

Cell-to-cell movement of plant viruses

Insights from amino acid sequence comparisons of movement proteins and from analogies with cellular transport systems

Brief Review

A. R. Mushegian¹ and E. V. Koonin²

¹Department of Plant Pathology, University of Kentucky, Lexington, Kentucky

²National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, U.S.A.

Accepted June 21, 1993

Summary. Cell-to-cell movement is a crucial step in plant virus infection. In many viruses, the movement function is secured by specific virus-encoded proteins. Amino acid sequence comparisons of these proteins revealed a vast superfamily containing a conserved sequence motif that may comprise a hydrophobic interaction domain. This superfamily combines proteins of viruses belonging to all principal groups of positive-strand RNA viruses, as well as single-stranded DNA containing geminiviruses, double-stranded DNA-containing pararetroviruses (caulimoviruses and badnaviruses), and tospoviruses that have negative-strand RNA genomes with two ambisense segments. In several groups of positive-strand RNA viruses, the movement function is provided by the proteins encoded by the so-called triple gene block including two putative small membrane-associated proteins and a putative RNA helicase. A distinct type of movement proteins with very high content of proline is found in tymoviruses. It is concluded that classification of movement proteins based on comparison of their amino acid sequences does not correlate with the type of genome nucleic acid or with grouping of viruses based on phylogenetic analysis of replicative proteins or with the virus host range. Recombination between unrelated or distantly related viruses could have played a major role in the evolution of the movement function. Limited sequence similarities were observed between i) movement proteins of dianthoviruses and the MIP family of cellular integral membrane proteins, and ii) between movement proteins of bromoviruses and cucumoviruses and M1 protein of influenza viruses which is involved in nuclear export of viral ribonucleoproteins. It is hypothesized that all movement proteins

of plant viruses may mediate hydrophobic interactions between viral and cellular macromolecules.

Introduction

It is generally accepted that plant viruses exploit plasmodesmata to move from initially infected cells, which are usually negligible in amount, into neighbouring healthy cells. Without cell-to-cell movement productive virus infection is not established. Also it is known that specific movement proteins encoded by virus genomes are essential for cell-to-cell spread. Frequently, a plant virus protein is referred to as movement protein if i) it is not a capsid protein and ii) disruption of the coding sequence of this protein abolishes infection in whole plants but has no effect on virus replication in protoplasts. Putative movement proteins have been identified in this manner in about 15 different plant virus groups (see [4, 19, 40] for recent reviews). In some cases, additional virus proteins are also required for movement, e.g. coat protein [13, 52, 58] or a specific segment in a replication protein [55], but these phenomena are not universal. On the other hand, a specific movement protein or at least movement domain is thought to be present in all viruses that are able to spread from cell to cell [4].

Apart from this genetic evidence for the movement function, the current knowledge of mechanisms which facilitate intercellular spread of plant virus genomes is insufficient. Recently certain properties of some movement proteins, notably of the 30 kDa movement protein of tobacco mosaic virus (TMV), were characterized *in vitro* and *in vivo* and models for movement of plant viruses have been prompted by these studies [16, 19]. Additionally, movement proteins were subjects of comparative sequence analysis, which allowed to delineate common motifs in movement proteins from otherwise unrelated groups implying existence of common features among many movement proteins [33, 43].

In this work, we sought to critically analyze currently available biochemical and ultrastructural data on movement proteins. We then investigated whether sequences of movement proteins could provide information on their function based on similarities with other viral or cellular proteins.

What do we know about movement protein functions?

Modification of plasmodesmatal size exclusion limit

Fluorescent probes injected into a plant cell are either transported into neighbouring cells via plasmodesmata or retained in the initially injected cell. The crucial determinant for the exclusion of a particle from plasmodesmatal movement is its hydrodynamic (Stokes) radius [51]. Studies of transgenic tobacco plants constitutively expressing TMV 30 kDa movement protein (MP + plants) revealed that certain molecules, which were too large to be transported in wild type plants, gained such an ability in MP + tobaccos. Specifically, FITC-labelled dextran (M_r 9.4 kDa) moved from injected to neighbouring cells only in plants transformed with and expressing TMV 30 kDa protein but not in control MP-

plants [57]. Interestingly, plasmodesmata in MP+ and MP- plants appeared to be structurally indistinguishable [19, 57].

These data indicate that movement protein of TMV functionally modifies plasmodesmata. It is not understood yet whether these modifications and the changes required for moving virus genome through plasmodesmata are the same or different. Virus nucleic acid would not be expected to exist in naked form in the cell; rather, it seems reasonable that, like all nucleic acids in eukaryotic cells, it is associated with cellular and/or virus proteins throughout its lifetime. While Stokes radius of a 10 kDa dextran molecule is estimated to be about 3.1 nm [20, 57], ribonucleoproteins (RNPs) are clearly larger, many of them having sizes within the range of 10–50 nm [41], which would exclude them from movement even in MP+ plants described in [57].

Recently, plasmodesmatal permeability in *Nicotiana clevelandii* leaves infected with tobacco rattle virus (TRV) has been investigated [20]. Movement protein of TRV, the 29 kDa protein, shares high sequence similarity with TMV movement protein, and both proteins are considered to be functionally very similar. Simultaneous injection of virus and fluorescent tracers into cells of leaf trichomes allowed to observe movement of the labels under the conditions when the virus itself was known to move in the same cells. It has been shown that FITC-labelled dextran (M_r 4.4 kDa) moved from cell to cell only in TRV-infected cells and synchronously with the virus. However, Lucifer Yellow-labelled dextran (M_r 10 kDa) was excluded from plasmodesmatal movement even in infected cells. Thus, certain molecules with the dimensions resembling the smallest of the known RNPs did not move under conditions when virus RNPs readily moved. Apparently, either an extremely thin virus-specific RNP should be postulated (see below), or it might be concluded that slight increase in plasmodesmatal permeability induced by movement proteins hints at some important circumstances but is not sufficient to explain cell-to-cell movement of the virus genome.

Binding of movement proteins to single-stranded nucleic acids

TMV 30 kD movement protein has been overexpressed in *E. coli*, and purified protein has been shown to bind single-stranded (ss) DNAs and RNAs non-specifically and cooperatively [14]. A model has been proposed, in which movement protein binds genomic RNA stoichiometrically to unfold and shape it into a thin complex [14, 16, 17]. Electron microscopy of such complexes formed in vitro revealed thin structures at the limit of microscope resolution; it was concluded that, as it was barely seen under EM, the complex was thin enough to fit in plasmodesmata with size exclusion limit of 3 nm [17]. Additionally, ss nucleic acid-binding properties have been reported for movement proteins of cauliflower mosaic caulimovirus (CaMV) [15], red clover necrotic mosaic dianthovirus (RCNMV) ([47], D. Cookmeyer, pers. comm.) and alfalfa mosaic virus (AIMV) [40]. Interestingly, P1, the movement protein of cauliflower mosaic virus, has been shown to preferentially bind RNA rather than DNA

[15]; it was speculated that 35 S transcript, a genome length virus replication intermediate, might be the form which is transported from cell to cell.

These *in vitro* observations offer new insights into a mechanism of virus cell-to-cell movement. However, it should not be forgotten that virus RNAs *in vivo* are associated with proteins, as revealed for TMV-specific RNAs [22]; it is thought that such an association is maintained throughout the whole life of the RNPs in the cell; a mechanism of stripping virus RNA from these proteins has not yet been proposed. Also, as binding of movement proteins to RNA is thought not to be sequence-specific [14, 17], it could be anticipated that the majority of movement protein will bind to cellular RNAs, which in the case of TMV infection are in excess over virus RNAs at the time when 30 kD protein is transiently expressed [18]. The “ss nucleic acid-binding” model, therefore, has to be modified to deal with these difficulties.

Intracellular localization of movement proteins – are there two morphologically distinct pathways for virus movement?

