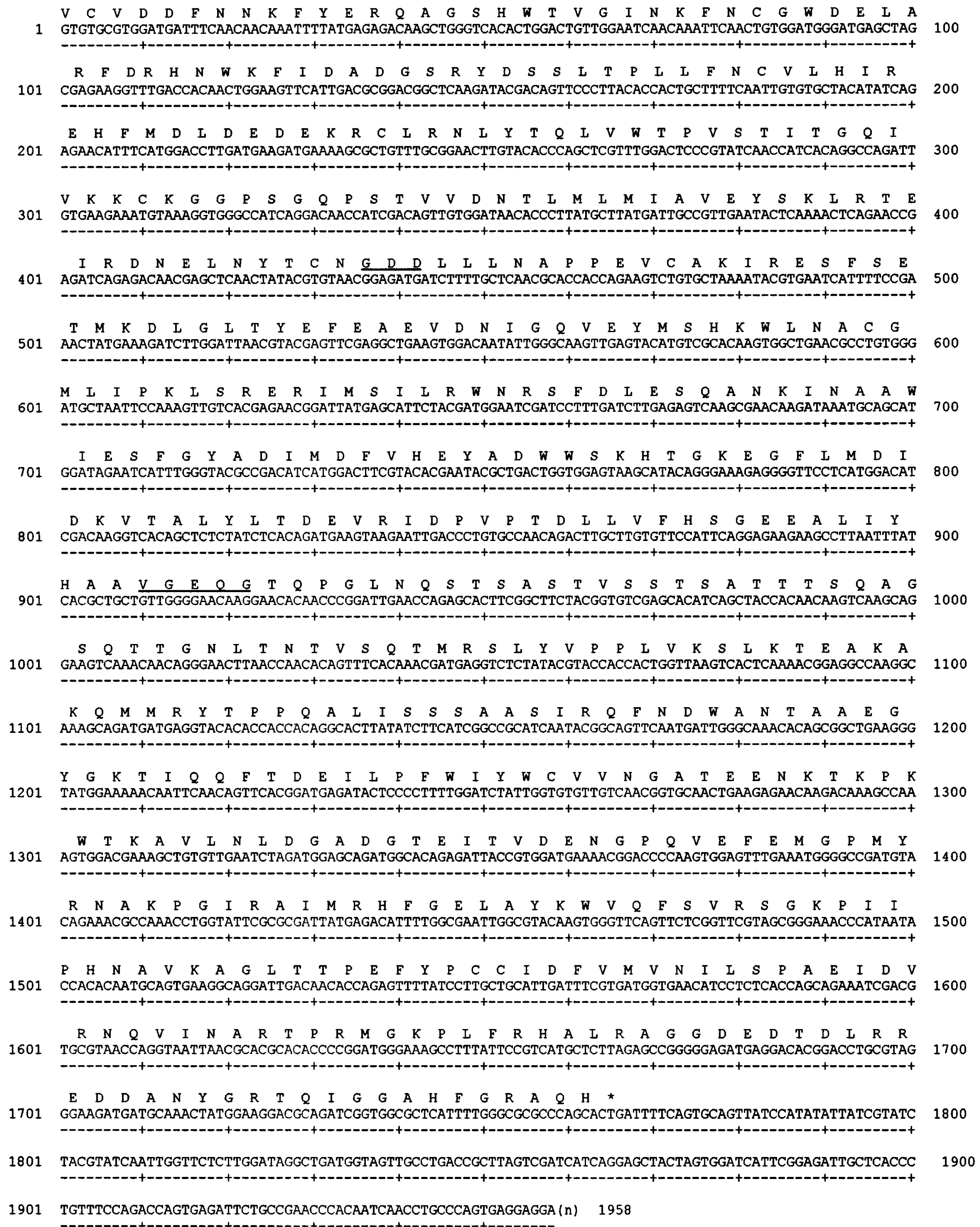

FIG. 1. The 3′ terminal, 1958 nucleotide sequence of the Pakistani virus cDNA with the derived amino acid sequence. The GDD motif in the replicase and the putative protease cleavage site VGEQ/G are underlined.

fragments using a Tekmar TM50 sonicator with 4× 20-s pulses at 50% power. The ends of these fragments were repaired using T4 DNA polymerase and ligated into *Sma*I endonuclease digested, dephosphorylated pBluescript SK+ vector (Stratagene). The ligation products were transformed into *Escherichia coli* XL1-Blue cells (Stratagene).

