Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 8259–8263, September 1992
Microbiology

Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: Delineation of an additional group of positive-strand RNA plant and animal viruses

(RNA-dependent RNA polymerase/RNA helicase/protease/methyltransferase/polyprotein organization)

EUGENE V. KOONIN*†, ALEXANDER E. GORBALENYA‡, MICHAEL A. PURDY§, MIKHAIL N. ROZANOV¶‖, GREGORY R. REYES**, AND DANIEL W. BRADLEY§

*Institute of Microbiology, Academy of Sciences, 117811, Moscow, Russia; ‡Institute of Poliomyelitis and Viral Encephalitides, Academy of Medical Sciences, 142782, Moscow Region, Russia; §Hepatitis Branch, Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control, 1600 Clifton Road N.E., Atlanta, GA 30333; ¶Institute of Molecular Biology, Academy of Sciences, Moscow, Russia; and **Genelabs Inc., 505 Penobscot Drive, Redwood City, CA 95064

Communicated by Sherman M. Weissman, March 17, 1992

ABSTRACT **Computer-assisted comparison of the nonstructural polyprotein of hepatitis E virus (HEV) with proteins of other positive-strand RNA viruses allowed the identification of the following putative functional domains: (*i*) RNA-dependent RNA polymerase, (*ii*) RNA helicase, (*iii*) methyltransferase, (*iv*) a domain of unknown function ("X" domain) flanking the papain-like protease domains in the polyproteins of animal positive-strand RNA viruses, and (*v*) papain-like cysteine protease domain distantly related to the putative papain-like protease of rubella virus (RubV). Comparative analysis of the polymerase and helicase sequences of positive-strand RNA viruses belonging to the so-called "alpha-like" supergroup revealed grouping between HEV, RubV, and beet necrotic yellow vein virus (BNYVV), a plant furovirus. Two additional domains have been identified: one showed significant conservation between HEV, RubV, and BNYVV, and the other showed conservation specifically between HEV and RubV. The large nonstructural proteins of HEV, RubV, and BNYVV retained similar domain organization, with the exceptions of relocation of the putative protease domain in HEV as compared to RubV and the absence of the protease and X domains in BNYVV. These observations show that HEV, RubV, and BNYVV encompass partially conserved arrays of distinctive putative functional domains, suggesting that these viruses constitute a distinct monophyletic group within the alpha-like supergroup of positive-strand RNA viruses.**

Hepatitis E virus (HEV) is the causative agent of the enterically transmitted form of non-A, non-B hepatitis, a disease of significant epidemiologic importance (1, 2). HEV has been found in stools from hepatitis patients as naked isometric virus-like particles of about 32–34 nm in diameter (3, 4). It has been shown that the HEV genome is a single-stranded polyadenylylated RNA of about 7.5 kilobases (5). Recently the HEV genome has been cloned as cDNA (5, 6), and its complete nucleotide sequence has been determined (7, 8). The positive-sense RNA of HEV encompasses three large open reading frames (ORFs): the largest ORF (5′ end) consists of 1693 codons, the second ORF (3′ end) is composed of 660 codons, and the third ORF consists of 123 codons that overlap the first and the second ORFs. By analogy with other positive-strand RNA animal viruses, such as alphaviruses (9), rubella virus (RubV), (10) and caliciviruses (11), it was tempting to speculate that the 660-codon ORF encodes the capsid protein(s) of HEV, whereas the product of the largest 5′ ORF should be the nonstructural polyprotein, which is probably cleaved to yield functional viral proteins (7, 8). The function of the smallest ORF is not known, but there are indications that it is expressed in infected humans (12).

Comparative analyses of the sequences of nonstructural proteins encoded by positive-strand RNA viruses have yielded a considerable collection of functional domains with varying degrees of conservation. Various subsets of this "pool" are combined in polyproteins (or large noncleaved proteins) of different viruses to form partially conserved arrays (13–16). Knowledge of the sequence motifs conserved in each of these domains and of their typical arrangements may now allow assignment of putative functions to a reasonable fraction of the nonstructural protein sequences from newly characterized but unclassified viruses.

