

Hypothesis

A possible consensus sequence for VPg of viruses in the family *Comoviridae*M.A. Mayo^{a,*}, C. Fritsch^b^aScottish Crop Research Institute, Invergowrie, Dundee, DD2 5DA, UK^bInstitut de Biologie Moleculaire des Plantes, 12 rue du General Zimmer, 67084 Strasbourg Cedex, France

Received 19 September 1994

Abstract Alignment of the amino acid sequences of the genome-linked proteins (VPg) of nepoviruses and comoviruses suggests that viruses within the family Comoviridae have VPg which share a consensus sequence.

Key words: Genome-linked protein; VPg; Nepovirus; Comovirus

Plant viruses in the family *Comoviridae* are classified in the genera *Comovirus*, *Nepovirus* and *Fabavirus* [1], largely according to which type of invertebrate vector is involved in their transmission. A characteristic feature of viruses in the genera *Comovirus* and *Nepovirus* is that the two RNA molecules which make up the virus genome have a small genome-linked protein (VPg) attached to their 5'-ends and a poly(A) tract at their 3'-ends. The genome RNA each contain one large open reading frame which encodes a large polyprotein from which functional proteins are released by the action of virus-coded proteases. The VPg is cleaved from the central part of the polyprotein encoded by RNA B of comoviruses [2] or RNA-1 of nepoviruses [3,4].

The amino acid sequences of the VPg of grapevine fanleaf nepovirus (GFLV) [3] and tomato back ring nepovirus (TBRV) [4] have been determined directly. The amino acid sequence of the VPg of tomato ringspot nepovirus (TomRSV) was inferred from the sequence of the RNA-1 polyprotein [5], and we have deduced that of grapevine chrome mosaic nepovirus (GCMV) VPg on the basis of the strong sequence similarity between the RNA-1 polyproteins of GCMV and TBRV [6]. These four sequences are aligned in Fig. 1. With the insertion of two gaps it was possible to discern a consensus sequence of:

$E/D-x_{(1-2)}-Y-x_{(2)}-R-N-x_{(4-5)}-R$

The sequences of the VPgs from three comoviruses for which the RNA-B sequences are known [7] are more similar than are those of the four nepovirus VPgs. Fig. 1 compares the alignments of the nepovirus and comovirus VPgs. Although the nepovirus VPgs are relatively diverse in sequence, some similarities are detectable among all the sequences if a gap is inserted in all the nepovirus sequences (Fig. 1). The broader consensus sequence applicable to both nepoviruses and comoviruses is:

$E/D-x_{(1-3)}-Y-x_{(3)}-N-x_{(4-5)}-R$

Some similarities seem to cut across taxonomic lines, in that the VPgs of TBRV and cowpea severe mosaic comovirus

(CPSMV) share a larger consensus (additional matches shown as dashed lines in Fig. 1) of:

$K-x_{(4)}-E/D-x_{(2-3)}-Y-R-x-R-N-I-P-x_{(3)}-R$

Other features common to all the VPgs of comovirid viruses are a basic charge and a hydrophilic nature. It is likely that some or all of the features shared by all the VPgs are related to the role played by VPg in virus replication. For some, but not all, nepoviruses the presence of a VPg on the RNA molecules is thought to stabilize their infectivity against nuclease action [8,9]. Picornavirus RNA is also linked to a VPg and it has been proposed that, by analogy with the replication of poliovirus RNA [10], a VPg-pU molecule acts as a primer for RNA transcription during replication of comoviruses [11]. In addition to giving clues as to the biological role(s) of VPg, the proposed consensus sequence may also prove useful in locating the VPg in polyprotein sequences of other nepoviruses or comoviruses.

Nepoviruses

GFLV	SEPRLE-ERYSPNRNF-VSRISKIRG
TomRSV	STIPSGSYAD--VYNARNMTRVFRPQ
GCMV	AHSVYSADGGD-RGYRSRNIPINHRYSYAR
TBRV	AQQKSSSQE-GGYRARNIPIHHRYAYAK

Comoviruses

CPSMV	SRKPNRFDVAQYRYRNIPLRKNWAEGQ
CPMV	SRKPNRFDMQYRYNNVPLKRRVWADAQ
RCMV	SRKPNRFEVQYRYKNVPLTRRSWGNAQ

Fig. 1. Alignment of the complete amino acid sequences of nepovirus and comovirus VPgs. Letters in reverse contrast show the proposed consensus residues in the VPgs of nepoviruses and where these match the sequences of comovirus VPgs. Vertical bars indicate extra matches between the VPg of tomato black ring nepovirus (TBRV) and cowpea severe mosaic comovirus (CPSMV). Other virus acronyms are GFLV (grapevine fanleaf nepovirus), TomRSV (tomato ringspot nepovirus), GCMV (grapevine chrome mosaic nepovirus), CPMV (cowpea mosaic comovirus) and RCMV (red clover mottle comovirus).

*Corresponding author. Fax: (44) (382) 562 426.

Acknowledgements: We thank D. Rochon for access to data on TomRSV RNA-1 prior to its publication.

References

- [1] Mayo, M.A. and Martelli, G.P. (1993) *Arch. Virol.* 133, 496–498.
- [2] Wellink, J., Rezelman, G., Goldbach, R. and Beyreuther, K. (1986) *J. Virol.* 59, 50–58.
- [3] Pinck, M., Reinbolt, J., Loudes, A.M., Leret, M. and Pinck, L. (1991) *FEBS Lett.* 284, 117–119.
- [4] Hemmer, O., Greif, C., Dufoucq, P., Reinbolt, J. and Fritsch, C. (1994) *Virology* (in press).
- [5] Rott, M.E., Gilchrist, A., Lee, L. and Rochon, D. (1994) *J. Gen. Virol.* (in press).
- [6] Le Gall, O., Candresse, T., Brault, V. and Dunez, J. (1989) *Nucleic Acids Res.* 17, 7795–7807.
- [7] Chen, X. and Bruening, G. (1992) *Virology* 191, 607–618.
- [8] Mayo, M.A., Barker, H. and Harrison, B.D. (1982) *J. Gen. Virol.* 59, 149–162.
- [9] Barker, H. and Mayo, M.A. (1982) *Annu. Rep. Scot. Crop Res. Inst. for 1981*, p. 109.
- [10] Takeda, M., Kuhn, R.J., Yang, C.F., Takegami, T. and Wimmer, E. (1986) *J. Virol.* 60, 43–53.
- [11] Eggen, R. and Van Kammen, A. (1988) in: *RNA Genetics*, vol I (Ahlquist, P. and Domingo, E., eds) pp. 49–69, CRC Press, Boca Raton, FL.