Nucleotide Sequencing

Clones were sequenced by the dideoxynucleotide chain termination method (Sanger *et al.,* 1977) using the Sequenase v. 2.0 sequencing kit (United States Biochemical). The sequence was compiled using DNASIS (Hitachi, 1991, 1992).

```
             1                                                  50
BYMV     SGQPSTVVDN TLVLMTAFLY AYIHKTGDRE LALLNERFIF VCNGDDNKFA
WSSMV    SGQPSTVVDN TLVLMVSFLY AYIHKTGDTE LLKLDERFVF VCNGDDNKFA
MacMV    SGQPSTVVDN TLVLMISFYY AYAVKTRDYT FDKIDERFVF VCNGDDNKFA
NLV      SGQPSTVVDN TLVLMLSFYY AYAKKTNDLA FEHIDERFRF VCNGDDNKFS
MDMV-A   SGQPSTVVDN TLMVIIAFNY TLLS..CGVV LEKADEVCRM YANGDDLLLA
SrMV-H   SGQPSTVVDN TLMVIIAFNY TMLS..CGIE ADMIDEICKM YANGDDLLLA
PPV      SGQPSTVVDN TLMVILAMTY SLLK..LGYH PDTHDCICRY FVNGDDLVLA
PVY      SGQPSTVVDN SLMVVLAMHY ALIK..ECVE FEEIDSTCVF FVNGDDLLIA
AgMV     SGQPSTVVDN TLMVVLAMQY AIAK..HGLS PESTDSFIRY FVNGDDLVIG
HoMV     SGQPSTVVDN TLMVVLAMQY AISK..YGLG LESTDTYIRY FANGDDLLIA
RgMV     SGQPSTVVDN TIMVIIAMQY AISK..AEFP AGRLRDQIRY FANGDDLVVA
WSMV     SGQPSTVVDN TMCLIIAMEY CRMR..VEKD HEHRMRI.LY VCNGDDLLIN
BrSMV    SGQPSTVVDN TLCLMIAMEY ARQR..AISD GHLNMQM.RY VCNGDDLLIN
SPMMV    SGQPSTVVDN TLIILMIAMEY AIAK..VFVT ...RPDI.KY VCNGDDLLIN
SCSMV    SGQPSTVVDN TLMLMIAVEY SKLR..TEIR DN...EL.NY TCNGDDLLLN

             51                                                100
BYMV     ISPQFDEEFG HDFSPELVEL GLTYEFDDIT SDICENPYMS LTMVKTPFGV
WSSMV    ISPEFNAQFG HDFSPELIEL GLTYEFDDIT DDICENPYMS LTMVRTPFGI
MacMV    VSPEFVKEFG GSFTDEIAQL GLHYEFDELT TDITANPYMS LTMIDIGGRI
NLV      VSPEFVAEFG GSFGDEISEL GSLTSLMILT PDIMRNPYMS LTIVQVGDRI
MDMV-A   VNPTHVNIL. DEFGKHFAAL GLNFDFESRT SDKSELWFMS TRGIKYEE..
SrMV-H   IRPDYEHFI. DNFSKHFADL GLNFDFTSRT RDRTELWFMS TRGIKIDN..
PPV      VHPAYESIY. DELQEHFSQL GLNYTFATKT ENKEELWFMS HKGVLYDD..
PVY      VNPEKESIL. DRMSQHFSDL GLNYDFSSRT RRKEELWFMS HRGLLIEG..
AgMV     IAPEKVDLL. DTVAQSFGRL GLNYDFSSRV EKREDLWFMS HQGKLIND..
HoMV     IHPDHEALL. DTLKEHFGEL GLNYDFTSRT RDRSELWFMS HQGKLIDD..
RgMV     VEPSLSDKI. SSFSASFAEL GLSYDFSNKV NVRSELQFMS HTGKLIDG..
WSMV     ADTKDKDFIQ QYFADYMREL ELNYSFDEAY RSIEEVEYMS HTFMKRNS..
BrSMV    ANEEAKDVVQ GKYEQYIKEL ELNYCFDDAF QSIEGVEFMS HKFMLRNG..
SPMMV    CPRSTANAIS EHFKDVFADL SLNYDFDHVC DKITDVDFMS HSFMWLDTEQ
SCSMV    APPEVCAKIR ESFSETMKDL GLTYEFEAEV DNIGQVEYMS HKWL..NACG

             101                                               150
BYMV     GF..SLPVER IIAIMQWSKK GGVLHSYLAG ISAIYESFNT PKLFKSIYAY
WSSMV    GF..SLPVER IVAIMQWARR GGVLHSYLAG ISAIYESFNT PKLFKSIYAY
MacMV    GF..QLNPER ILGIVQWIKK GGIVHAAQAA FAAMIESFND PDLFCVMHSY
NLV      GF..QLNPER IVGIVQWIKK GGVLHAAQAA FAAMVEAFND PSLFTVMHTY
MDMV-A   MYIPKLEKER IVAILEWDRS LLPQYRLEAI CAAMVEAWGY KDLLHEIRKF
SrMV-H   MYIPKLEQER IVAILEWDRS LLPQYRLEAI CAAMVESWGY PQLLHEIRKF
PPV      MYIPKLEPER IVSILEWDRS NEPIHRLEAI CASMVEAWGY KELLREIRKF
PVY      MYVPKLEEER IVSILQWDRA DLPEHRLEAI CAAMIESWGY SELTHQIRRF
AgMV     MYIPMLERER VVAILEWDRS HEPEFQMGAI NAAIIESWGD DELIYQVRKY
HoMV     LYIPMLERDR VVAILEWDRS HEPEFQMDAI NAAIIEAWGD DELIHHIRKF
RgMV     MYIPMLERER ICAILEWSRS DEPQFQLDAI SAAMIEAWDD DYLLYQIRRY
WSMV     MY.PKLKRER IVAILEWQRS KEPKAIQSAI IAAYVEAFGY DEFTEMIEEF
BrSMV    IYIPKLARHR IVAILEWQRS AEPQAIKSAI LAACVEAFGY DDLTELIREY
SPMMV    MYIPKLDKER IVAILEWERS DEQFRTRSAL NAAYIESFGY EDLMTEIEKF
SCSMV    MLIPKLSRER IMSILRWNRS FDLESQANKI NAAWIESFGY ADIMDFVHEY

             151                                               200
BYMV     LLWLTEEHEA EILAAMTQSS TALPIPSMLD VYRLHYGDD. ...EIW....
WSSMV    LLWLTEEYEA DILAAMKDTA TALPIPSMLD VYRLHYGGC. ...DIE....
MacMV    LVWLLVTYRS ELVYAMHNDL VSVVYMDPCQ VFALHYNDSE DVRE.W....
NLV      LVWLLVTHKD VLLYAQENGL GSVCYMDPCQ VFALHNGSSK GLEDVK....
MDMV-A   YAWLLEMQPF SNLAKEGSA. ...PYIAESA LRNLYTGAKV SEDELNVYAR
SrMV-H   YAWILEMQPF ATLAKEGLA. ...PYIAETA LRNLYTGEGI KEGELDVYYT
PPV      YSWVLEQAPY NALSKDGKA. ...PYIAETA LKKLYTDTEA SETEIERYLE
PVY      YSWLLQQQPF ATIAQEGKA. ...PYIASMA LRKLYMDRAV DEEELRAFTE
AgMV     YNWLLEQEPY KSLADAGKA. ...PYLAETA LTKLYTDVDA SQEILDLYQN
HoMV     YSWLLEQEPY KSLADCGRA. ...PYLAETA LRKLYTDQDA TQELLDLYAN
RgMV     YSWLLDEEPY KSIAELGHA. ...PYLAEAA LKALYTGKNP DAELIAIYER
WSMV     ......AQEV SAVWPDFK.. ....LPSRQE VEDLYLTG.. TRTDLGEEIK
BrSMV    ......AISL EPVWGSF... ....LPTDGE IEQLYFEG.. IAKQEVARCL