Biochemical fractionation of infected cells revealed that in many virus-plant interactions movement proteins are found in cell wall-enriched fraction. Immunogold labelling has confirmed these observations and demonstrated localization of TMV, A1MV, CaMV, and RCMV movement proteins in plasmodesmatal channels or in cell wall near plasmodesmata [36, 40].

Movement protein of cowpea mosaic comovirus (CPMV), the 58 kD/48 kDa protein, has been found to form tubular structures perpendicular to and intruding into plasma membrane and cell wall. These tubules have been found both in whole leaves and in protoplasts [56]. Virions have been observed inside the tubules [56]. It has been speculated that the tubular inclusions are the structures required for movement.

Cucumber mosaic virus (CMV) movement requires the 32 kDa movement protein [7]. This protein has not been found to associate with the cell wall or from tubules; instead, it was located in nucleoli of infected cells [37].

Observations on movement proteins of TMV and CPMV have given rise to the idea of two different types of virus movement, “tobamo-type” and “como-type”. It has been proposed that the tobamo-type is characterized by movement protein localization in the plasmodesmatal area of the cell wall and by the virus infectious entity capable of cell-to-cell movement in the absence of the viral capsid protein. Como-type of movement was thought to require both movement protein and coat protein and to involve formation of tubular structures protruding from cell wall into cytoplasm and formed by the movement protein [19, 56]. Complementation data argued, however, that there is certain compatibility among both types of movement as tobamoviruses were able to complement comovirus movement [4].

When all available ultrastructural data are considered, the picture becomes not so clear-cut. It should be expected that RCMV, a comovirus closely related to CPMV, moves according to the como-type; instead, movement protein of

RCMV was found exclusively in plasmodesmata of the infected plants [40] which had been proposed to be a tobamo-type feature. CaMV movement protein is found in cell wall fraction and, more specifically, in plasmodesmata [2, 36]; this seemingly supports the attribution of caulimovirus movement to the tobamo-type. However, tubular structures of unknown composition have been observed in the cytoplasm of CaMV-infected cells [36] and recently it has been shown that P1, the movement protein of CaMV, forms tubular structures on the surface of inoculated protoplasts [49]. These findings seem to be more consistent with the como-type action of the caulimovirus movement protein.

The above discussion shows that it may be premature to draw too sharp a distinction between different types of movement based on ultrastructural observations. It also should be remembered that we see movement proteins at the sites where they are most easily observed and/or are retained for longer periods, but not necessarily at the sites of their essential activity.

Mapping of movement protein functions by reverse genetic approaches

With full-length DNA copies of numerous virus genomes available, attempts have been made to genetically dissect movement protein coding sequences. In most cases, lengthy deletions have been introduced, and effect of these deletions on the overall infectivity or individual properties of the movement proteins has been investigated. It has been shown that even small changes near the N-terminus of TMV 30 kD movement protein abolish virus infectivity [6]; at the C-terminus, up to 33 of the 268 amino acid residues of this protein could be deleted without apparent effect on infectivity or cell wall localization of the movement protein. When 55 C-terminal amino acids were deleted, virus moved slowly though movement protein was still found in plasmodesmata [6]. Deletion of 74 C-terminal residues (195 to 268) was lethal, and, at the same time, movement protein was no longer found in plasmodesmata. It has been noted that essential residues from 195 to 214 could comprise a domain with a specific function, or their deletion might simply cause misfolding of the protein [6].

In another study, C-terminal and internal deletions have been introduced into TMV movement protein to determine location of ss nucleic acid-binding site(s) [17]. Initially, it was thought that residues 65 to 86 constitute RNA-binding domain [14]. Later, this domain has been shown to rather decrease protein solubility [17]. Up to 84 amino acids from the C-terminus of TMV movement protein could be eliminated without effect on ss nucleic acid binding capacity of the protein [17]. In the remaining part, deletion of amino acids 111–125 abolished ss nucleic acid binding, but when the gene segment encoding these amino acids was fused to an unrelated sequence, the resulting himeric protein did not bind to ss nucleic acids [17].

Movement protein of AIMV associates with cell wall upon infection and in transgenic tobacco plants; however, this did not happen when N-proximal amino acids (12–77) have been deleted [24].

Strains of TMV have been characterized that are temperature sensitive in

cell-to-cell movement or break the host resistance which is based on plant genes restricting virus movement. Involvement of mutations in 30 kD movement protein cistron could be implied, and, indeed, in many cases nucleotide changes responsible for the altered phenotype were shown to be confined to the movement protein gene [11]. Many of these mutations were scattered around the middle one-third of 30 kD movement protein and resulted in a change in polarity of the encoded amino acid [11]. As no known function could be mapped to the area, these data are difficult to interpret.

Classification of movement proteins based on computer-assisted analysis of amino acid sequence similarities

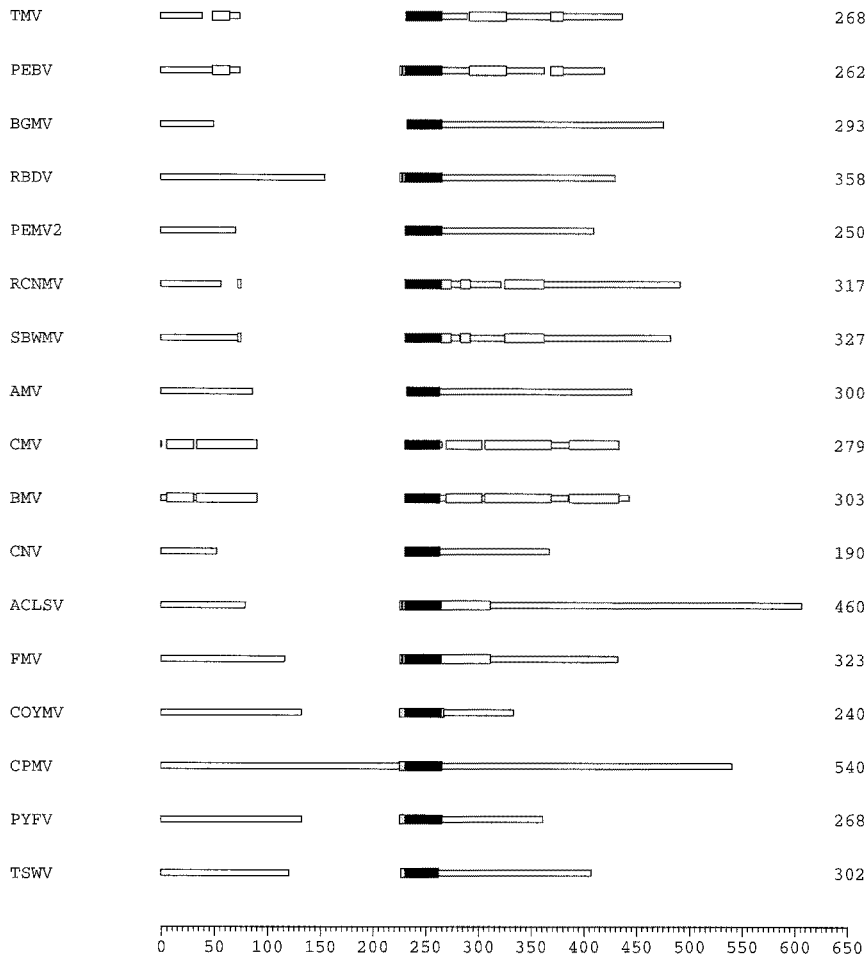
With representative sequences of plant virus genomes from many groups available, efforts were made to identify regions of similarity between known movement proteins and to predict movement functions for uncharacterized ORFs. Early observations made by this approach have been reviewed previously (e.g. [4]) and included strong similarity between tobamovirus and tobavirus movement proteins [29] as well as less pronounced but significant local similarity between segments of tobamovirus and caulimovirus movement proteins [30]. Later, similarities between movement proteins of tricornaviruses and dianthoviruses were observed [59]. In fact, the movement function for the respective proteins of tobaviruses, caulimoviruses and dianthoviruses has been implied by these observations and only subsequently confirmed by site-directed mutagenesis [54, 59, 60].

