Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 9156–9160, October 1992
Evolution

Coat protein gene duplication in a filamentous RNA virus of plants

VITALY P. BOYKO*, ALEXANDER V. KARASEV†‡, ALEXEY A. AGRANOVSKY*§, EUGENE V. KOONIN†¶, AND VALERIAN V. DOLJA*‖

*A. N. Belozersky Laboratory at Moscow State University, Moscow 119899, Russia; †Institute of Microbiology, Russian Academy of Sciences, Moscow 117811, Russia; and ¶National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894

Communicated by Robert J. Shepherd, June 12, 1992

ABSTRACT **Computer-assisted analysis revealed a striking sequence similarity between the putative 24-kDa protein (p24) encoded by open reading frame (ORF) 5 of beet yellows closterovirus and the coat protein of this virus encoded by the adjacent ORF6. Both of these proteins are closely related to the homologous proteins of another closterovirus, citrus tristeza virus. It is hypothesized that the genes for coat protein and its diverged tandem copy have evolved by duplication. Phylogenetic analysis using various methods for tree generation suggested that the duplication was already present in the genome of the common ancestor of the two closteroviruses. The genes for p24 and coat protein of beet yellows closterovirus were cloned, transcribed, and translated *in vitro* yielding products of the expected size. It was shown that p24 is translated starting from the first of the two alternative AUG codons located near the 5′ terminus of ORF5. The presence of a single protein species in beet yellows closterovirus virions and the near identity of the amino acid composition of this protein with the composition of the ORF6 but not the ORF5 product indicated that p24 is not a major virion component. Most of the amino acids that are conserved in the coat proteins of filamentous viruses of plants are retained also in p24. These observations suggest that p24 may share some structural and functional features with the coat protein but probably fulfills a distinct function in virus reproduction.**

RNA viruses have small genomes ranging in size from 3000 to 30,000 nucleotides (nt). It is frequently conjectured that further increase of the genome size is precluded both by the fragility of long RNA molecules and by the low fidelity of RNA replication, this imposing severe constraints on the content of RNA genomes (1). However, recent findings suggest that some of the RNA viruses [e.g., coronaviruses (2)] may have more complex genome organization than previously suspected.

Among the plant RNA viruses, the most peculiar gene arrangement probably has been revealed in beet yellows virus (BYV). BYV is the type member of the closterovirus group. It has flexuous filamentous particles containing a single 22-kDa coat protein species and a single molecule of positive-sense RNA of ≈14,500 nt (3–5). BYV virion RNA possesses a cap structure at the 5′ end and lacks a poly(A) tail at the 3′ end (5). The sequence of the 3′-terminal 6746 nt has been determined (6). This portion of BYV genome encompasses an array of eight open reading frames (ORFs) encoding (in the 5′ → 3′ direction) a C-terminal fragment of the putative RNA polymerase (incomplete ORF1), a 6.4-kDa hydrophobic protein (ORF2), a 65-kDa homologue of cellular 70-kDa heat shock proteins HSP70 (ORF3), a 64-kDa protein distantly related to 90-kDa heat shock proteins HSP90 (ORF4), a 24-kDa protein (ORF5), a 22-kDa coat protein (ORF6), a 20-kDa protein (ORF7), and a 21-kDa protein (ORF8) (Fig. 1*A* and refs. 6–8). BYV infection is accompanied by the formation of at least five subgenomic RNAs involved in the expression of the 3′-proximal genes (9). Recent computer-assisted analysis of the amino acid sequences of the coat proteins of the plant RNA viruses with filamentous particles clearly showed that coat proteins of potexviruses, carlaviruses, potyviruses, bymoviruses, BYV, and another alleged closterovirus, apple chlorotic leaf spot virus, constitute a distinct family with a characteristic pattern of conserved amino acid residues (10).

Here we show that the 24-kDa protein (p24) encoded by ORF5 of BYV is a diverged copy of the coat protein. The p24, however, has not been detected in the virus particle and may perform a function in virus multiplication other than virion formation. Thus we report here coat-protein gene duplication among viruses with the helical type of capsid structure.