SPMMV    ......AH.F WAKKHGLND. ...VLMEREK VRSLYVDENF DASRFEKFYP
SCSMV    ......A.DW WSKHTGKEG. ...FLMDIDK VTALYLTD.. .....EVRID

             201                                               250
BYMV     ........LQ AADPLTDAQK EAAHTAAADR ARLDLADADR RRKVEADRVE
WSSMV    ........LQ AADTQTDAQK EAARVAAADK ARADAADAAR KQKVEADRVE
MacMV    ........FD EDDESSDDED EEPTQVLQMD AETLAKDGEA KKEKDEKERE
NLV      ........FD NEDESADEDD GNITPDLELQ MDVGNLIPE. .KEKNSQNVN
MDMV-A   QFFDDLPNYL AD.EIIDVKH QAG....EN. VDAGQKT... ..........
SrMV-H   QFLKDLPEYI ED.ELIDVRH QAG....GGT VDAGATTAEA TAQAQRDAAA
PPV      ...AFYDDFN DDGESNVVVH QADEREDEEE VDAGKPSVVT APAATSPILQ
PVY      MMVALDDEF. .ELDSYEVHH QAN.....DT IDAGGSNKKD AKPEQGS...
AgMV     YVFE.LPQFD E.PTLVYHEA DDPTKLLRTI LQVRITRQTR KQEMDKKPQQ
HoMV     SMLE.LPTFD E.PTLVFHEA DDKT...... .......QPK AKHQQNPPPS
RgMV     AMLN.TPPTE DRPTKVVHEA .NVTAASSAA TQTSATSPTV TSTSGASTLT
WSMV     ECGEQYCVYE SSEAATDAVL AAANAGTGSA SSSGSTQSSQ SASTASGSGS
BrSMV    AGVDDVCKFE SAASGTN... ...EAVDEVL KAAGDDEALA RANAAATSG.
SPMMV    ESFSPFDVYV EPHASTSKTI EELQQEMEDL DA........ ..........
SCSMV    PVPTDLLVFH SGEEAL.... ........IY HAAVGEQGTQ PGLNQSTSAS

             251                                               300
BYMV     AARVKKAADA VLKPVTLTAT RMPTEDDGKL KTPS...... ..........
WSSMV    AARVKKAAA. ..DTANLTAT KVTATEDGKV TTDS...... ..........
MacMV    KAEQRRVEVE KARAEKAQVS DGAKEPQPEI KGNE...... ..........
NLV      TSDGGKNASN SATGESSKPP ENKAGKGAD. ..QG...... ..........
MDMV-A   ......DAQK EAEKKAAE.. .....EKKAK EAEAKQK... ........ET
SrMV-H   KAQRDADAKK KADDEAAERQ RQDAAAKKKA DDDAKAKADA IVKQNQIADA
PPV      PPPVIQPAPR TTASMLN..P IFTPATTQPA TKPVSQVSGP QLQTFGTYGN
PVY      ....IQPNP. .......... .......... .......... ..........
AgMV     LVNNKIEPIR KMLRINKTQA NQRKSNNGER PKAANDDKSK QLAKAQSKEV
HoMV     DSGNKSGA.. .......... ...TDNNSGK DKSG.DDKSK QLAKAQTGEK
RgMV     SSGTTSAPIA STTPPVPTAT TPPTGTTAPT TPAV...RAA NLDLAGHRKA
WSMV     SQSGSGAAQT QSNNVSVMAG LDTGGAKTGQ GSGSKGTGGS FT........
BrSMV    .......ATT PAQNV...GA GTTTPAKATP QSGRRPSFGS LI........
SPMMV    .......... ........DT TITVVQRETQ KAGIRDQIEA LR........
SCSMV    .......... .........T VSSTSATTTS QAGSQTTGN. ..........

             301                                               350
BYMV     .......... GARIPSSAA. ....DGNWSV PATKQVNAGL TLKIPLNK..
WSSMV    .......... GTKRTSAAA. ....EVTWTL PTMKQANAGL KLRIPIAK..
MacMV    .......... DVEQPASDPE EKEEEVKWVM PSINPNRGSN AIPTVNGKKL
NLV      .......... DVDPPQGDPL VDDEVVEWVI PKMSPNIGTS PIPVINGKRL
MDMV-A   KEKSTEKTGD GGSIGKDKDV DAGTSGSVSV PKLKAMSKKM RLPQAKGKNI
SrMV-H   KKKADDEAAR KAQNQKDKDV DVGTSGTVAV PKLKAMSKKM KLPQAKGKNI
PPV      EDASPSNSNA LVNTNRDRDV DAGSVGTFTV PRLKAMTSKL SLPKVKGKAI
PVY      .......... ..NKGKDKDV NAGTSGTHTV PRIKAITSKM RMPTSKGATV
AgMV     VRQNNEKRVM NSGGDDADVT IKDETKTFVI PKVEVLNKNV RMPKFKGKAM
HoMV     KRQENEKRVM QSGGDDADVS LKDDNKTFEV PTVDVINKKL RMPKYKGKAL
RgMV     KANGESQLNV RGENDDEDVP AASE...FAL PRLPTLGAKI RVPKFKGVIV
WSMV     ....SNPVRT GGRATDVQDQ TPG..LVFPA PKITT.KAIY MPKTVRDKIK
BrSMV    ....DNPI.G GNGVQDVADR TSG..IVFPV PTRKS.TSLY LPPKVKLRAT
SPMMV    ....AQQIVR PPEAQLQPDV TPAQIVTFEP PRVTGFGALW IPRQQRNYMT
SCSMV    .......... .......... .......LTN TVSQTMRSLY VPPLVKSLKT

             351                                               400
BYMV     ......LKSV PKSVMEHNNS VALESELKAW TDAVRTSLGI TTDEAWIDAL
WSSMV    ......LKSV PKSVMQHDNS IALDSELTAW ADAVRTSLGI TTDEAWQNTL
MacMV    WKR.GILKHI PKQQYDASTT KATSAQLAAW VEAVKKDLKI RNDDAWSIVL
NLV      WKR.GVLKSI PKMMFNTTST MATQAQLTSW VEEVKQALAL KTDDAWTVVI
MDMV-A   LH.LDFLLKY KPQQQDLSNT RATRAEFDRW YEAVQKEYEL DDTQ.MTVVM
SrMV-H   LH.LDFLLGY KPQQQDISNT RATRDEFDRW YDALQKEYEL DDTQ.MTVVA
PPV      MN.LNHLAHY SPAQVDLSNT RAPQSCFQTW YEGVKRDYDV TDDE.MSIIL
PVY      PN.LEHLLEY APQQIDISNT RATQSQFDTW YEAVRMAYDI GETE.MPTVM
AgMV     VN.VDHLLVY KPDQRDLSNK RATQKQVDNW VEKVAKDYGV EESS.MDIII
HoMV     VN.ADHLMKY SPDQRDLSNK RATQRQLDNW VENIVKDYNV DEGK.IDIVL
RgMV     LN.KDHLIKY TPDQRDLSNT RATQEQFEKW YSGVRNEVEK TDEE.MALLL
WSMV     PEMINNMIKY QPRTELIDNR YATTEQLNTW IKEASEGLDV TEDVFINTLL
BrSMV    PERIEKVRKY LPDPQQIDLR YSTQQELNDW IKASADGLGQ TEEAFIDNIL
SPMMV    PSYIEKIKAY VPHSNLIESG LASEAQLTSW FENTCRDYQV SMDVFMSTIL
SCSMV    EAKAKQMMRY TPPQALISSS AASIRQFNDW ANTAAEGYGK TIQQFTDEIL
```