More detailed analyses performed by Melcher [43] and by ourselves [33] revealed longer (ca. 150 amino acids) segments of sequence similarity between movement proteins from different virus groups. Interestingly, one family of movement proteins included proteins encoded by viruses with positive-strand RNA genomes, single-strand DNA geminiviruses and by pararetroviruses with virion DNA, namely caulimoviruses [33].

In the time past after the publication of these analyses, further significant increase in the number of sequenced virus genomes was achieved and it has been argued that the current collection of such sequences may represent the majority of the existing virus groups [32, 34]. With this in mind, we undertook a systematic comparison of the sequences of the movement proteins with each other and with the current sequence databases.

“30 K superfamily”

Using the BLAST program [3], statistically highly significant similarities were observed only between proteins of viruses belonging to a single group or two closely related groups. Further comparisons were done using programs MA-CAW [53a] and OPTAL [27] to generate multiple alignments according to the progressive alignment strategy [21], i.e. starting with pairs of most closely related sequences and proceeding to incorporate more and more distant ones.



b

TMV	tobamovirus	66	VCLAGLVVTGEWNLPDNCRGGVSVCLVDKRMER	169	P03583
PEBV	tobravirus	81	LRIVGIQIKVTSHLPRDTSQFLQIDLIDSRLTD	148	X14006G
BGMV	geminivirus (2)	50	MKIDHCVIEYRQQVPINATGSSVVVEIHDKRMTD	210	P06001
RBDV	idaevirus	161	IHLNRLLLI AVLPAVGKGTPTGARIKIRDARLDD	164	S55890G
PEMV2	separate group	73	LSVEKVVVFTPHVPDDVLGEVEIWLHDSILPH	144	S53233G
SBWMV	furovirus	78	MSYSCVLLNVIPTVPMGYAGTVEVSLDLSGLSP	215	L07937G
RCNMV	dianthovirus	62	LNETHVVL SLAPTIGVAIPGHVTVELINPNVEG	222	P02595
AMV	ilarvirus	87	VHSGAIYLLYMPNLIKSSSGSITLKL FNEATGE	180	P10838
CMV	cucumovirus	85	MSVPQVLCVTRTVSTDAEGSLKIYLADLGDK	161	P03604
BMV	bromovirus	93	MNVPRIVCFLVRTTDSAESG SITVSLCDSGKAA	177	P03603
CNV	tombusvirus	55	ITVRRVRIKIIFLVSRKAGVSGKLYLRDISDTT	102	P15185
ACLSV	capillovirus	86	VHWGALSI SIDALFRKNAGVSGWCYVYDNRWET	341	P27739
CPMV	comovirus	232	IHTGAIETIIQSFASPESDL MGGFLLVDSLHTD	274	P03599
PYFV	separate group	139	MAIDSIRVAVTSLMHQGD SRECIMYLCDRRFKD	?	D14066G
FMV	caulimovirus	123	IHLGAVKILLTAQFRQGD TSVKMALIDRIVN	147	P09520
COYMV	badnavirus	139	IHTGVMLVRIQILHRKFAGT MALIVFRDTRWSD	?	X52938G
TSWV	tospovirus	126	MMISRIVIWVCTTIP-NPTGKLVVALIDPNMPS	132	S48091G
	consensus		&xx&xxU&U&xxx&xxxx&GSU&&&&&Dxx&xx		