MATERIALS AND METHODS

Construction of cDNA Clones. Two plasmids, pBY24 and pBYCP, were constructed containing individual p24 and coat protein genes under the control of T7 RNA polymerase promoter (Fig. 1). pBYCP contained a *Bam*HI–*Xho* I fragment of clone x19 inserted into plasmid pTZ18 (Pharmacia); pBY24 contained the insert of clone r9 introduced into plasmid pTZ18 between *Pst* I and *Asp*718 sites. The origin and restriction maps of clones x19 and r9 have been reported (6). Plasmid pA103 containing the potato virus X (PVX) coat protein gene under transcriptional control of a T7 promoter (11) was kindly provided by S. Morozov (A. N. Belozersky Laboratory at Moscow State University, Moscow).

Transcription and Translation *in Vitro*. Transcription of pBY24, pBYCP, or pA103 plasmids, linearized with *Hin*dIII, *Pst* I, or *Hin*dIII, respectively, was performed using T7 RNA polymerase (Biopol, Moscow) as described (11). Uncapped transcripts were translated in rabbit reticulocyte lysates or in Krebs-2 ascites extracts as described (5). Translation products were analyzed in 8–20% polyacrylamide gradient gels (12). Oligodeoxynucleotides complementary to the BYV RNA sequence downstream of the first AUG codons of the p24 and coat protein genes were used for translation arrest in Krebs-2 ascites extracts as described (13). Oligo I (5′-dTGAAGTGGATTAAGTC) was complementary to nt 4209–4224 and oligo II (5′-dGCAGTCTTCTCCGTGC) was complementary to nt 4977–4992 of the sequenced portion of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: BYV, beet yellows virus; CTV, citrus tristeza virus; nt, nucleotides; ORF, open reading frame; PVX, potato virus X.

‡Present address: Department of Plant Pathology, University of California, Riverside, CA 92521.
§Present address: Federal Biological Research Center for Agriculture and Forestry, Institute for Biochemistry and Plant Virology, Messeweg 11/12, D-3300 Braunschweig, Federal Republic of Germany.
‖To whom reprint requests should be addressed at the present address: Department of Biology, Texas A&M University, College Station, TX 77843-3258.

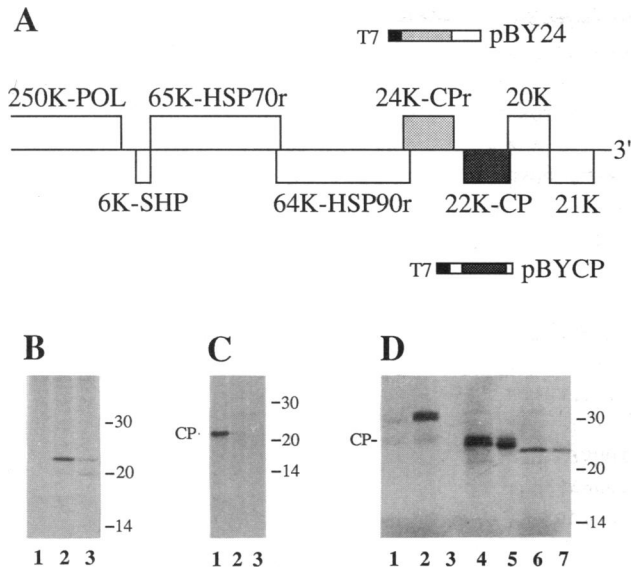

FIG. 1. Genetic map of the 3'-terminal half of the BYV genome with the partial diagrams of expression plasmids pBY24 and pBYCP (*A*) and *in vitro* expression of BYV p24 and coat protein (*B–D*). (*A*) Open and shaded boxes show ORFs 1–8 encoded in the BYV RNA (6, 7). K, kDa; POL, RNA polymerase; SHP, small hydrophobic protein; HSP70r, 70-kDa cellular heat shock protein-related product; HSP90r, 90-kDa cellular heat shock protein-related product; CPr, coat protein-related product (light shading); CP, coat protein (dark shading); T7, phage T7 RNA polymerase promoter (solid boxes). (*B–D*) Autoradiograms show SDS/PAGE analysis of translation products directed by the respective T7 RNA polymerase transcripts in Krebs-2 (*B* and *C*) or reticulocyte lysate (*D*) cell-free systems. (*B*) Lanes: 1, no RNA added; 2, pBY24 (ORF5) transcript; 3, pBY24 transcript plus oligo I. (*C*) Lanes: 1, pBYCP (ORF6) transcript; 2, pBYCP transcript plus oligo II; 3, no RNA added. (*D*) Lanes: 1, no RNA added; 2, pA103 (coat protein gene of PVX) transcript; 3, same as in lane 2 after immunoprecipitation with anti-BYV immunoglobulins; 4, pBYCP transcript; 5, same as in lane 4 after immunoprecipitation with anti-BYV immunoglobulins; 6, pBY24 transcript; 7, same as in lane 6 after immunoprecipitation with anti-BYV immunoglobulins. Positions of BYV coat protein (CP) and molecular mass markers (in kDa) are indicated.