FIG. 2. The alignment of the derived peptide sequences of viruses representing the different genera of the family Potyviridae. The different data subsets, NIb (positions 1–216) and coat protein (positions 357–579), are underlined. Viruses used in this analysis are barley yellow mosaic virus (BYMV), wheat spindle streak mosaic virus (WSSMV), *Maclura* mosaic virus (MacMV), narcissus latent virus (NLV), maize dwarf mosaic virus strain-A (MDMV-A), sorghum mosaic virus strain-H (SrMV-H), plum pox virus (PPV), potato virus-Y (PVY), *Agropyron* mosaic virus (AgMV), *Hordeum* mosaic virus (HoMV), ryegrass mosaic virus (RgMV), wheat streak mosaic virus (WSMV), brome streak mosaic virus (BrSMV), sweet potato mild mottle virus (SPMMV), and sugar cane streak mosaic virus (SCSMV).

Multiple Sequence Alignments

GenBank accession numbers and sources of the virus sequences used in the alignments are shown in Table 1. The sequence of SCSMV has been submitted to GenBank and has been assigned Accession No. U75456.

Historically, comparative sequence analysis of these viruses has focused on the coat protein gene or portions thereof. This has proven especially useful for differentiating viruses at the species and strain levels (Shukla *et al.,* 1994). To make the best use of the available data and to determine if the different genes suggest alternative phylogenetic topologies, we divided the individual sequences into subsets of (a) the entire derived peptide sequence from a highly conserved SGQP motif in the replicase (NIb) to the C-terminus of the coat protein (amino acids 109 to 588 of SCSMV peptide); (b) the coat

```
        401                                                450
  BYMV  IPFIGWCCNN GTSDKHAENQ VMQ....... .......IDS GKGAVTEMSL
 WSSMV  IPFLGWCCNN GASDKHSENQ KMQ....... .......VDA GKATLSEVSL
 MacMV  TAWCIWCANN GTSSEVDTNQ DME....... ........SD SLGKVQTVRI
   NLV  TNWCIWCANN GTSSEVDTSQ TME....... .......IRD GFGKVQAIPI
MDMV-A  SGLMVWCIEN GCSPNINGV. .......... ......WTMM DGDEQRTFPL
SrMV-H  SGLMVWVIEN GCSPNINGV. .......... ......WTMM DGDEQRKFPL
   PPV  NGLMVWCIEN GTSPNINGM. .......... ......WVMM DGETQVEYPI
   PVY  NGLMVWCIEN GTSPNVNGV. .......... ......WVMM DGNEQVEYPL
  AgMV  NGFMVWALDN GTSPNITGT. .......... ......WIMM DKEEQREYPI
  HoMV  NGFMVWALDN GTSPNISGT. .......... ......WLMM DGEKQKEYPL
  RgMV  NGFMVWCMEN GTSPDLSGS. .......... ......WTMM EGEEQISYPL
  WSMV  PGWVYHCIIN TTSPENRALG TW.RVVNNAG .........K DNEQQLEFKI
 BrSMV  PGWIVHCIVN TTSSENRKAG SW.RCVTNAG .........T ADEEQVLYDI
 SPMMV  PAWIVNCIIN GTSQERTNEH TW.RAVIMAN .........M EDQEVLYYPI
 SCSMV  PFWIYWCVVN GATEENKTKP KWTKAVLNLD GADGTEITVD ENGPQVEFEM

        451                                                500
  BYMV  SPFIVHARMN GGLRRIMRNY SDETVLLI.. ...TNNKLVA HWSMKHGASA
 WSSMV  SPFIVHARLH GGLRRIMRAY SDETVLLI.. ...SEGKLVP RWAMRHGASA
 MacMV  DSFVEPAIEN GGLRKIMRLL FRYHSGNL.. ...GQRGKND SLWNQAGFTE
   NLV  EVFVNPAVEN GGLRKIMRHF SGITHEIL.. ...KAGKRMT AWGNKRGFTE
MDMV-A  KPVIENA..S PTFRQIMHHF SDAAEAYIEY R.NSTEKYMP RYALQRNLTD
SrMV-H  KPVIEYA..S PTFRQIMHHF SDAAEAYIEY R.NSTERYMP RYGLQRNLTD
   PPV  KPLLDHA..K PTFRQIMAHF SNVAEAYIEK R.NYEKAYMP RYGIQRNLTD
   PVY  KPIVENA..K PTLRQIMAHF SDVAEAYIEM R.NKKEPYMP RYGLIRNLRD
  AgMV  EPLVRHA..Q PTLRQIMMHL SDTATGYIVL R.NTKERYMP GYGLKRNLND
  HoMV  EPIVKHA..Q PTLRQIMMHF SDAATAYIVL R.NTKGRYMP GYGLKRNLND
  RgMV  GPFCRHA..Q PTLRSIMAHF SDAATAYVVL R.NQKSRYMP RYGLKRGLND
  WSMV  EPMYKAA..K PSLRAIMRHF GEGARVMIEE SVRIGKPIIP RGFDKAGVLS
 BrSMV  EPMYSAA..N PTMRAIMRHF SDLARLVIAE SFKQGRPLIP KGYIKAGVLD
 SPMMV  KPIIINA..Q PTLRQVMRHF GEQAVAQYMN SLQVGKPFTV KGAVTAGYAN
 SCSMV  GPMYRNA..K PGIRAIMRHF GELAYKWVQF SVRSGKPIIP HNAVKAGLTT

        501                                                550
  BYMV  NA..KYAFDF FVPRSWMNPQ DIEVSKQARL AALGTGTYNT MLTSDTTNLR
 WSSMV  NA..AYAFDF FVPRPWMNPQ DIEISKQARL AALGTGTNNT MLTSDTTNLR
 MacMV  KAMTTLPFDF VEVTKTTPKT VKEQLAQAKI AAIGHGTRRA MVTDGSVHGN
   NLV  KSMIPYAFDY YVVTNTTPKT VREQLAQSKA AAIGSGVTRK MVLDGNIQGS
MDMV-A  FSLARYAFDF YEISSRTPVR AKEAHMQMKA AAVRGSNTRM FGLDGNVGEA
SrMV-H  YNLARYAFDF YEITSRTPAR RREAHMQMKA AAVGGSNTRM KGLDGNVGES
   PPV  YSLARYAFDF YEMTSTTPVR AREAHIQMKA AALRNVQNRL FGLDGNVGTQ
   PVY  MGLARYAFDF YEVTSRTPVR AREAHIQMKA AALKSAQPRL FGLDGGISTQ
  AgMV  MSLAPYAFDF YEITSETPNR VREAHLQMKA AAIRGKVNRT FGLDGTVSSG
  HoMV  MNLAPYAFDF YGITSETPNR AREVHMQMKA AAIRGKVNRP FGLDGPVASG
  RgMV  YSLAPYAFDF YEITSTSPLR ARERHAHMKA AAIRGKASRM FGLDGNVSAQ
  WSMV  INNIVAACDF IMRGADDTPN FVQVQNSVAV NRLRGIQNKL FAQARLSAGT
 BrSMV  ASSAAAACDF VVRDRHDTAT FVQVQNQVLV NRVSGITNRL FAQAMPSAGA
 SPMMV  VQDAWLGIDF LRDTMKLTTK QMEVKHQIIA ANVTRRKIRV FALAAPGDGD
 SCSMV  PEFYPCCIDF VMVNI.LSPA EIDVRNQVIN ARTPRMGKPL FRHALR.AGG

        551                  579
  BYMV  KTTNHRVLDS DGHPELT... .........
 WSSMV  KTTNHRVLDT DGHPELT... .........
 MacMV  KTSYERHVDT DNDESEHGKD IDYRPHLS.
   NLV  HASYERHVDT DNSEYEHGND VDQRPYLT.
MDMV-A  HENTERHTAG DVSPNMHSLL GVQQGH...
SrMV-H  QENTERHTAG DESRNMHSLL GVQQHH...
   PPV  EEDTERHTAG DVNRNMHNLL GVRGV....
   PVY  EENTERHTTE DVSPSMHTLL GVKNM....
  AgMV  SEDTERHTVD DVKHGTHSFY GAGMN....
  HoMV  SEDTERHTVD DVKHGTHSFY GAGMN....
  RgMV  SENTERYTVE DVNTRVHSLS GANML....
  WSMV  NEDNSRHDAD DVRENTHSFN GVNALA...
 BrSMV  NEDMARHDAQ DAAEGIHNLG GARAF....
 SPMMV  ELDTERHVVD DVARGRHSLR GAQLD....
 SCSMV  DEDTDLRRED DANYGRTQIG GAHFGRAQH
```

FIG. 2—*Continued*

protein without the highly variable N-terminal region (amino acids 370 to 588); (c) the conserved core of the NIb gene (amino acids 109 to 309); and (d–f) the corresponding nucleic acid sequences of each subset.