Comparison of the amino acid sequence of the putative RNA-dependent RNA polymerase of HEV with those of other positive-strand RNA viruses revealed the closest similarity with RubV and, unexpectedly, with beet necrotic yellow vein virus (BNYVV), a plant furovirus. Phylogenetic analysis showed that the polymerases of these three viruses comprised a compact group within the so-called "alpha-like" supergroup of positive-strand RNA viruses (17). With this finding, it was of interest to initiate a detailed comparison of the putative genome products of HEV and other positive-strand RNA viruses in order to identify functional domains in the HEV polyprotein and to gain further insight into the relationships between HEV, RubV, and BNYVV.

METHODS

Amino Acid Sequences. The HEV sequence (Burma strain) was from ref. 7, the RubV sequence was from ref. 10, and the EAV sequence was from ref. 18. All other sequences were from the Swiss-Prot data bank (Release 16).

Comparative Sequence Analysis. Initial pairwise comparison of amino acid sequences was done using the program DOTHELIX (19), which generates complete diagonal local similarity plots. The results produced by this program were used to delineate the boundaries of the regions to be aligned by the program OPTAL as described (16) using the amino acid residue comparison matrix MDM78. OPTAL generates multiple sequence alignments in a stepwise manner and calculates

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: HEV, hepatitis E virus; ORF, open reading frame; RubV, rubella virus; BNYVV, beet necrotic yellow vein virus; EAV, equine arteritis virus; nsP*x*, nonstructural protein *x*.

†To whom reprint requests should be addressed at: National Center for Biotechnology Information, National Library of Medicine, Building 38A, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD 20894.

‖Present address: Institut Jacques Monod, 2 Place Jussieu, 75005, Paris, France.

adjusted alignment scores as the number of standard deviations (SD) over the mean of 25 random simulations. Search of amino acid sequences for conserved motifs was done using the program SITE (20). Programs DOTHELIX and SITE are modules of the GeneBee package for biopolymer sequence analysis (21).

RESULTS AND DISCUSSION

Identification of Functional Domains in the Nonstructural Polyprotein of HEV. Functional mapping of the nonstructural polyprotein of HEV employed both direct sequence comparison and search for specific conserved motifs. Local similarity searches were done with nonstructural proteins of all other positive-strand RNA viruses. In accord with the results of the polymerase comparisons (17), these searches revealed the highest level of correlation between the nonstructural polyproteins of HEV and RubV. This was the only comparison that revealed four segments of high similarity (over 6.5 SD) grouping close to a single diagonal on the local similarity plot (data not shown). The highest conservation was observed for one of the previously described motifs in the RNA-dependent RNA polymerase, whereas the three remaining segments corresponded to additional conserved domains (see below). Comparison of the HEV polyprotein with the large nonstructural protein encoded by RNA 1 of BNYVV revealed segments of significant similarity predominantly within the polymerase. In addition, we searched the HEV polyprotein sequence for segments with the best correspondence to the amino acid motifs conserved in positive-strand RNA viral RNA helicases, various types of proteases, and putative methyltransferases. These searches resulted in tentative identification of the following functional domains in the nonstructural polyprotein of HEV (described in the order of decreasing sequence conservation).

RNA-dependent RNA polymerase domain. The putative polymerases of HEV, BNYVV, and RubV (Fig. 1) form a distinct tight group within the so-called supergroup III of viral RNA polymerases (17). Alignment of the HEV polymerase sequence with either RubV or BNYVV polymerase yielded highly significant scores of >10 SD. Despite the impressive overall conservation among the three viral polymerases, HEV polymerase encompassed a striking deviation from the supergroup III consensus, having alanine instead of the conserved cysteine in motif VII (Fig. 1).