T

Fig. 1. The conserved sequence motif in a vast superfamily of plant virus movement proteins. **a** Schematic representation of sequence conservation in movement proteins. The black box shows the principal conserved motif; additional regions of sequence conservation in related proteins are shown by open boxes. The scheme was generated by the MACAW program. **b** The principal conserved motif in a representative set of movement proteins. The motif was extracted from the alignment generated by the MACAW program. One sequence was included from each group of viruses. The consensus (for symbols see legend to Fig. 2) includes amino acid residues conserved in at least 9 out of the 17 aligned sequences. SWISSPROT or where not available, Gen Bank (G) accession number accompanies each sequence

```

consensus          SxUYxDxSxFxxxxxxxx&xx@xxxxxxxxxxxxxxxxGxxUSxx&x
                    G   N T   A
ACLSV 50K (25-151)  SRIYSDISPFKKASDLMINEFVFK-VMPEDIAGDGFRLA-SIP
PVT 34K (16-148)  GAIYTN-SAFHNLKT-KWHVYKSECSIGLDLDPNGKIIISKDIP
ASGV 38K (26-150)  SELYKDATFFKPDVLCIKRFESN-VKVSSRSGDGLVLS-DFK

U&xxxEUxxUxxxxxxxx&&H&G&x&x&xUx&x&x&xxxxUxGxxxU&Dxxxxxx
ACLSV 50K  VIPSSEVQAVLRKRESTNYVHWGALSISIDALF--NAGVSGWCYVYDNRWETF
PVT 34K  LFDQBEIDNIMK--DDKQVPHLGAFFVGLVAHFVDEEVEGLVSIIDKRRRTDL
ASGV 38K  LLDDTEIDSIRKKS NKYKYLHYGVILVGIKAMPLNFRMGMEGRVIVYDACLDP-

xxAxUxxxx&x&xxxx&xxxPxxxUxxx
G
ACLSV 50K  EQAMLQKFHFNLDSGSATLVTSPNFVSLDD
PVT 34K  RRATLACRRIKVFNGRCAFMMKPNFVSRKED NA
ASGV 38K  KRGHICSYLKFESDCCYFGLRPEHCLSTTD S47260G

consensus          xNR&ExxKxxxxxxxxUxS&xxxx&xKxxxxxxxx
                    K D   A
RBDV 39K (127-342) FNKFKFKYKGNNEFSLVEASV-----FDKLR-KDDS-P
RCNMV 35K ( 22-223) LNRyTDW-KCRSGVSEAPLI PASMMSKI TDYAKTFA-KGNSVA
SBWMV 37K ( 42-236) ENRYVE--KALTQPGVVTMMDAW---TNWTKTKKEEGTPYN

cons  Ux&xxUUUxUxPxUxxGxxGxxxUxUxDxxUxxxxUxxxxxxxxGxxxxCx&x
                    A N S
RBDV IHLNRLLI AVLPAVGGKTPGTARIKIRDARLDDGYGELFSSE---NRVDSGYIYICIN
RCNMV LNYTHVVL SLAPTIGVAIPGHVTVELINPNVEGPFQVMSGQT--LSWSPGAGKFCILM
SBWMV MSYSCVLLNVIPVPMGYAGTVEVSLDLSGLSPLENVI PDQTMMLLGKSPHVMCFF

cons  Ux@SUxxxxxxxxxxxxFxxUxxxxxxxxxxxxYS
                    A
RBDV VGYSPKSEIDYKINIDFAGVPIKDGKSPIWKAAFS
RCNMV I-FSVHH-QLN-SDHEPFR-VRI TNGI P--TKKSYA
SBWMV MHYSIPLNDKGRAVKLAFAK-IDAEMASKGMSVMNVYS

consensus          xxUxxExxLGxxxUxxDxxxxxxxxUPxxxxxxxxUxxxxxx
                    N
CMV 3A (33-209) KKMATECDLGRHHWMRADNAISVRPLVPQVTSNNLLPFFKSGYD
BMV 3A (41-224) KEIAABECKLGCYNLKSNEPRNYIDLVPKSHVSAWLSWATSKYD
PEMV 22K (24-197) KEL-EETNLGVITPVRANEKVITITPLLP-PKTQGRVSVLKRKRF-
CNV 21K (13-173) NELYKEATLG--NKLTVNVGMEDVE-VP-LLPSNFLT--KV---

cons  xxxxxxxG&UxVxxUxxx&xxxxxxxxxxxxUxLxUxxxxxxxxxxxxxxxxxxUxx
CMV AGELRSKGYMSVPQVLCVAVTRVSTDAEGSLKIYLDLGDKE---LSPIDGQCVTLHN
BMV KGELPSRGMNVPRIVCFVLRVTD SAESGSITVSLCDSGKAARAGVLEAIDNQEATIQL
PEMV RSTRNTGGLLSEKVVVVFVTPHVPDDVLEGEVEIWLHDSILPH----LGSVGR-LKCLK
CNV RVSLSG-GYITVRRVRIKIIPVSRKAGVSGKLYLRDISDT-----TGRKLHCTELL-

cons  xxxxxxUx&xxxx&xxxUxxxxGxxxxxxxxFx&xxxxxxxx&xxxxxxxxxxxxWxxx
                    S
CMV HELPALISFQ-PTYDCPM-ELVG---NRHRCFAVVVER-HGYIGYGGTASVCSNWQAO
BMV SALPALIALT-PSYDCPM-EVIGDSDGRNRCFGIATQL-SGVVGTGSAVATHAYWQAN
PEMV SEGPKLLAFY-PPYSIALGDSIS---GQPRFSIATELPEGNFAPGCS PFLFLMWSR
CNV -DLGKEIRLTMQHLDLFSV-SARS---DVPVIVGFEDLV-SPYLE-GRELF SVSLRWQ--

cons  &xxxxxxYxxxxxxxxUx&xxxxUx
CMV FSSKNNNYTHAAAGKTLVLPYNRLA
BMV FKAKPNNYKLGPAITIMVMPFDRLR
PEMV IEAVTHNY-LSRPPRALPICRTMVR
CNV FGLSAQCYSLPPARWKVMYQEDALK

consensus          xxxxxxxQxxxxxxxxxxxxxxxxxxxxxxxx&UxUxxUxUxxxxxxx
                    E
PYFV PPRO (110-246) EPLKDRAEITPTAASKLDKLTLEAHRKFNMAIDSIRVAVTSMLHQG
** * ** * * * * * * * * * * * * * * * * * * * * * * * * * * * *
FMV P1 ( 88-231) LPLITKQEIQ---QKLMKIDPSVRSKISMIHLGAVKILLTAQFRQG
SoCMV P1 ( 74-212) LATINKEQIT---KRISKIEBKQRRLIRYVHISTLQVLIKSTFLKG
CoYMV 194K (111-240) MMVVNGSQVD-RSFIQESSFEVLSRTGIEFHIGVMLVRIQILHRKF
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
TSWV NSm (102-230) LFGVNGKQNA---NKVIKICPTWDSRKQYMMISIRVIVWCPTIP-N

xxxxxxUx&xDxx&xxxxxxUxxxxxxxxxxxxxxxx@xxxx&x&xxxxxxxxxxxxxxxx&xx
PYFV DSRECI MYLCDRRFKDLLGALALIGFTLPGMQTHVYKTRMMAFSRKEAIAADRLQL-Y
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
FMV IDTSVKMALIDDRIVNRKDSLLGAARGNLA-YGKFMFTVYPKFALSLQSKNLDKTLFSFIH
SoCMV LDTPLELTLRDNRLNLEESKIAVGHGNLK-YGKMKFDVNLQGLSLKDLDDLDRSTILNY
CoYMV AGTMALIVFRDTRWSD-DRAVLAAMEIDLSEGNQIVY-VLPDIMMTIK--SFYRHIQI-C
: * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
TSWV PTGKLVVALIDPNMPS-GKQVILKGGQTTI--DPICFV FYL NWSIPKMNTPENCCQL-H

x&xxxxxxxxxxxxxxxx&xxxxxxxxxxxxxxxx
PYFV LVYKGALETRQN-TPITVNVRTSLIFGSNAE
: : * * * * * * * * * * * * * * * * * * * * * * * * * * * *
FMV QFER-KDLMKTGD-KVFTVTVLYGAL TNSHH
SoCMV KFLR-RNFMKEGN-HAFSISYRINYALSNHH P15631
CoYMV VMTKGYDQWQGED-NLLITRGLTG-RLSNT-S
:* * * * * * * * * * * * * * * * * * * * * * * * * * * * *
d TSWV LMCS-QEYKKGVSFGSMYSWTKEFGDSPRAD

```


It was demonstrated that movement proteins of viruses belonging to 17 groups contain a weak but confidently identified conserved motif consisting of approximately 30 amino acid residues (Fig. 1). This motif includes a nearly invariant aspartic acid residue (replaced by asparagine in dianthoviruses and ilarviruses) preceded by a region enriched in hydrophobic and nonpolar residues. In some of the movement proteins this region was predicted to form a transmembrane helix using the algorithm of Rao and Argos ([50], data not shown; see also discussion below). Remarkably, this conserved motif unites viruses with single-stranded DNA genomes (geminiviruses of group II), pararetroviruses with double-stranded DNA genomes (caulimoviruses and badnaviruses), negative-strand RNA viruses with ambisense genome segments (tospoviruses), and numerous groups of positive-strand RNA viruses. Among the latter, all the three major phylogenetically defined divisions [31, 34] are represented, namely rubi-like viruses (tobamoviruses, tobnaviruses, cucumoviruses, bromoviruses, ilarviruses, capilloviruses, furoviruses, raspberry bushy dwarf virus), picorna-like viruses (comoviruses, parsnip yellow fleck virus, pea enation mosaic virus), and flavi-like viruses (dianthoviruses and tombusviruses).

Several groups of movement proteins showed significant similarity on longer sequence stretches. In Fig. 2 we show the amino acid sequence alignments for four of these groups, delineation of which included identification of new putative movement proteins. Such new identifications are: i) 38 kDa protein of apple stem grooving virus, which is related to the putative movement proteins of the other capilloviruses, namely apple chlorotic leaf spot virus and potato virus T (Fig. 2 a); ii) 37 kDa protein of soil-borne wheat mosaic furovirus related to the movement proteins of dianthoviruses (Fig. 2 b); iii) 27 kDa protein of pea enation mosaic virus and 21 kDa protein of tombusviruses related to the movement proteins of bromoviruses and cucumoviruses (Fig. 2 c); and iv) NSm protein of

Fig. 2. Amino acid sequence alignments for groups of proteins including newly identified putative movement proteins. For each group of boundaries of the conserved segments were determined using the MACAW program and the alignment itself was constructed hierarchically using the OPTAL program. In the consensus *U* indicates a bulky aliphatic residue (*I, L, V, or M*), *a* indicates an aromatic residue (*F, Y, or W*) & indicates a bulky hydrophobic residue (aliphatic or aromatic) and *x* indicates any residue. The conserved motif of the “30 K superfamily” is highlighted by bold typing. **a** Putative movement proteins of capilloviruses. *ACLSV* Apple chlorotic leaf spot virus; *PVT* potato virus T; *ASGV* apple stem grooving virus; **b** movement protein of red clover necrotic mosaic dianthovirus (*RCNMV*), putative movement proteins of raspberry bushy dwarf virus (*RBDV*) and soil-borne wheat mosaic furovirus (*SBWMV*); **c** movement proteins of cucumber mosaic cucumovirus (*CMV*) and brome mosaic bromovirus (*BMV*), putative movement proteins of pea enation mosaic virus (*PEMV*) and cucumber necrosis tombusvirus (*CNV*); the similarity between 21 kDa proteins of tombusviruses and movement proteins of bromoviruses and cucumoviruses has been independently reported by U. Melcher (pers. comm.); **d** putative movement proteins of caulimoviruses, badnaviruses, parsnip yellow fleck virus, and tomato spotted wilt tospovirus; asterisks show identical residues and colons show similar residues between the sequences of *PYFV* and *FMV*, and *CoYMV* and *TSWV*

tospoviruses and the N-terminal domain of the parsnip yellow fleck virus polyprotein related to the putative movement proteins of caulimoviruses and badnaviruses (Fig. 2 d; the similarity between caulimovirus movement proteins and N-terminal domains of badnavirus polyproteins was noted by U. Melcher and cited in [8]). Even among these smaller groups, the latter three brought together proteins of very different viruses. Group iv) is particularly notable in that it compiles proteins of a positive-strand RNA virus, negative-strand (ambisense) RNA viruses, and DNA-containing pararetroviruses. Curiously, the array of domains, namely N-movement domain-capsid protein(s)-replicative domains-C in the polyproteins of such seemingly unrelated viruses as parsnip yellow fleck virus and the badnaviruses is very similar. A relevant observation from our previous work is the relationship between movement proteins of two component geminiviruses (BL1/BC1) with those of tobamoviruses and tobamoviruses [33].