Multiple sequence alignments of each dataset were generated using the PILEUP program of GCG (Program manual for the Wisconsin Package, Version 8, 1994). To test the effects of different gap penalty weights on tree topologies and to optimize the alignments, eight gap initiation and extension penalties that varied from strict to lenient by a factor of 0.9 (PILEUP, GCG) were tested.

Phylogenetic Analysis

Each partitioned dataset, as well as the full-length sequence alignments, were analyzed by parsimony (PAUP v. 3.1.1; Swofford, 1993) and maximum likelihood (PUZZLE v. 2.5.1; Strimmer and von Haeseler, 1996) The PUZZLE program utilized quartet puzzling of 1000 steps, with BYMV and NLV selected as the outgroup taxa for rooting purposes. Two models of sequence evolution (Jones *et al.,* 1992; Dayhoff *et al.,* 1978) were used for the protein data sets. Models of evolution applied to the nucleotide data sets were those of Hasegawa *et al.* (1985), Tamura and Nei (1993), and Schöniger and von Haeseler (1994).

Evaluation of Tree Topology

Bootstrapping (100 replicates) was performed using PAUP (input sequences randomized each replication). Bremer support or decay indices (Bremer, 1988, 1994; Källersjö *et al.,* 1992) for the parsimony trees were calculated using AutoDecay v. 3.03 (Eriksson, 1996). Overall strength of phylogenetic signal was estimated by Relative Apparent Synapomorphy Analysis (RASA v. 2.0.0 d16; Lyons-Weiler *et al.,* 1996) and PTP tests (Archie, 1989; Faith and Cranston, 1991) of 100 replicates. A PTP estimate is included for comparative purposes, even though it is likely an inappropriate estimate of phylogenetic signal (Carpenter *et al.,* 1998). To check for possible outgroup convergence and sources of phylogenetic signal/noise, the datasets were manipulated as to outgroup and ingroup taxa.

RESULTS AND DISCUSSION

Sequence Analysis

The 3′ terminal 1958 nucleotide sequence and the derived amino acid sequence of the Pakistani virus are shown in Fig. 1. The nucleotide sequence has a single open reading frame which begins at the 5′ end and continues for 1765 bases. The ORF is followed by a 3′ nontranslated region (NTR) of 192 bases and a polyadenylated terminus. The proteins encoded in this portion of the potyviral polypeptide correspond to the C-terminal portion of the nuclear inclusion protein (NIb), which is the putative viral replicase, and the capsid protein (CP).

Analysis of the derived amino acid sequence reveals the presence of the motif, S-3X-T-3X-NT-(18-37)X-GDD. This is the highly conserved, putative active site of the RNA-dependant RNA polymerase of positive-strand RNA viruses (Kamer and Argos, 1994). The virus lacks the motif, MVWCIENG, that is conserved in the core region of the capsid proteins of many potyviruses. This motif is also absent in WSMV and BrSMV, but is present in RgMV and SrMV-H and partially conserved in AgMV and HoMV (Salm *et al.,* 1996a). The new isolate also lacks the DAG motif in the coat protein

which is reported to be necessary for aphid transmission of potyviruses (Atreya *et al.,* 1990).

Efforts to sequence the N-terminus of the purified capsid protein were unsuccessful, suggesting that the terminus may be blocked. The N-terminus of the capsid protein of WSMV is also reported to be blocked (Niblett *et al.,* 1991). However, a potential proteolytic cleavage site between the NIb and the capsid protein VGEQ/G (Dougherty *et al.,* 1989; Gotz *et al.,* 1995) occurs at amino acid 307. This indicates a putative capsid protein of 281 amino acids (35.9 kDa), consistent with SDS–PAGE analysis (data not shown).

These results suggest that the Pakistani virus is related to the rymovirus WSMV. A Western blot revealed that antiserum raised to this sugar cane virus weakly crossreacts with WSMV capsid protein but not with AgMV capsid protein (data not shown). The sequence analysis, the lack of serological relationships between this new virus and other sugar cane- and cereal-infecting potyviruses, and the Western blot results suggest that the new virus is indeed a rymovirus related to WSMV. We have not been able to demonstrate transmission of this virus by the eriophyid mite *Aceria tosichella* vs. Keifer that transmits WSMV. It is possible that a different species of mite transmits this virus (Sithanantham *et al.,* 1972); however, we are unable to test this. To show the relationship with the rymovirus WSMV and to avoid potential confusion with the sugar cane mosaic group of potyviruses, we propose naming this virus sugar cane streak mosaic virus (SCSMV).

Multiple Sequence Alignments

Figure 2 is the alignment of the derived peptide sequences with the internal subsets (NIb and CP) underscored. Gap initiation and extension penalties used with PILEUP (Program manual for the Wisconsin Package, Version 8, 1994) were 2.95 and 0.177, respectively. The alignment is likely excessively stringent but tree topologies of alignments with substantially more gaps are still congruent. For space considerations, the DNA alignments are not shown but are available upon request.

Phylogenetic Analysis

The current taxonomy of the family Potyviridae is shown in Table 1, with the vectors associated with each genus. The bymoviruses have generally been used as the outgroup in previous phylogenetic studies and are placed “alone on the basal branch” (Ward *et al.,* 1995). Historically the overwhelming majority of available sequence data and subsequent phylogenetic analyses have been within the *Potyvirus* genus. Recent additions to the databases have enabled us to analyze the phylogeny of SCSMV and the entire potyvirus family more critically. In this analysis, the *Bymovirus* genus is represented by BYMV and WSSMV. MacMV and NLV are members in a probable new genus *Macluravirus* (Badge *et al.,* 1997). The *Rymovirus* representatives are WSMV, BrSMV, AgMV, HoMV, and RgMV and the *Ipomovirus* is SPMMV. Representatives of the genus *Potyvirus* are PPV, PVY, SrMV-H, and MDMV-A. SrMV-H and MDMV-A are potyviruses that infect corn and sorghum, as does SCSMV, and PVY is the “type” species of the potyvirus family.

Figure 3 depicts the phenogram derived by maximum likelihood analysis of the full-length peptide sequence alignment. The bootstrap, PUZZLE support, and decay indices are shown for each node. The parsimony analysis yielded a single tree with similar topology to the ML analysis. All of the different datasets (full length, NIb and coat protein, each as DNA and amino acid sequences) from each algorithm (parsimony and ML) produced trees that were congruent (data not shown). The only topological variation was that the node of PPV and PVY formed an unresolved trichotomy with MDMV-A and SrMV-H by parsimony analysis in the NIb peptide dataset. The branch support indices, together with the congruent analyses of the different datasets, indicate relatively strong support for the overall tree topology.

PTP analysis and RASA both revealed significant hierarchic signal in all manipulated data sets tested (inclusion/exclusion of various ingroup/outgroup taxa and rooted vs unrooted constraint trees). These analyses suggested that the NIb portion of the sequence was the most signal rich of the three regions tested (full 2 kb, NIb, coat protein). Differences in signal content between rooted and unrooted RASA also suggested possible outgroup convergence. To ensure that this was not problematic, alignments were constructed that excluded the bymoviruses and rooted instead with the macluraviruses. These alignments required fewer gaps but still resulted in identical topological arrangements to those produced by alignments that included the bymoviruses (data not shown).