BYV genome (6). Immunoprecipitation of ^{35}S- and ^{14}C-labeled translation products was as described (14).

Computer-Assisted Amino Acid Sequence Analysis. Amino acid sequences of coat protein and p24 of BYV (6) and the coat protein of citrus tristeza virus (CTV) (15) were inferred from the corresponding nucleotide sequences. Sources of the sequences of coat proteins of other filamentous plant RNA viruses are given in ref. 10. Amino acid data base screening was performed using the nonredundant data base created at the National Center for Biotechnology Information (National Institutes of Health, Bethesda, MD) and the BLAST program (16). Local sequence similarity analysis was performed using the DOTHELIX program (17), and multiple alignments were generated in the hierarchical fashion using the OPTAL program (18). Cluster dendrograms were obtained by the TREE program implementing the unweighted pair group maximum averages clustering algorithm (19). Parsimony trees were generated by the PROTPARS program in the PHYLIP package (20). Rate-independent trees were produced by a program implementing the so-called maximal topological similarity principle allowing construction of a tree in which the set of nearest neighbor species quartets is best compatible with that derived from the distance matrix (21). Protein secondary structure was predicted by the PROTEIN2 program implementing the modified algorithm of Garnier and coworkers (22). Programs DOTHELIX and PROTEIN2 are components of the GENEBEE program package (23).

RESULTS

Expression of BYV p24 and Coat Protein *in Vitro*. To study the expression of the two genes encoded by ORF5 and ORF6, we constructed recombinant plasmids pBY24 and pBYCP, containing the complete sequences of the respective genes downstream from the T7 promoter (Fig. 1*A*). After transcription the resulting RNAs were translated in Krebs-2 (Fig. 1 *B* and *C*) or reticulocyte lysate (Fig. 1*D*) cell-free systems. In both *in vitro* systems, pBY24 and pBYCP transcripts directed the synthesis of products with the apparent molecular masses close to the predicted molecular mass for ORF5 and ORF6 products (Fig. 1 *B* and *C*).

As there are two AUG codons in the 5' region of ORF5, both in a context favorable for translation (6, 24), we attempted to identify the actual initiator of the p24 product synthesized *in vitro*. To this end, oligo I, which is complementary to the portion of the p24 gene between the first and the second AUG codons, was added to the Krebs-2 translation mixture with the respective RNA transcript. Due to the presence of endogenous RNase H activity in Krebs-2 extract (13), this should lead to cleavage of RNA and generation of a truncated mRNA. As expected, a dramatic reduction in the intensity of the p24 band and appearance of a new band of 20-kDa were observed (Fig. 1*B*, lane 3). We concluded that the synthesis of p24, at least *in vitro*, is initiated at the 5'-proximal AUG codon of ORF5. Our previous data on the translation of denatured double-stranded forms of subgenomic RNAs isolated from BYV-infected plants (9) have suggested that ORF5 translation *in vivo* also starts from the first AUG codon. Oligo II complementary to the 5' portion of the ORF6 nearly completely arrested the synthesis of coat protein (Fig. 1*C*, lane 2). No additional products became visible as ORF6 possesses no internal AUG codons.

Immunoglobulins isolated from the polyclonal anti-BYV serum precipitated the ORF6-directed coat protein (Fig. 1*D*, lanes 4 and 5) but not the distantly related PVX coat protein (negative control, Fig. 1*D*, lanes 2 and 3). A weak immunoprecipitation of the ORF5-directed protein (p24) was also observed (Fig. 1*D*, lanes 6 and 7). At least partially this may be attributed to the existence of common antigenic determinants in the BYV coat protein and p24.