All the proteins containing the conserved motifs shown in Fig. 1 should be considered a single vast, and highly diverged superfamily of plant virus movement proteins. We would like to coin the nickname “30 K superfamily”, after the most thoroughly studied movement protein of tobamoviruses, and also because most of the proteins of this superfamily have the size around this value (Fig. 1 a).

Other types of movement proteins

Several groups of positive-strand RNA viruses, namely potexviruses, carlaviruses, hordeiviruses, and a group including beet necrotic yellow vein virus, *Nicotiana velutina* mosaic virus, and peanut clump virus, encompass the triple gene block that consists of two small membrane proteins and a putative RNA helicase [34, 44]. It has been demonstrated that disruption of any of the triple block genes abolished the cell-to-cell movement of hordeiviruses, potexviruses and BNYVV [5 a, 25 a, 49 a]. The putative helicases in the triple block comprise a distinct group within superfamily I of DNA and RNA helicases and contain several unusual amino acid substitutions in the conserved helicase motifs [26, 34]. Unlike the genes coding for the proteins of the “30 K superfamily”, the triple block is found only in positive-strand RNA viruses. Carmoviruses and necroviruses encode a pair of small proteins, one of which is predicted to be an integral membrane protein. It has been shown that disruption of either of these genes precluded cell-to-cell movement of turnip crinkle carmovirus [28]. These two genes perhaps may be considered a “truncated triple block” although it was hard to demonstrate this at the level of sequence comparison due to the small size of the hydrophobic proteins.

A third type of movement protein has been revealed in tymoviruses. A 75 kDa proline-rich protein that is encoded in the 5' portion of the genomic RNA and overlaps with the replicative polyprotein gene is essential for the movement of TYMV in infected plants [10]. This protein may have specific secondary and tertiary structure similar to that of other proline-rich filamentous proteins.

Interestingly, upon screening of the amino acid sequence databases with the sequence of the tymovirus movement protein the highest (albeit moderate) similarity was observed with the plant cell wall-associated protein extensin (data not shown).

Recently it has been shown that cell-to-cell movement of geminivirus MSV is dependent on the small protein V1 [9]. This protein and the related proteins of other one component geminiviruses did not show appreciable sequence similarity to movement proteins of other viruses. Interestingly, however, alignment of their amino acid sequences revealed the conservation of a highly hydrophobic central domain terminating at an invariant aspartic acid residue and thus distantly resembling the conserved motif of the 30 K superfamily (Fig. 3). This domain was flanked by two proline-rich domains showing some similarity to extensins (Fig. 3). Thus, in a sense the movement proteins of one component geminiviruses combined the 30 K superfamily and the tymovirus structural themes.

For several virus groups, in which at least one member has been completely sequenced, the movement function has not been genetically mapped and identification of a putative movement proteins by amino acid sequence comparison was very uncertain if possible at all. Among these, uncharacterized proteins of nepoviruses and closteroviruses showed a very limited sequence similarities to some members of the 30 K superfamily, in particular to caulimovirus movement proteins [33].

Evolutionary implications of amino acid sequence comparisons

The observations discussed in the previous sections show that classification of movement proteins based on comparison of their amino acid sequences does not correlate with the type of genome nucleic acid or with grouping of viruses based on phylogenetic analysis of replicative proteins [31, 34, 61]. Also, we were unable to notice any correlation between grouping of movement proteins and biological properties of the viruses encoding these proteins (e.g. host ranges). Recombination between unrelated or distantly related viruses could have played a major role in the evolution of the movement function. This is clearly the preferential explanation for situations when related movement proteins are found in viruses with different type of genome nucleic acid (e.g. the striking similarity between the putative movement proteins of caulimo/badnaviruses

consensus	xxxxxxxxxxxx@xxx&xxxx&xxx&xUxxx&Lx@x&xKDxU&UxKAxxxxxxxxEUX@Gxxxxx		
		R	R
PASV 11K 18	PSAAPSAGRLPWSRVGGEIVFTFVSVLALYLLWLWVLRKDCILLLLKAQRGRSTEELIFGPGERP	27	X60168G
WDV 10K 17	TPSIPGSSDYAWRTFVFTFGLLIAVGVAVLAYTLFLKDLILVCKAKKQRTEEIGYGNTPAR	10	P06849
TYDV V1 27	PQKAGESSEHFFSKVVVALIVLFAVGIIVYLAYTLFLKDLILLLKAKKQRTTTEIGFGNTPLR	12	M81103G
DSV V1 12	PQSPTQAPSLPWSRLGEIATLTLAVLCIYLLYIWLRLDILFVVRNRRGRVAEELEFGPAETR	33	M23022G
SCSV V1 19	PGAAPSSSGLPWSRVGEIATFTFVAVLALYLLWSVWGRDLLLVLKARRGGTTEELTFGPRERH	30	M82918G
MSV 11K 13	PTAAPSTGGVPWSRVGE <u>VALLSFVALICFYLLYLWVLRDLILVLKARQGRSTEELIFGGQAVD</u>	25	P14992

Fig. 3. Alignment of the putative movement proteins of monopartite geminiviruses. The hydrophobic region predicted to form a transmembrane helix is underlined. Consensus: for symbols see legend to Fig. 2

and parsnip yellow fleck virus shown in Fig. 2 d). In some cases recombinational transfer of movement protein genes appears very likely also between remote groups of positive-strand RNA viruses (e.g. dianthoviruses and soil borne wheat mosaic furovirus; Fig. 2 c).

Another equally important trend in the evolution of movement protein genes is the apparent high rate of mutational change leading to the very limited level of sequence conservation as featured above. Even within compact groups of viruses that share common genome organization and functional properties (e.g. tobamoviruses), the divergence between the movement protein sequences is much higher than that between the principal replicative domains although comparable to that between capsid proteins.

Searching for movement protein functions

Analogies with cellular transport systems

It has been proposed that plasmodesmata are plant analogues for certain animal cell-cell contacts, namely gap junctions, based mainly on studies in permeability of both to fluorescent dye-labelled molecules [5, 51]. Details of protein organization of plasmodesmata are poorly understood, although a protein serologically related to a major gap junctions protein, connexin, was reportedly found in plasmodesmata and gene encoding this plant protein was cloned [42]. In animal kingdom, gap junctions are important for electrical coupling of cells in various tissues [51]. Interestingly, permeability of both plasmodesmata and gap junctions for some low molecular weight tracer dyes is down-regulated by activation of protein kinase C [5]. However, movement of macromolecules, in particular nucleoproteins, through gap junctions, has not been described.