These data confirm the relationships of the macluraviruses and the ipomovirus within the family Potyviridae (Badge *et al.,* 1997; Colinet *et al.,* 1996). In this analysis the ipomovirus SPMMV is shown to be a sister taxon to SCSMV, WSMV, and BrSMV. The macluraviruses are sister taxa to the bymoviruses. Therefore both of these virus groups clearly should be included within the family Potyviridae.

The phylogenetic analysis also confirms that SCSMV is indeed related to the rymoviruses WSMV and BrSMV. Each dataset produced a topology promoting SCSMV as the sister taxon to these two viruses. It is also evident that the current rymovirus genus, defined by mite transmission, is actually comprised of two independant lineages. RgMV, HoMV, and AgMV are sister taxa to the aphid-transmitted potyviruses while the other rymoviruses WSMV, BrSMV, and now SCSMV are members of a clade that also includes the ipomovirus

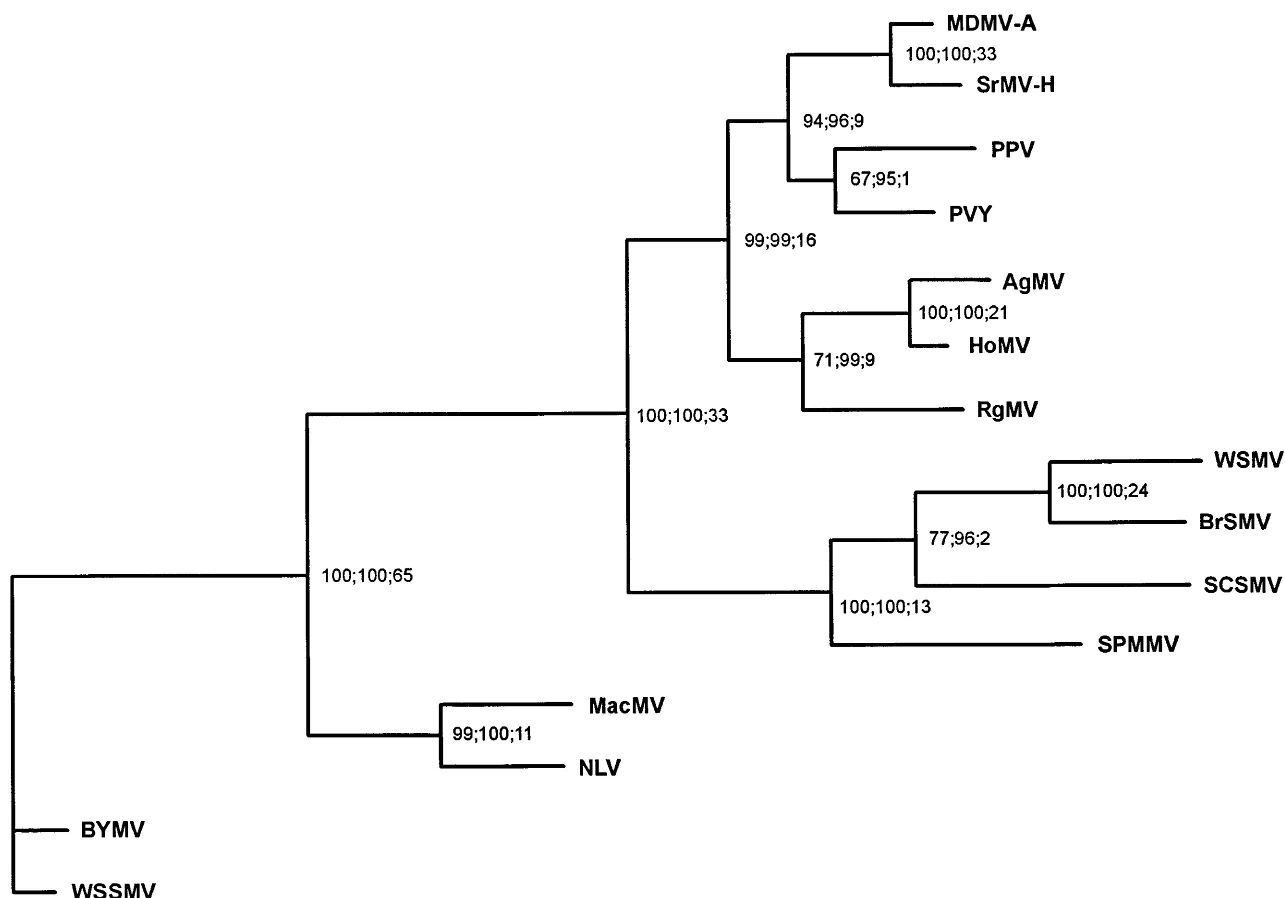

FIG. 3. Maximum likelihood phenogram of the full-length derived amino acid sequences representative of each genus in the family Potyviridae. Bootstrap, PUZZLE support values, and decay indices are shown at each node. Each dataset, analyzed by maximum likelihood and maximum parsimony methods, produced topologies congruent with this phenogram except the PVY-PPV node, which could not be unambiguously resolved in the NIb dataset by parsimony analysis.

SPMMV. Clearly AgMV is more closely related to the aphid-transmitted PVY than it is to WSMV, another mite-transmitted virus, and WSMV shares a more recent common ancester with the whitefly-transmitted SPMMV than with AgMV. These data confirm the previous results of Salm *et al.* (1996b) and show that the current taxonomic arrangement of the mite-vectored *Rymovirus* genus is paraphyletic.

Salm *et al.* (1996b) proposed that this group of mite-transmitted viruses (WSMV and BrSMV) should comprise a new genus *Whestrevirus,* and the addition of a third species (SCSMV) in this taxon would add support to their proposal. We agree, but feel that the name *Whesmovirus* is more appropriate for this genus given that these viruses historically were within the genus *Rymovirus* and to maintain the convention of using the mosaic symptom characteristic in the nomenclature (i.e., *Bymovirus, Ipomovirus, Rymovirus*). Therefore, SCSMV should be classified with WSMV and BrSMV in the genus *Whesmovirus* on the basis of its molecular and serological properties as well as the recovered phylogenetic relationships.

The evolution of specific vector relationships, i.e., mite transmission, has apparently occurred multiple times among these viruses. This has important biological implications. If two groups of viruses have the same vector yet are genetically distant from each other, there may be fundamental differences in their virus/vector/host relationships. They may be transmitted by different mechanisms and interact with different host factors during infection and replication. Plant cultivars resistant to one group of viruses may not be resistant to a different group vectored by the same agent. These are important and interesting matters for plant breeders and control strategies and warrant further study.