The presence of the doublet band of BYV coat protein (Fig. 1*D*, lanes 4 and 5) was due most likely to the artifacts of *in vitro* translation in the reticulocyte cell-free system. The doubling was not observed when the Krebs-2 system (cf. Fig. 1 *C*, lane 1, and *D*, lanes 4 and 5) or certain batches of reticulocyte lysate (data not shown) were used for translation. The somewhat anomalous relative mobility of BYV coat protein and p24 seen on Fig. 1*C* might be also dependent on translation and/or electrophoresis artifacts. In line with this, and with the previous observations (25), the coat protein of another filamentous RNA virus, PVX, having a calculated molecular mass of 26-kDa migrated in our experiment as a 31-kDa protein (Fig. 1*D*, lane 2).

BYV p24 Protein Is Related to the Coat Proteins of Closteroviruses. Unexpectedly, computer-assisted analysis of the amino acid sequences of the BYV gene products, using the local similarity search program DOTHELIX, demonstrated a statistically significant similarity between the p24 and coat protein (data not shown). When the sequences of p24 and coat protein were compared with the nonredundant amino acid sequence data base, the only protein showing a significant relatedness to both of these sequences of BYV was the coat protein of another closterovirus, CTV. Analysis of the resulting alignment of the three proteins revealed the conservation of 23 out of 206 aligned amino acid residues in all three sequences, with additional 39 positions occupied by functionally similar residues (Fig. 2) and yielded scores of 10.2 and 8.8 standard deviations for comparisons of nearly

```
consensus         +   +  +            +       +   +          +
BYV CP  ( 3-193) AIAT-FENVSLAD--QTC---LHGEDCDKLRKNFEECLKLK--GVPEDNL
                 *: * *: : **:  :     *   : * *   : * :*     : ::::
BYV p24 (20-216) AMETFFNSYDLAEYSEVNPNKLNRKETDELLGVIRERFKSEL-VITDEDF
                  : *  * :      :     *   :   *   :   :*     : * *
CTV CP  (38-221) TLIT-MNDVR-----Q-----LSTQQNAALNRDLFLTLKGKHPNLPDKDK
                 αααααααααα                ααααααααααααα          α
                 1         10        20        30        40

consensus  ++ ++++ +    +S          +             +    +         +
BYV CP   GIALGLCLYSCATIGTSNKVNVQPTS-TFIKASFGGGKELYLTHGELNSF-LGSQKL
            *:  *:  * * ** ***:      :      ***  :*    :  * *  : :
BYV p24  VKHLAFALIRAANITTSVKVNYVGAY-EYTI----GGKK-FLVKDAW-VFPLIKECM
            :*  * * *  ::*:  :  :    **      *    *    *  : :    :
CTV CP   DFRIAMMLYRLAVKSSSLQSDDDATGITYTR----EGVEVDLSDKLWTDVVFNSKGI
         ααααααααααααα              ββββ         ααααααααααααααααα
         50        60        70        80        90        100

                          f                   w          rf        f
consensus        +|R| ++  y      +            y      +   ky   f|D|Fy
BYV CP   LE-GKPNKL|R|CFCRTFQKDYISLRKEYRGKLPPIARA-NRHGLPAEDHYLAA|D|FI
             *** :|*| ** **:  ** : : :  **  : *   : *:*    :*:*|*|*:
BYV p24  KKFNKPNPV|R|TFCATFEDAYIVIARSLP-KL-FLNRTIGKRGIPSGYEFLGA|D|FL
            *: * :|*| :  * :  *: : *     * :  * :   *** ** :* *|*|**
CTV CP   G--NRTNAL|R|VWGRTNDALYLAFCRQNR-NLSYGGRPLD-AGIPAGYHYLCA|D|FL
         α         ββββ   αααααααα                       ββββαα ααα
             110       120       130       140       150       160

                      e
consensus  +          d ++  +                    +   +
BYV CP   STSTE-LTDLQQSRLLLA-RENATHTEFSS-ESPVTSLKQLGRGLG
         : ::  * * :   :* * *        ** :  : ** :*** *
BYV p24  TATSVCLNDHEKAIVLQASRAAIDRAVSSSVDGKIVSLFDLGR-LS
         *:    * * * *:::** :  : :      *  :  : :**: :
CTV CP   TGAG--LTDLECAVYIQA-KEQLLKKRGAD-DVVVTNVRQLGK-FN
          ααα     ααααααααααααααα          βββββ
                  170       180       190       200
```