A case of movement of a nucleoprotein through a channel in a membrane is represented by nucleocytoplasmic trafficking of RNPs (reviewed in [38, 45]). Among the features of the nucleocytoplasmic export of cellular RNAs revealed thus far, several might be relevant to the mechanisms of movement of plant virus genomes. In particular, it has been shown that at all stages of nucleocytoplasmic export, mRNA is attached to specific protein frameworks, being transferred from nucleoskeleton to nuclear pore complex to cytoplasm [12]; specific proteins in mRNP are thought to chaperone association-dissociation of the mRNA to and from these frameworks [1, 53]. Components of mRNP activate enzymes within the nuclear pore complex which are essential for energy-dependent translocation of RNP through nuclear pore; two of these enzymatic activities are ATPase and RNA helicase [53]. It is tempting to speculate that mechanistic analogies may exist between the nucleocytoplasmic transport of cellular RNPs and intracellular and/or intercellular movement of viral RNPs. More specifically, the parallel between the RNA helicase in the nuclear pore and the putative viral RNA helicase in the triple block that is involved in virus movement (see above) certainly is provocative. It should be noted that plant viruses, which undergo some stages in the nucleus of the infected cells (i.e.

pararetroviruses and geminiviruses), should be able to perform both nucleocytoplasmic movement and cell-to-cell movement of their genomes; one may wonder to what extent these two processes are similar and whether movement proteins are involved in both of them.

Clues from sequence similarities

Comparison of the amino acid sequences of dianthovirus movement proteins with amino acid sequence databases revealed a marginally significant but provocative similarity with membrane proteins of the MIP family [48]. With the BLAST score of 67 and probability of random matching of 0.23, the similarity of the movement protein of RCNMV with tonoplast membrane protein TIP 1 from Arabidopsis was the highest in the entire database except for the similarities with the homologous proteins of the other dianthoviruses. Analysis of the multiple alignment of the dianthovirus movement proteins with a representative set of the MIP proteins (Fig. 4) showed that the region of similarity included the conserved motif of the "30 K superfamily" described above and the characteristic conserved motif of the MIP family [48]. Each of these motifs consisted of a putative transmembrane helix succeeded by a loop containing conserved amino acid residues (Fig. 4). The drastic difference between the two types of proteins is that MIP proteins contain six transmembrane helices whereas in the virus movement proteins the region shown in Figs. 1 and 4 is the only such predicted helix. Nevertheless, we believe that the observed sequence similarity may be functionally relevant indicating that the conserved motif of the "30 K superfamily" may mediate membrane interaction. For many proteins of this superfamily, membrane interaction is compatible with the observed localization in the cell wall [40]. However, caution is due in the interpretation of these observations as the hydrophobicity of the conserved region in different movement proteins differed significantly and not for all of them a membrane-spanning helix was confidently predicted (Fig. 1 b and data not shown). It is possible that

consensus		xxUxxx&xxxxUxxxGxUxxx&NPAUxxxxUxxxxxxxxxxxxxxxxUUxxxxxxxxUxxxx	
		A N	
GLPF <i>E. coli</i>	43	ISVIWGLGVAMAIYLTAGVSGAHLNPAVTIALWLFACFDKRVIPFIVSQVAGAFCAALVYGLY	153 P11244
WCH1 human	51	VSLAFGLSIATLAQSVGHSAGHLNPAVTLGLLLSCQISIFRALMYIIAQCVGAIVATAILSGIT	153 P29972
MIP bovine	41	VALAFGLALATLVQTVGHISGAHVNPVTFAPLVGSQMSLLRAFICYIAAQLLGAVAGAAVLYSVT	155 P06624
NO26 soybean	72	IATVWGLVLTVLVYTVGHISGGHFNPAVTIAFASTRRPPLIQVPAYVVAQLLGSILASGTLRLLF	135 P08995
TIPG <i>A. thal.</i>	60	AAVAHAFGLFVAVSVGANISGGHVNPAVTFGAFIGGNI TLLRGILYWI AQLLGSVVAACLILKFAT	126 P25818
TIPA <i>A. thal.</i>	68	VALAHAFALFAAVSAAITNVSGGHVNPAVTFGALVGGRTAIRAIYWI AQLLGA LILACLLRLPT	135 P26587
TIP1 tobacco	58	<u>VAVAHAFALFVGVSTIANISGGHLNPAVTLGLAVGGNI TLLTGFFYWI AQLLGSVVAACLILKYVT</u>	127 P21653
		* : : ** : * : : : ** * * : * : * : *	
RCNMV	65	<u>THVVLSLAPTIGVAIPGHVTVVELINPNVEGPPFQVMSGQTLWSWSPGAGKPCLMIFSIVHHQLNSDHE</u>	187
CRSV	65	<u>SDLVLSLGA TVSVGTPGSLVVELINPNMDDGPFQLVQGQSLWSWSPGSGRPCLMIFSIIHQLTADAE</u>	174 M88589G

Fig. 4. A conserved sequence motif in the movement proteins of dianthoviruses and the MIP family of membrane proteins. Asterisks show identical residues and colons show similar residues in the sequences of the movement protein of RCNMV and TIP1. The consensus (for symbols see legend to Fig. 2) includes the conserved residues, with one possible exception in the MIP-related proteins. The hydrophobic regions predicted to form transmembrane helices are underlined. The motif was extracted from an alignment generated using the MACAW program

in some of these proteins the conserved motif may be involved in hydrophobic interactions other than membrane spanning.

As mentioned above, membrane localization also has been proposed for the triple block proteins and strongly predicted for the movement proteins of monopartite geminiviruses.

The observations on the involvement of influenza virus M1 protein in nuclear export of virus ribonucleoproteins [39] prompted us to directly compare the M1 sequence to those of plant virus movement proteins. Moderate similarity was revealed between M1 and the movement proteins of bromoviruses and cucumoviruses (Fig. 5). The region of similarity included the conserved motif of the "30 K superfamily", with the invariant aspartic acid residue and the preceding hydrophobic stretch. M1 protein chaperones newly formed influenza virus RNPs from the nucleus to the cytoplasm; upon formation of virions, this protein is thought to provide the link between the RNP and the virion membrane (reviewed in [35]). Both of these functions are apparently based on the ability of M1 to interact both with RNP and the membrane providing a provocative analogy with the plant virus movement proteins. Very recently, a mutation in M1 protein gene resulting in impaired nucleocytoplasmic movement was mapped within the region of similarity to bromovirus and cucumovirus movement proteins [23].

Finally, our previous observations indicated a weak sequence similarity between a domain of cellular 90 kDa heat shock proteins with some of the movement proteins, specifically those of caulimoviruses [33].

A common denominator for all these observations may be that the function

```

consensus      UxUxxUxxxUxxxxxSxxxxGxUxxxLxDxxxxxxxLxxUxDxxx
                T      A
CMV 3A (86-212) MSVPQVLCVTRTVSTDA-EGSLKIYLADLGDKE---LSPI-D-GQ
BMV 3A (94-227) MNVPRIVCFVTRTDSAE-SGSITVSLCDSGKAARAGVLEAI-D-NQ
CCMV 3A (95-225) LNIPRILLYLVRTSHSST-VGSITIRLVDTYSASDSCLEAI-D-GQ
                * : : ** : * : * ** * : : : *** : :
INFA M1 ( 3-139) L-LTEVETVYVLSIVPSGPLKAEIAQRLEDVDFAGKNTD-LEALMEWLK

cons          xxxUxxxxxxxUx&xxxT&xxxxExxxxxxxRxx&xxxxxxxGxxxxxxx
CMV 3A CVTLHNHELPALISFQPTYDCPEMLVG--NRHRCFAVVVERHGYIGYGGTTASVCS-
BMV 3A EATIQLSALPALIALTPSYDCPEMEVIGDSDGRNRCFGIATQLSGVVGTGSAVTHA-
CCMV 3A EFTVDLSSLPCMIGFSPTYDCKLEMVD--GRRRCFGIVTELNGVIGEGHTVAMVHA-
                : : ** * * * ** * : ** :
INFA M1 TRPILSPLTKGILGFVFTLTVPSERGL--QRRRFVQNALNGNDSNNMMDRAVKLYRK
                ▲

cons          xxxxxxPxxxKxxxxxYxxxxxAxx&xxx@xRUxxxx
                G
CMV 3A -NWQAQFS-SKNN--NYTHAAAGKTLVLPYNRLAEHS
BMV 3A -YWQANFK-AKPN--NYKLHGFPATIMVMPFDRLRQLD
CCMV 3A -YWKAMFR-TKPG--NYTRVKPAAKFIAPFDRLKQLS
                * * * * * * : * : :
INFA M1 LKREITPHGAKEVALSYSTGALATCMGLIYNRMGTVT P10918