In the past, the vector-based taxonomy of the Potyviridae was thought to correlate well with sequence data and phylogenetic analyses (Ward *et al.,* 1995). However, questions have previously been raised concerning how these viruses are distinguished at the genus level (Zettler, 1992). In the Sixth Report of the International Committee of Taxonomy of Viruses the editors state "in the future genera will not stand where evidence is obtained of distinct phylogenies among

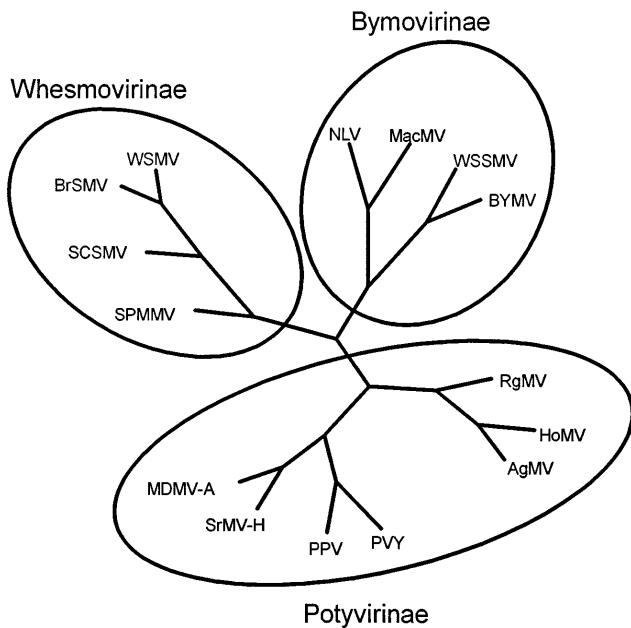

FIG. 4. An unrooted tree showing the three subfamily lineages within the family Potyviridae. Subfamily Potyvirinae contains the genera *Potyvirus* (MDMV-A, SrMV-H, PPV, PVY) and *Rymovirus* (RgMV, HoMV, AgMV). Subfamily Bymovirinae consists of the genera *Bymovirus* (BYMV, WSSMV) and *Macluravirus* (MacMV, NLV). The Whesmovirinae is the *Whesmovirus* genus which includes WSMV, BrSMV, SCSMV, and the ipomovirus SPMMV.

member species" (Murphy *et al.,* 1995). Using the new sequence data available, our analysis indicates that the use of vector species as a primary taxonomic character is misleading and obscures rather than represents the evolutionary history of the Potyviridae. Therefore, we propose restructuring the taxonomy of the Potyviridae family to more accurately represent evolutionary history.

It is clear from the phylogenetic reconstruction that members of the Potyviridae can be placed into three distinct lineages (bymoviruses and macluraviruses, impomovirus and whesmoviruses, potyviruses and rymoviruses). Therefore, we propose dividing the Potyviridae into subfamilies that correspond to these lineages (Fig. 4). Subfamily level of classification is used in the families Poxviridae, Herpesviridae, Parvoviridae, and Paramyxoviridae, and we feel this is also appropriate for a family as large and diverse as the Potyviridae. The subfamily Potyvirinae would be comprised of the genera *Potyvirus* (MDMV-A, SrMV-H, PVY, PPV) and *Rymovirus* (RgMV, HoMV, AgMV). Subfamily Bymovirinae would contain the genera *Bymovirus* and *Macluravirus.* Subfamily Whesmovirinae would consist of the new genus *Whesmovirus* (WSMV, BrSMV, SCSMV, and the current *Ipomovirus* SPMMV). By creating a subfamily level of taxonomic hierarchy and separating the paraphyletic groupings into different genera, i.e., *Whesmovirus* and *Rymovirus,* the taxa are grouped monophyletically according to their inferred histories. It must be noted that in this classification, vector type remains consistent with phylogeny at the genus level, except in the case of the ipomovirus SPMMV, which we have included in the *Whesmovirus* genus.

It has been reported that several virus families apparently have evolved congruently with their hosts and vectors. Codivergent evolution has been proposed for bunyaviruses, poxviruses, tymoviruses, and tobamoviruses (Eldridge, 1990; Ward *et al.,* 1995; Lartey *et al.,* 1996). There also has been speculation that the potyviruses coevolved with both their hosts and vectors (Ward *et al.,* 1995). This scenario is not supported by the phylogeny as inferred in this study. Instead, the data show no clear relationships of potyviral lineages with either host or vector type. All of the proposed subfamilies contain viruses that infect monocotyledonous and dicotyledonous hosts and have more than one type of vector. If fungal transmission is assumed to be ancestral then there have been at least four abrupt changes of vector type during the evolution of this family. Aphid, whitefly, and in two independent lineages, mite transmissability have been acquired. Similar scenarios can be derived by assuming that the ancestral virus was transmitted by one of the other vector types, thus supporting the contention that these viruses have switched vectors several times during their evolution.

It is important to keep in mind that viruses, particularly RNA viruses, have the capacity to evolve rapidly. The quasispecies population structure (Domingo *et al.,* 1996), gene reassortment and duplication, modular evolution and recombination, and insertion of host or vector sequences into the genome are all potential factors in the evolution of the potyviruses that could complicate phylogenetic interpretation. Thus it seems possible that recombination events may have occurred, perhaps with other virus families, providing a mechanism for such drastic changes in vector transmission. For instance, it appears that the bymoviruses may have gained fungal transmission by recruitment of an entire set of genes encoded as a separate genomic RNA (Dessens and Meyer, 1996). As new DNA sequence data become available, it may be possible to determine exactly where the changes relating to vector switching have occurred and the mechanisms involved in the evolution of these viruses. In any event, it is evident that the phylogenetic relationships and the evolutionary history of this important virus family are more complex than have been previously recognized.

ACKNOWLEDGEMENTS

The authors thank Drs. Michael V. Graves, T. Jack Morris, and Tom Powers for reviewing the manuscript. The patience and technical help of Tim Harris and the encouragement of Dr. Ellen Ball were critical in this research and are greatly appreciated.

Note added in proof. The International Committee on the Taxonomy of Viruses is currently reevaluating the taxonomy of the genus *Rymovirus* (P. Berger, personal communication).

REFERENCES

Archie, J. W. (1989). A randomization test for phylogenetic information in systematic data. *Syst. Zool.* **38:** 239–252.

Atreya, C. D., Raccah, B., and Pirone, T. P. (1990). A point mutation in the coat protein abolishes aphid transmissability of a potyvirus. *Virology* **178:** 161–165.

Badge, J., Robinson, D. J., Brunt, A. A., and Foster, G. D. (1997). 3′-terminal sequences of the RNA genomes of narcissus latent and maclura mosaic viruses suggest that they represent a new genus of the *Potyviridae. J. Gen. Virol.* **78:** 253–257.

Brakke, M. K., Skopp, R. N., and Lane, L. C. (1990). Degradation of wheat streak mosaic virus capsid protein during leaf senescence. *Phytopathology* **80:** 1401–1406.

Bremer, K. (1988). The limits of amino acid sequence data in Angiosperm phylogenetic reconstruction. *Evolution* **42:** 795–803.

Bremer, K. (1994). Branch support and tree stability. *Cladistics* **10:** 295–304.

Carpenter, J. M., Goloboff, P. A., and Farris, J. S. (1998). PTP is meaningless, T-PTP is contradictory: A reply to Trueman. *Cladistics* **14:** 105–116.

Colinet, D., Kummert, J., and Lepoivre, P. (1996). Molecular evidence that the whitefly-transmitted sweetpotato mild mottle virus belongs to a distinct genus of the *Potyviridae. Arch. Virol.* **141:** 125–135.