FIG. 2. Alignment of the amino acid sequences of p24 and coat proteins (CPs) of BYV and CTV closteroviruses. Asterisks, identical amino acid residues; colons, similar residues. The grouping of amino acid residues by physicochemical similarity is as follows: G,A; S,T; D,E,N,Q; K,R; I,L,V,M,F,Y,W. "Consensus" includes amino acid residues that are conserved in 16 sequences of coat proteins of filamentous plant viruses, including 15 sequences aligned previously (10), and the CTV coat protein sequence (uppercase type) or in at least 12 out of 16 sequences (lowercase type). Plus signs denote hydrophobic residues (I,L,V,M,F,Y,W,C,A) conserved in at least 12 sequences. The positively charged arginine residue and the negatively charged aspartic acid residue thought to form a salt bridge (10) are boxed. α-Helix (α) and β-strand (β) assignments are shown where the prediction agreed for at least two sequences. The positions of the aligned regions in the proteins are indicated in parentheses.

complete sequence of p24 with those of the coat proteins of BYV and CTV, respectively.

Upon inspection of the recently reported sequence of an internal RNA portion of another BYV strain (26), we came across the same gene duplication. Moreover, the sequence of the CTV genome upstream to the coat protein gene has been determined and was kindly provided by H. Pappu, E. J. Anderson, and C. L. Niblett (personal communication). Analysis of this sequence revealed the presence of an ORF coding for a 26-kDa protein (p26), which is related to BYV p24 and to coat proteins of BYV and CTV. The similar arrangement of two adjacent ORFs in BYV and CTV genomes suggested that the coat protein gene duplication was common to the two closteroviruses.

Coat Protein and p24 of BYV May Share Similar Structural Features. Inspection of the alignment shown in Fig. 2 suggested that p24 and the coat proteins of BYV and CTV may be structurally related and may share some common properties. Indeed, against the background of the considerable overall divergence (62 out of 206 aligned residues, or 30.0%, were identical or similar in p24 and the two closterovirus coat proteins), nearly all of the consensus amino acid residues typical of the coat proteins of the filamentous viruses were retained in p24. Where a deviation from the consensus was seen, it was usually found both in p24 and in the coat proteins of BYV and CTV (Fig. 2). The conserved residues were mostly hydrophobic, probably being involved in ensuring the proper folding of the coat proteins (10). The three residues, which are strictly conserved in all proteins of this family, Ser-66, Arg-115, and Asp-158, were conserved also in p24 (Fig. 2). The latter two residues are thought to form a functionally important salt bridge in the coat proteins of all filamentous viruses (10). Also, secondary structure predictions indicated that the main structure elements, or at least the α-helical type of structure, were conserved in the coat proteins and in p24 (Fig. 2). The p24 of a British strain of BYV (26) encompassed 4 amino acid substitutions as compared to our sequence, whereas 10 substitutions were found in the coat proteins. When the sequences of p24 and coat protein of BYV were compared with the homologous sequences of the CTV proteins, very close levels of similarity were revealed for each pair of proteins (data not shown). These observations emphasized the importance of the conservation of p24 structure for the reproduction of closteroviruses.

Phylogenetic Analysis. It has been shown that the coat proteins of all filamentous positive-strand RNA plant viruses represent one phylogenetically compact protein family (10). The alignment of the closterovirus coat proteins with their diverged tandem copies (p24 and p26) was fitted into the previously generated complete alignment of this family (10), and the resulting alignment was subjected to phylogenetic analysis using various tree-generating algorithms. It was

demonstrated that p24, p26, and the coat proteins of BYV and CTV constitute a distinct group within this family of viral structural proteins (Fig. 3). This is fully compatible with the hypothesis that the genes for p24 and p26 of closteroviruses and their coat proteins evolved by duplication. Moreover, the observed tree topology strongly suggested that the duplication was already present in the genome of the common ancestor of BYV and CTV.

DISCUSSION

In this paper we showed that the proteins encoded by the adjacent ORF5 and ORF6 in the RNA genome of BYV closterovirus are related to each other. A similar duplication was observed also in another closterovirus, CTV (H. Pappu, E. J. Anderson, and C. L. Niblett, personal communication). The products of both ORFs of BYV were expressed *in vitro* by transcription of the respective cloned genes and translation of the resulting transcripts. Comigration of the ORF6 product with the BYV coat protein during polyacrylamide gel analysis and its precipitation by the anti-BYV immunoglobulins clearly demonstrated that this ORF encodes the coat protein. Efficient *in vitro* expression of the RNA coding for ORF5 product (p24) as well as the tentative identification of the corresponding abundant subgenomic mRNA in BYV-infected plants (9) suggested that this protein too might be expressed *in vivo* to a significant level.