```

Fig. 5. Amino acid sequence alignment of the movement proteins of cucumoviruses and bromoviruses and influenza virus M1 protein. Asterisks show identical amino acids residues and colons show similar residues in the sequences of the putative movement protein of CCMV and M1 protein. The alignment was generated using the OPTAL program. Position of phenylalanine to serine change in M1 protein resulting in defect of vRNP export is shown by an arrowhead. Consensus: for symbols see legend to Fig. 2

of plant virus movement proteins is determined by domains mediating hydrophobic interactions. Taking into account the limited sequence conservation and the absence of strictly invariant amino acid residues, it is very unlikely that these proteins (with the exception for the putative helicase in the triple block) have any enzymatic activity.

Conclusions

Among the known groups of plant viruses [25], movement proteins so far have been identified genetically or by sequence comparison for about 75 percent. The variety of these proteins could be reduced to only three types, namely the "30 K superfamily", with the distantly similar proteins of monopartite geminiviruses; the triple block; and the proline-rich proteins of tymoviruses. Obviously, future studies may lead to characterization of movement proteins unrelated to any of these groups.

Different types of movement proteins may affect different aspects in virus-cell interaction leading to virus spread in planta. However, it appears likely that in many cases these proteins mediate hydrophobic interactions between viral and cellular components, in a general analogy with molecular chaperone action [16, 33].

As discussed in the first part of this review, a large fraction of experimental data on plant virus movement proteins is very hard to be interpreted unequivocally. Conceivably, this reflects the highly complex nature of virus-plant interaction. In this situation, it appears reasonable to focus further experiments on probing the specific functions of conserved domains revealed by amino acid sequence comparison. Hopefully, this approach will highlight new important aspects of plant virus lifestyle.

Note added in proof

Recently, it has been shown that TMV movement protein expressed by transgenic tobacco behaves as an integral membrane protein [Moore PJ, Fenczik CA, Deom CM, Beachy RN (1992) Developmental changes in plasmodesmata in transgenic tobacco expressing the movement protein of tobacco mosaic virus. *Protoplasma* 170: 115–127].

Acknowledgements

We appreciate helpful discussions with Dr. V. V. Dolja and Dr. U. Melcher. We would like to thank Drs. D. Cookmeyer, S. A. Demler, R. Kormelink, U. K. Melcher, N. Olszewski and Yu. Shirako for communicating data prior to publication. A. M. is grateful to Dr. R. J. Shepherd for constant support and encouragement.

References

1. Agutter PS (ed) (1991) *Between nucleus and cytoplasm*. Chapman and Hall, New York
2. Albrecht H, Geldreich A, Menissier de Murcia J, Kirchnerr D, Mesnard J-M, Lebeurier G (1988) Cauliflower mosaic virus gene I product detected in cell wall-enriched fraction. *Virology* 163: 503–508

3. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410
4. Atabekov JG, Taliansky ME (1990) Expression of a plant virus-coded transport function by different viral genomes. *Adv Virus Res* 38: 201–248
5. Baron-Epel O, Hernandez D, Jiang L-W, Meiners S, Schindler M (1988) Dynamic continuity of cytoplasmic and membrane compartments between plant cells. *J Cell Biol* 106: 715–721
- 5a. Beck DL, Guilford PJ, Voot DM, Andersen MT, Forster RL (1991) Triple gene block proteins of white clover mosaic potexviurs are required for transport. *Virology* 183: 695–702
6. Berna A, Gafny R, Wolf S, Lucas WJ, Holt CA, Beachy RN (1991) The TMV movement protein: role of the C-terminal 73 amino acids in subcellular localization and function. *Virology* 182: 682–689
7. Boccard F, Baulcombe D (1993) Mutational analysis of cis-acting sequences and gene function in RNA 3 of cucumber mosaic virus. *Virology* 193: 563–578
8. Bouhida M, Lockhart BEL, Olszewski N (1993) An analysis of the complete nucleotide sequence of a sugarcane bacilliform virus genome infectious to banana and rice. *J Gen Virol* 74: 15–22
9. Boulton MI, Pallaghy CK, Chatani M, MacFarlane S, Davies JW (1993) Replication of maize streak virus mutants in maize protoplasts: evidence for a movement protein. *Virology* 192: 85–93
10. Bozart CS, Weiland JJ, Dreher TW (1992) Expression of ORF-69 of turnip yellow mosaic virus is necessary for virus spread in plants. *Virology* 187: 124–130
11. Calder VL, Palukaitis P (1992) Nucleotide sequence analysis of the movement genes of resistance breaking strains of tomato mosaic virus. *J Gen Virol* 73: 165–168
12. Carter KC, Bowman D, Carrington W, Fogarty K, McNeil JA, Fay FS, Lawrence JB (1993) A three-dimensional view of precursor messenger RNA metabolism within the mammalian nucleus. *Science* 259: 1330–1335
13. Chapman S, Kavanagh T, Baulcombe D (1992) Potato virus X as a vector for gene expression in plants. *Plant J* 2: 549–557
14. Citovsky V, Knorr D, Schuster G, Zambryski P (1990) The P 30 movement protein of tobacco mosaic virus is a single-stranded nucleic acid binding protein. *Cell* 60: 637–647
15. Citovsky V, Knorr D, Zambryski P (1991) Gene I, a potential cell-to-cell movement locus of cauliflower mosaic virus, encodes an RNA-binding protein. *Proc Natl Acad Sci USA* 88: 2476–2480
16. Citovsky V, Zambryski P (1991) How do plant virus nucleic acids move through intercellular connections? *BioEssays* 13: 373–379
17. Citovsky V, Wong ML, Shaw AL, Venkataram Prasad BV, Zambryski P (1992) Visualization and characterization of tobacco mosaic virus movement protein binding to single-stranded nucleic acids. *Plant Cell* 4: 397–411
18. Dawson WO (1990) Relationship of tobacco mosaic virus gene expression to movement within plant. In: Pirone TP, Shaw JG (eds) *Viral genes and plant pathogenesis*. Springer, Wien New York, pp 39–52
19. Deom CM, Lapidot M, Beachy RN (1992) Plant virus movement proteins. *Cell* 69: 221–224
20. Derrick PM, Barker H, Oparka KJ (1992) Increase in plasmodesmatal permeability during cell-to-cell spread of tobacco rattle virus from individually inoculated cells. *Plant Cell* 4: 1405–1412
21. Doolittle RF (1986) *Of URFs and ORFs. A primer on how to analyze derived amino acid sequences*. University Science Books, Mill Valley
22. Dorokhov Yul, Alexandrova NM, Miroshnichenko NA, Atabekov JG (1984) The