RNA helicase domain. The putative RNA helicase of HEV obviously contains all the seven conserved segments typical of the so-called helicase superfamily I (refs. 22 and 23; Fig. 2*A*). Comparison with other positive-strand RNA viral helicases belonging to this superfamily showed the highest overall similarity with the helicase of BNYVV (8.6 SD), whereas somewhat less pronounced similarity (6.8 SD) was observed with the RubV helicase. In addition, a region of striking resemblance spanning the C-terminal motifs IV–VI (15.2 SD for the alignment shown in Fig. 2*B*) was revealed between the putative helicases of HEV and equine arteritis virus (EAV), which belongs to a recently described superfamily of positive-strand RNA viruses, together with corona- and toroviruses (18). The N-terminal domains of the helicases of HEV and EAV were much less closely related. Comprehensive comparative analysis of the helicases of alpha-like viruses using various methods for phylogenetic tree construction revealed grouping between HEV, RubV, and BNYVV (E.V.K. and M.N.R., unpublished observations).

Putative methyltransferase domain. We have shown recently that a degree of conservation could be revealed in the N-terminal regions of large replicative proteins of all viruses belonging to the alpha-like supergroup of positive-strand RNA viruses (24). Functional studies on nonstructural protein 1 (nsP1) harboring the conserved N-terminal domain in

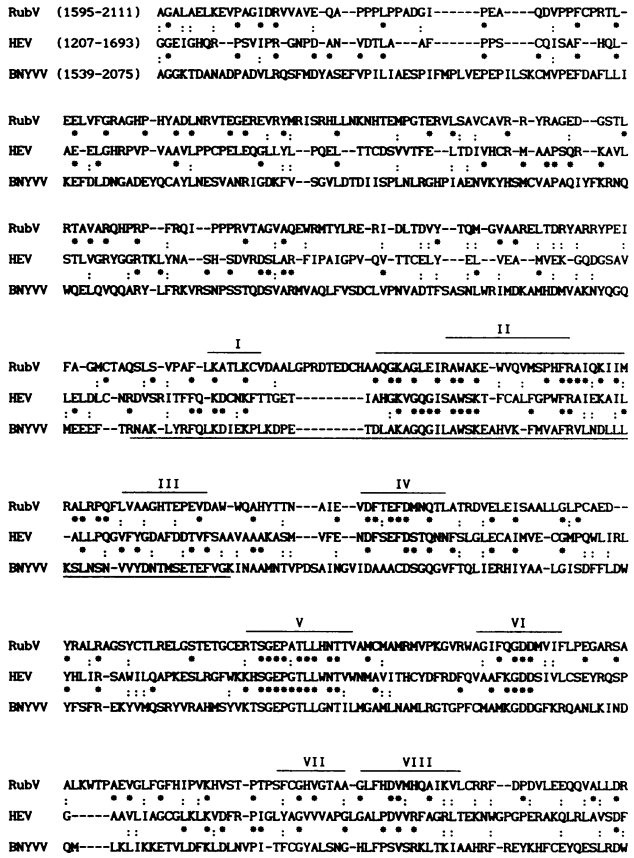

FIG. 1. Alignment of the putative RNA-dependent RNA polymerase domains of HEV, RubV, and BNYVV. The conserved motifs of positive-strand RNA viral RNA polymerases are designated I–VIII as in ref. 17. Asterisks indicate identical amino acid residues, and colons indicate related residues. Grouping of amino acid residues by physico-chemical similarity was as follows: G and A; S and T; D, E, N, and Q; K and R; I, L, V, and M; and F, Y, and W. The region of highest similarity between HEV and RubV revealed upon local similarity search is overlined, and the region of highest similarity between HEV and BNYVV is underlined. In this figure and Figs. 2, 4, and 5, the locations of the aligned sequences in viral polyproteins are indicated by the numbers of the first and last residues in parentheses.

alphaviruses have shown that it encompasses at least the methyltransferase and guanylyltransferase activities required for capping of viral plus-strand RNA (25, 26). Mutational analysis of nsP1 showed that motif II (Fig. 3) may be directly involved in the methyltransferase activity (27).

Surprisingly, the putative methyltransferase domain of HEV showed the closest similarity to those of tricornaviruses (alignment in Fig. 3 scoring 13.3 SD), not to the respective sequences of RubV or BNYVV, although a clear resemblance to the latter was also retained (data not shown).