Examination of the pattern of conserved amino acid residues in the alignment of p24 with the two closterovirus coat proteins (Fig. 2) indicated that p24 might retain the general spatial folding and some crucial properties of the coat proteins, e.g., the ability to form specific particles with the viral RNA. On the other hand, there is convincing evidence that p24 cannot be a major component of BYV virions. (*i*) A single protein species has been revealed in the BYV virion that comigrates in a polyacrylamide gel with the *in vitro*-expressed ORF6 product, whereas the *in vitro* products of ORF5 and ORF6 could be readily separated (ref. 9 and Fig. 1*D*). (*ii*) The deduced amino acid composition of the coat protein but not of p24 closely mirrored the composition of the protein from purified BYV virions (4,6).

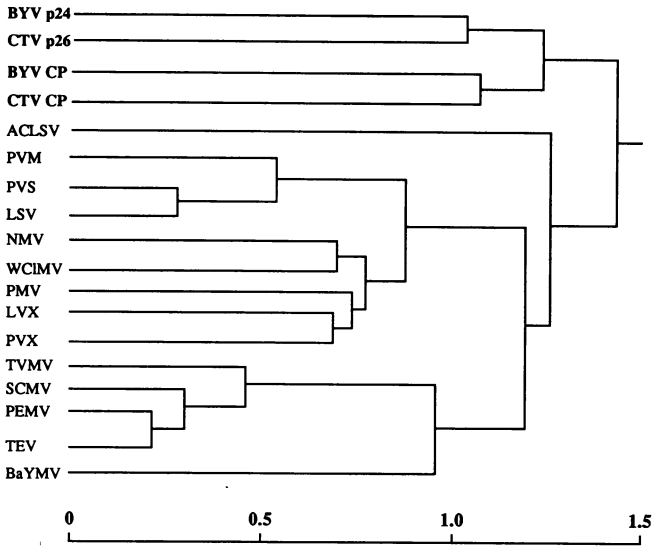

FIG. 3. Cluster dendrogram depicting the relationships between coat proteins of filamentous positive-strand RNA plant viruses. The scale in the bottom shows the evolutionary distance calculated by a modification of the formula of Feng *et al* (27). The parsimony and maximal topological similarity algorithms produced very similar tree topologies (data not shown). Inclusion of the sequences of BYV p24, CTV coat protein, and p26 has led to a minor modification of the branching order as compared to the previously published dendrogram (10). CP, coat protein; BaYMV, barley yellow mosaic virus; LSV, lily simptomless virus; LVX, lily virus X; NMV, narcissus mosaic virus; PEMV, pepper mosaic virus; PMV, papaya mosaic virus; PVM, potato virus M; PVS, potato virus S; SCMV, sugar cane mosaic virus; TEV, tobacco etch virus; TVMV, tobacco vein mottling virus; WClMV, white clover mosaic virus; ACLSV, apple chlorotic leaf spot virus.

All this cannot dismiss the possibility that p24 may be a minor component of the virus particle. However, provided that the capsids of all elongated plant RNA viruses appear to contain only one species of coat protein, it is tempting to propose an alternative function for p24. It can be speculated that p24 transiently associates with the viral RNA, yielding particles that are distinct from virions and serve as the form of the virus spread between cells in the infected plants. The involvement of such nonvirion ribonucleoprotein complexes in cell-to-cell movement of tobacco mosaic virus has been demonstrated (for reviews, see refs. 28 and 29). Naturally, other possibilities, like involvement of p24 in the aphid transmission of BYV, cannot be also excluded.

These findings highlight the duplication of the coat protein gene in viruses with elongated particles. Two groups of positive-strand RNA viruses with icosahedral capsids (namely, animal picornaviruses and plant comoviruses) appear to encode three and two diverged copies of the coat protein, respectively (30, 31). There are, however, two obvious distinctions between these cases of duplication and the coat protein gene duplication in closteroviruses. (*i*) Different species of icosahedral virus coat proteins are present in the virions in equimolar amounts. (*ii*) The similarity between the coat proteins of picorna- and comoviruses is manifested mainly at the tertiary structure level, whereas the level of amino acid sequence similarity is not comparable to that observed for p24 and coat protein of BYV.