- informosome-like virus-specific ribonucleoprotein (vRNP) may be involved in the transport of tobacco mosaic virus infection. *Virology* 137: 127–134
23. Enami K, Qiao Y, Fukuda R, Enami M (1993) An influenza virus temperature-sensitive mutant defective in the nuclear-cytoplasmic transport of the negative-sense viral RNAs. *Virology* 194: 822–827
 24. Erny C, Schoumacher F, Jung C, Gagey M-J, Godefroy-Colburn T, Stussi-Garaud C, Berna A (1992) An N-proximal sequence of alfalfa mosaic virus movement protein is necessary for association with cell walls in transgenic plants. *J Gen Virol* 73: 2115–2119
 25. Francki RIB, Fauquet CM, Knudson DL, Brown F (eds) (1991) Classification and nomenclature of viruses. Fifth Report of the International Committee on Taxonomy of Viruses. Springer, Wien New York (Arch Virol [Suppl] 2)
 - 25a. Gilmer D, Bouzoubaa S, Guilley H, Richards K, Jonard G (1992) Efficient cell-to-cell movement of beet necrotic yellow vein virus requires 3' proximal genes located on RNA 2. *Virology* 189: 40–47
 26. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1988) A novel superfamily of nucleoside triphosphate-binding motif-containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS Lett* 239: 16–24
 27. Gorbalenya AE, Blinov VM, Donchenko AP, Koonin EV (1989) An ATP-binding motif is the most conserved sequence in a highly diverged group of proteins involved in positive strand RNA replication. *J Mol Evol* 28: 256–268
 28. Hacker DL, Petty ITD, Wei N, Morris TJ (1992) Turnip crinkle virus genes required for RNA replication and virus movement. *Virology* 186: 1–8
 29. Hamilton WDO, Boccara M, Robinson DJ, Baulcombe DC (1987) The complete nucleotide sequence of tobacco rattle virus RNA-1. *J Gen Virol* 70: 963–968
 30. Hull R, Saedler J, Longstaff M (1986) The sequence of carnation etched ring virus DNA: comparison with cauliflower mosaic virus and retroviruses. *EMBO J* 5: 3083–3090
 31. Koonin EV (1991) The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J Gen Virol* 72: 2179–2207
 32. Koonin EV (1992) Virus evolution: time for Sturm und Drang. *Semin Virol* 3: 311–313
 33. Koonin EV, Mushegian AR, Ryabov EV, Dolja VV (1991) Diverse groups of plant DNA and RNA viruses share related movement proteins that may possess chaperone-like activity. *J Gen Virol* 72: 2895–2903
 34. Koonin EV, Dolja VV (1993) Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit Rev Biochem Mol Biol* (in press)
 35. Lamb RA (1989) Genes and proteins of the influenza viruses. In: King RM (ed) *The influenza viruses*. Plenum Press, New York, pp 1–87
 36. Linstead PJ, Hills GJ, Plaskitt KA, Wilson IG, Harker CL, Maule AJ (1988) The subcellular localization of gene I product of cauliflower mosaic virus is consistent with a function associated with virus spread. *J Gen Virol* 69: 1809–1818
 37. MacKenzie DJ, Tremaine JH (1988) Ultrastructural location of non-structural protein 3A of cucumber mosaic virus in infected tissue using monoclonal antibodies to a cloned chimeric fusion protein. *J Gen Virol* 60: 2387–2394
 38. Maquat LE (1991) Nuclear mRNA export. *Curr Opin Cell Biol* 3: 1004–1012
 39. Martin K, Helenius A (1991) Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import. *Cell* 67: 117–130
 40. Maule AJ (1991) Virus movement in infected plants. *Crit Rev Plant Sci* 9: 457–473
 41. Mehlin H, Daneholt B, Skoglund U (1992) Translocation of a specific premessenger

- ribonucleoprotein particle through the nuclear pore studied with electron microscope tomography. *Cell* 69: 605–613
42. Meiners S, Xu A, Schindler M (1991) Gap junction protein homologue in *Arabidopsis thaliana*: evidence for connexins in plants. *Proc Natl Acad Sci USA* 88: 4119–4122
 43. Melcher U (1990) Similarities between putative transport proteins of plant viruses. *J Gen Virol* 71: 1009–1018
 44. Morozov SYu, Dolja VV, Atabekov JG (1989) Probable reassortment of genomic elements among elongated RNA-containing plant viruses. *J Mol Evol* 29: 52–62
 45. Nigg EA, Bauerle PA, Luhrmann R (1991) Nuclear import-export: in search of signals and mechanisms. *Cell* 66: 15–22
 46. Osman TAM, Buck KW (1991) Detection of the movement protein of red clover necrotic mosaic virus in a cell wall fraction from infected *Nicotiana clevelandii* plants. *J Gen Virol* 72: 2853–2856
 47. Osman TAM, Hayes RJ, Buck KW (1992) Cooperative binding of the red clover necrotic mosaic movement protein to single stranded nucleic acids. *J Gen Virol* 73: 223–227
 48. Pao GM, Wu L-F, Johnson KD, Hofte H, Chrispeels MJ, Sweet G, Sandal NN, Saler Jr MH (1991) Evolution of the MIP family of integral membrane transport proteins. *Mol Microbiol* 5: 33–37
 49. Perbal MC, Thomas CL, Maule AJ (1993) Cauliflower mosaic virus gene I product (P1) forms tubular structures which extend from the surface of infected protoplasts. *Virology* 195: 281–285
 - 49 a. Petty ITD, French R, Jones RW, Jackson AO (1990) Identification of barley stripe mosaic virus genes involved in viral RNA replication and systemic spread. *EMBO J* 9: 3452–3457
 50. Rao MJK, Argos P (1985) A conformation preference parameter to predict helices in integral membrane proteins. *Biochim Biophys Acta* 869: 197–214
 51. Robards AW, Lucas WJ, Pitts JD, Jongsma JD, Spray DC (eds) (1990) *Parallels in cell-to-cell junctions in plants and animals*. Springer, Berlin Heidelberg New York Tokyo
 52. Sacher R, Ahlquist P (1989) Effects of deletions in the N-terminal basic arm of brome mosaic virus coat protein on RNA packaging and systemic infection. *J Virol* 63: 4545–4552
 53. Schroder HC, Ugarkovic D, Langen P, Bachmann M, Dorn A, Kuchino Y, Muller WEG (1990) Evidence for involvement of a nuclear envelope-associated RNA helicase activity in nucleocytoplasmic RNA transport. *J Cell Physiol* 145: 136–146
 - 53 a. Schuler GD, Altschul SF, Lipman DJ (1991) A workbench for multiple alignments construction and analysis. *Proteins Struct Funct Genet* 9: 180–190
 54. Thomas GL, Perbal C, Maule AJ (1993) A mutation in cauliflower mosaic virus gene I interferes with virus movement but not virus replication. *Virology* 192: 415–421
 55. Traynor P, Young BM, Ahlquist P (1991) Deletion analysis of brome mosaic virus 2 a protein: effects on RNA replication and virus spread. *J Virol* 65: 2807–2815
 56. van Lent J, Storms M, van der Meer F, Wellink J, Goldbach R (1991) Tubular structures involved in movement of cowpea mosaic virus are also formed in infected cowpea protoplasts. *J Gen Virol* 72: 2615–2623
 57. Wolf S, Deom CM, Beachy RN, Lucas WJ (1991) Plasmodesmatal function is probed using transgenic tobacco plants that express a virus movement protein. *Plant Cell* 3: 593–604
 58. Wellink J, van Kammen A (1989) Cell-to-cell transport of cowpea mosaic virus requires both the 58/48 K proteins and the capsid proteins. *J Gen Virol* 51: 317–325
 59. Xiong Z, Kim KH, Giesman-Cookmeyer D, Lommel S (1993) The roles of the red clover necrotic mosaic virus capsid and cell-to-cell movement proteins in systemic infection. *Virology* 192: 27–32

60. Ziegler-Graff V, Guilford PJ, Baulcombe DC (1991) Tobacco rattle virus RNA-1 29 K gene product potentiates viral movement and also affects symptom production in tobacco. *Virology* 182: 144–155
61. Zimmern D (1988) Evolution of RNA viruses. In: Holland JJ, Domingo E, Ahlquist P (eds) RNA genetics. CRC Press, Boca Raton

Authors' address: Dr. A. R. Mushegian, Department of Plant Pathology, University of Kentucky, Lexington, KY 40546-0091, U.S.A.

Received June 9, 1993