"Y" domain. A region of significant similarity between HEV, RubV, and BNYVV spanned ≈200 amino acid residues downstream of the methyltransferase domain (Fig. 4). We coined the designation Y domain for these sequences. A subsequence of this region scored second highest upon comparison of the entire polyproteins of HEV and RubV (Fig. 4). Alignments of the sequences of this domain between HEV and either of the other two viruses scored over 8 SD, indicating a genuine relationship. In the polyproteins of HEV and RubV, the N terminus of the Y domain overlapped the

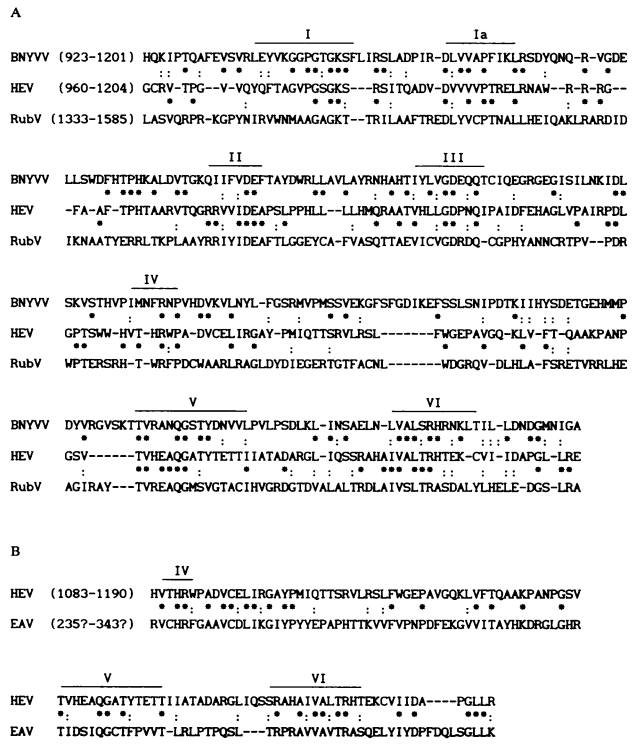

FIG. 2. Alignment of the putative RNA helicase domains of HEV, RubV, BNYVV, and EAV. The conserved motifs typical of the helicase superfamily I are designated I–VI as in ref. 16. (*A*) The sequences of the helicases of HEV, RubV, and BNYVV were aligned over regions I–VI. (*B*) The sequences of the helicases of HEV and EAV were aligned only within the regions spanning motifs IV–VI. The other designations are as in Fig. 1.

conserved motif III at the C terminus of the methyltransferase domain (compare Figs. 3 and 4). On the other hand,