Coat protein gene duplication has also been described in DNA-containing bipartite geminiviruses, with one of the copies not being incorporated into the virions (32, 33). Interestingly, the diverged nonvirion copy of the coat protein encoded in DNA-B of these viruses is apparently involved in their systemic spread (34). This phenomenon revealed in a quite different class of viruses shows a provocative analogy with the coat protein gene duplication in the closteroviruses.

The unique gene set of BYV (Fig. 1*A*) provides dramatic illustrations to the module principle of the evolution of RNA virus genomes (35, 36). The putative replication-associated protein of BYV (ORF1) is closely related to the respective proteins of tricornaviruses, which have spherical virions (7). The 6.4-kDa protein (ORF2), p24 (ORF5), and coat protein (ORF6) are related to the respective proteins of filamentous potexviruses, carlaviruses, apple chlorotic leaf spot virus, and potyviruses (refs. 6 and 10 and this paper). Finally, the 65-kDa product of ORF3 is closely related to cellular heat shock proteins of the HSP70 family (7). Analysis of the partial sequence of the CTV genome (H. Pappu *et al.*, personal communication) showed that this gene organization was at least partially conserved in both closteroviruses. Thus, the evolution of the closteroviruses probably involved multiple recombination events between related RNA viruses and between viral and cellular RNAs.

Duplication of nucleotide sequences with subsequent divergence is one of the main principles of the evolution of DNA genomes (37). On the other hand, only a few documented examples of gene duplication in RNA viral genomes could be added to those mentioned above, e.g., the three copies of the genome-linked protein of foot-and-mouth disease virus (38) and the 3C and 2A proteases of enteroviruses and rhinoviruses (39). This difference might be due both to the strict constraints on the size of RNA genomes and to their rapid evolution (40), which in many cases might have hampered the detection of duplication events. Thus, the relatively

high sequence conservation between the coat proteins of closteroviruses and their nonvirion diverged copies is fairly unexpected and adds to our understanding of the evolution of RNA viral genomes. Given the high evolutionary rate, it may be gathered that the duplication in the closterovirus genome is rather recent at the evolutionary scale and/or that the functional similarity between coat proteins and their copies has slowed down their divergence.

We are grateful to Drs. H. Pappu, E. J. Anderson, C. L. Niblett, and R. F. Lee for providing data before publication, to Dr. N. A. Lunina for excellent assistance at some stages of the work, to Dr. J. G. Atabekov for constant interest and encouragement, and to Dr. J. C. Carrington for useful discussions and critical reading of the manuscript.