Dayhoff, M. O., Schwartz, R. M., and Orcutt, B. C. (1978). A model of evolutionary change in proteins. *In* Atlas of Protein Sequence Structure" (M. O. Dayhoff, Ed.), Vol. 5, Suppl. 3, pp. 345–352. National Biomedical Research Foundation, Washington, DC.

Dessens, J. T., and Meyer, M. (1996). Identification of structural similarities between putative transmission proteins of *Polymyxa* and *Spongospora* transmitted bymoviruses and furoviruses. *Virus Genes* **12:** 95–99.

Domingo, E., Escarmís, C., Sevilla, N., Moya, A., Elena, S., Quer, J., Novella, I., and Holland, J. J. (1996). Basic concepts in RNA virus evolution. *FASEB J.* **10:** 859–864.

Dougherty, W. G., Cary, S. M., and Parks, T. D. (1989). Molecular genetic analysis of a plant virus polyprotein cleavage site: A model. *Virology* **171:** 356–364.

Eldridge, B. (1990). Evolutionary relationships among California serogroup viruses (Bunyaviridae) and *Aedes* mosquitoes (Diptera: Culicidae). *J. Med. Entomol.* **27:** 738–749.

Eriksson, T. (1996). AutoDecay. Distributed by the author.

Faith, D. D., and Cranston, P. S. (1991). Could a cladogram this short have arisen by chance alone?: On permutation tests for cladistic structure. *Cladistics* **7:** 1–28.

Gillaspie, A. G., Jr., Mock, R. G., and Smith, F. F. (1978). Identification of sugarcane mosaic virus and characterization of strains of the virus from Pakistan, Iran, and Camaroon. *Proc. Int. Soc. Sugar Cane Technol.* **16:** 347–355.

Gillaspie, A. G., Jr., and Mock, R. G. (1984). Sugarcane mosaic virus. A survey of strains from sugarcane with a set of sorghum inbred lines. *Sugar Cane* **2:** 1–3.

Gotz, R., and Maiss, E. (1995). The complete nucleotide sequence and genome organization of the mite-transmitted brome streak mosaic rymovirus in comparison with those of potyviruses. *J. Gen. Virol.* **76:** 2035–2042.

Hasegawa, M. H., Kishino, H., and Yano, K. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22:** 160–174.

Hennig, W. (1966). "Phylogenetic Systematics" (Davis, D. D. and Zangeil, R., Trans.). Univ. of Illinois Press, Urbana.

Hitachi Software Engineering Co. Ltd. (1991, 1992).

Jensen, S. G., and Hall, J. S. (1993a). Characterization of a krish-sorghum infecting potyvirus. (Abs.) *Sorghum Newslett.* **34:** 17.

Jensen, S. G., and Hall, J. S. (1993b). Identification of a new sorghum infecting species of potyvirus from sugarcane. (Abs.). *Phytopathology* **83:** 884.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *CABIOS* **8:** 275–282.

Källersjö, M., Farris, J. S., Kluge, A. G., and Bult, C. (1992). Skewness and permutation. *Cladistics* **8:** 275–287.

Kamer, G., and Agros, P. (1984). Primary structural comparison of RNA-dependent RNA polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* **12:** 7269–7282.

Lane, L. C. (1986). Propagation and purification of RNA plant viruses. *Methods Enzymol.* **118:** 687–696.

Lartey, R. T., Voss, T. C., and Melcher, U. (1996). Tobamovirus evolution: Gene overlaps, recombination, and taxonomic implications. *Mol. Bio. Evol.* **13:** 1327–1338.

Lyons-Weiler, J., Hoelzer, G. A., and Tausch, R. J. (1996). Relative apparent synapomorphy analysis (RASA) I: The statistical measurement of phylogenetic signal. *Mol. Bio. Evol.* **13:** 749–757.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D., (Eds. 1995). Virus taxonomy-classification and nomenclature of viruses. *Arch. Virol.* (**Suppl. 10**): 1–11. [Sixth Report of the International Committee on Taxonomy of Viruses].

Niblett, C. L., Zagula, K. R., Calvert, L. A., Kendall, T. L., Stark, D. M., Smith, C. E., Beachy, R. N., and Lommel, S. A. (1991). cDNA cloning and nucleotide sequence of the wheat streak mosaic capsid protein gene. *J. Gen. Virol.* **72:** 499–504.

Program manual for the Wisconsin Package, Version 8, September 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711.

Salm, S. N., Rey, M. E. C., Robertson, N. L., French, R., Rabenstein, F., and Schubert, J. (1996a). Molecular cloning and nucleotide sequence of the partial genomes of Agropyron and Hordeum mosaic viruses, two members of the *Rymovirus* genus in the taxonomic family *Potyviridae. Arch. Virol.* **141:** 2115–2127.

Salm, S. N., Rey, M. E. C., and Rybicki, E. P. (1996b). Phylogenetic justification for splitting the *Rymovirus* genus of the taxonomic family *Potyviridae. Arch. Virol.* **141:** 2237–2242.

Sanger, F., Nicklen, S., and Coulsen, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463–5467.

Schöniger, M., and von Haeseler, A. (1994). A stochastic model for the evolution of autocorrelated DNA sequences. *Mol. Phyl. Evol.* **3:** 240–247.

Shukla, D. D., Ward, C. W., and Brunt, A. A. (1994). "The Potyviridae," CAB International, Wallingford, UK.

Sithanantham, S., Muthuswamy, S., and Bhaskaran, T. L. (1972). A note on the role of the eriophyid mite *Aceria sacchari* Channabasavanna, (Eriophyoidea: acarina) in the transmission of streak virus disease of sugarcane. *Sci. Culture* **38:** 248–250.

Strimmer, K., and von Haeseler, A. (1996). Quartet puzzling: A quartet maximum likelihood method for reconstructuring tree topologies. *Mol. Biol. Evol.* **13:** 964–969.

Swofford, D. L. (1993). Phylogenetic analysis using parsimony (PAUP), version 3.1.1 University of Illinois, Champaign.

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10:** 512–526.

Ward, C. W., and Shukla, D. D. (1991). Taxonomy of Potyviruses: Current problems and some solutions. *Intervirology* **32:** 269–296.

Ward, C. W., McKern, N. M., Frenkel, M. J., and Shukla, D. D. (1992). Sequence data as the major criterion for potyvirus classification. *Arch. Virol.* (volno**Suppl. 5**): 283–297.

Ward, C. W., Weiller, G. F., Shukla, D. D., and Gibbs, A. J. (1995). Molecular systematics of the Potyviridae, the largest plant virus family. *In* "Molecular Basis of Virus Evolution" (A. Gibbs, C. Calisher, and F. García-Arenal, Eds.), pp. 477–500, Cambridge Univ. Press, Cambridge.

Zerbini, F. M., Koike, S. T., and Gilbertson, R. L. (1995). Biological and molecular characterization of lettuce mosaic potyvirus isolates from the Salinas Valley of California. *Phytopathology* **85:** 746–752.

Zettler, F. W. (1992). Designation of potyvirus genera: A question of perspective and timing. *Arch. Virol.* (volno**Suppl. 5**): 235–237.