1. Reanney, D. C. (1982) *Annu. Rev. Microbiol.* **36**, 47–73.
2. Lee, H.-J., Shieh, C.-K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzyahdzhyan, A. & Lai, M. M. C. (1991) *Virology* **180**, 567–582.
3. Bar-Joseph, M. & Murant, A. F. (1982) CMI/AAB Descript. *Plant Viruses*, **260**.
4. Carpenter, J. M., Kassanis, B. & White, R. F. (1977) *Virology* **77**, 101–109.
5. Karasev, A. V., Agranovsky, A. A., Rogov, V. V., Miroshnichenko, N. A., Dolja, V. V. & Atabekov, J. G. (1989) *J. Gen. Virol.* **70**, 241–245.
6. Agranovsky, A. A., Boyko, V. P., Karasev, A. V., Lunina, N. A., Koonin, E. V. & Dolja, V. V. (1991) *J. Gen. Virol.* **72**, 15–23.
7. Agranovsky, A. A., Boyko, V. P., Karasev, A. V., Koonin, E. V. & Dolja, V. V. (1991) *J. Mol. Biol.* **217**, 603–610.
8. Koonin, E. V., Mushegian, A. R., Ryabov, E. V. & Dolja, V. V. (1991) *J. Gen. Virol.* **72**, 2895–2903.
9. Dolja, V. V., Karasev, A. V. & Agranovsky, A. A. (1990) in *New Aspects of Positive-Strand RNA Viruses*, eds. Brinton, M. & Heinz, F. (Am. Soc. Microbiol., Washington), pp. 31–35.
10. Dolja, V. V., Boyko, V. P., Agranovsky, A. A. & Koonin, E. V. (1991) *Virology* **184**, 79–86.
11. Morozov, S. Yu., Miroshnichenko, N. A., Zelenina, D. A., Fedorkin, O. N., Soloviev, A. G., Lukasheva, L. I. & Atabekov, J. G. (1990) *Biochimie* **72**, 677–684.
12. Laemmli, U. K. (1970) *Nature (London)* **27**, 680–685.
13. Miroshnichenko, N. A., Karpova, O. V., Morozov, S. Yu., Rodionova, N. P. & Atabekov, J. G. (1988) *FEBS Lett.* **234**, 65–68.
14. Robinson, D. J., Mayo, M. A., Fritch, C., Jones, A. T. & Raschke, J. H. (1983) *J. Gen. Virol.* **64**, 1591–1599.
15. Sekiya, M. E., Lawrence, S. D., McCaffery, M. & Cline, K. (1991) *J. Gen. Virol.* **72**, 1013–1020.
16. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **28**, 256–268.
17. Leontovich, A. M., Gorbalenya, A. E. & Brodsky, L. I. (1990) *Biopolim. Kletka* **6**, 12–19.
18. Gorbalenya, A. E., Blinov, V. M., Donchenko, A. P. & Koonin, E. V. (1989) *J. Mol. Evol.* **28**, 256–268.
19. Sneath, P. & Sokal, R. (1973) *Principles of Numerical Taxonomy* (Freeman, San Francisco).
20. Felsenstein, J. (1989) *PHYLIP 3.2 Manual.* (Herbarium, Univ. of California, Berkeley).
21. Chumakov, K. M. & Yushmanov, Yu. S. (1988) *Mol. Genet. Mikrobiol. Virusol.* **3**, 3–9.
22. Biou, V., Gibrat, F., Levin, J. M., Robson, B. & Garnier, J. (1988) *Protein Eng.* **2**, 185–191.
23. Brodsky, L. I., Drachev, A. L., Tatuzov, R. L. & Chumakov, K. M. (1991) *Biopolim. Kletka* **7**, 10–14.
24. Lutke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F. & Scheele, G. A. (1987) *EMBO J.* **6**, 43–48.
25. Koenig, R. & Torrance, L. (1986) *J. Gen. Virol.* **67**, 2145–2151.
26. Brunstedt, J., Moseley, J. & Hull, R. (1991) *Virus Genes* **5**, 267–272.
27. Feng, D. F., Johnson, M. S. & Doolittle, R. F. (1985) *J. Mol. Evol.* **21**, 112–125.
28. Atabekov, J. G. & Dorokhov, Yu. L. (1984) *Adv. Virus Res.* **29**, 313–364.
29. Hull, R. (1989) *Annu. Rev. Phytopathol.* **27**, 213–240.
30. Palmenberg, A. C. (1989) in *Molecular Aspects of Picornavirus Infection and Detection*, eds. Semler, B. L. & Ehrenfeld, E. (Am. Soc. Microbiol., Washington), pp. 211–241.
31. Stauffacher, C. V., Usha, R., Harrington, M. Schmidt, T., Hosur, M. V. & Johnson, J. E. (1987) in *Crystallography in Molecular Biology*, eds. Moras, D., Drenth, J., Strandberg, B., Suck, D. & Wilson, K. (Plenum, New York), pp. 293–308.
32. Kikuno, R., Toh, H., Hayashida, H. & Miyata, T. (1984) *Nature (London)* **308**, 562.
33. Mohana Rao, J. K. (1985) *Biochem. Biophys. Res. Commun.* **130**, 892–896.
34. Davies, J. W. & Stanley, J. (1989) *Trends Genet.* **5**, 77–81.
35. Zimmern, D. (1988) in *RNA Genetics*, eds. Holland, J. J., Domingo, E. R. & Ahlquist, P. (CRC, Boca Raton, FL), pp. 211–240.
36. Morozov, S. Yu., Dolja, V. V. & Atabekov, J. G. (1989) *J. Mol. Evol.* **29**, 52–62.
37. Ohno, S. (1970) *Evolution by Gene Duplication* (Springer, Berlin).
38. Forss, A. & Schaller, H. (1982) *Nucleic Acids Res.* **10**, 6441–6450.
39. Bazan, J. F. & Fletterick, R. J. (1990) *Seminars Virol.* **1**, 311–322.
40. Steinhauer, D. & Holland, D. J. (1987) *Annu. Rev. Microbiol.* **41**, 409–433.