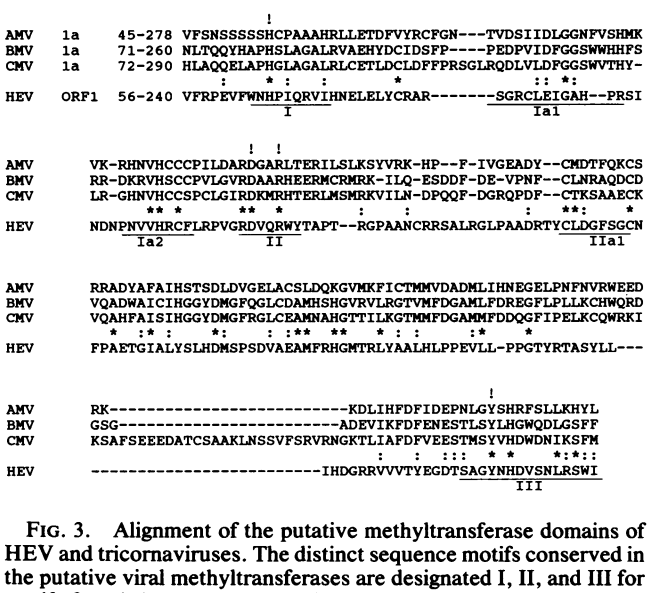

FIG. 3. Alignment of the putative methyltransferase domains of HEV and tricornaviruses. The distinct sequence motifs conserved in the putative viral methyltransferases are designated I, II, and III for motifs found throughout the entire alpha-like supergroup; and Ia1, Ia2, and IIa1 for motifs detected in a subset of the alpha-like viruses including alphaviruses, tobamoviruses, tobraviruses, hordeiviruses, tricornaviruses, HEV, BNYVV, and RubV (24). Exclamation marks denote amino acid residues conserved throughout the alpha-like supergroup, stars denote identities between HEV and the three tricornaviruses, and colons denote similar residues in HEV and tricornaviruses. The numbers of the first and last residues are given before the sequence. Viruses: AMV, alfalfa mosaic; BMV, brome mosaic; CMV, cucumber mosaic.

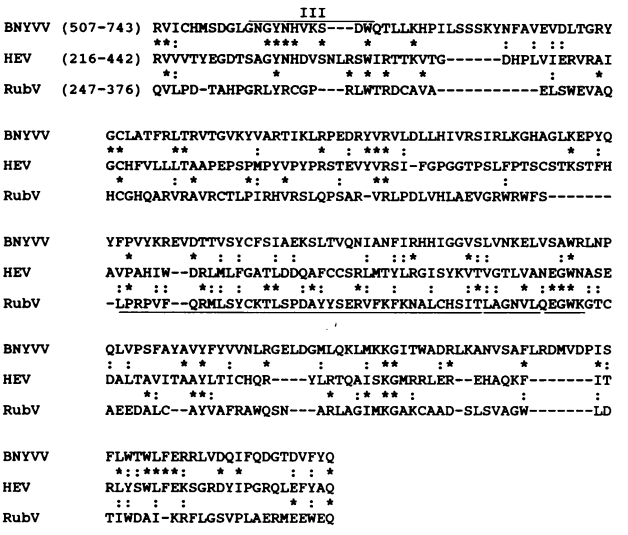

FIG. 4. Alignment of Y domains of HEV, RubV, and BNYVV. Motif III comprising the C-terminal extremity of the putative methyltransferase domain in HEV and RubV (see also Fig. 3 and text) is designated. The region of highest similarity between the HEV and RubV sequences revealed upon local similarity search is underlined. Stars indicate identical residues, and colons indicate similar residues.

BNYVV appeared to contain a duplication of an ≈100-amino acid residue segment starting at motif III (M.N.R., E.V.K., and A.E.G., unpublished observations). Thus in this virus the Y domain sequence was separated from the C terminus of the methyltransferase domain by 88 residues. No appreciable similarity, beyond motif III, could be revealed between the Y domain and the respective portions of large replicative proteins of other alpha-like viruses.

"X" domain. A conserved domain of unknown function flanks the papain-like protease domains in the polyproteins of positive-strand RNA animal viruses (28). A segment located within the X domain scored third highest upon comparison of HEV and RubV polyproteins (Fig. 5*A*). Pronounced conservation was observed among the X domains of HEV, RubV, and alphaviruses (Fig. 5*A*), with alignment scores of 10.7 SD (HEV vs. aligned sequences of alphaviruses) and 11.4 SD (RubV vs. aligned sequences of alphaviruses and HEV). Curiously, a less striking score of 5.0 SD was observed for the alignment of the X-domain sequences of HEV and RubV. A more distant, albeit reliable, relationship was found with X domains of coronaviruses (data not shown; see ref. 28).

In RubV and HEV, the X domain is preceded by a proline-rich region (alignment score, 5.1 SD; Fig. 5*B*). This portion of the polyprotein might constitute a flexible hinge between the X domain and the upstream domains.

Putative papain-like cysteine protease. Alphaviruses and RubV encode (putative) papain-like proteases mediating processing of nonstructural polyproteins (28). Conservation of the X domain, which has been found to date exclusively in association with (putative) viral papain-like proteases (28), strongly suggested that HEV encodes a related protease too. The results of mapping putative functional domains in HEV polyprotein left virtually no room to accommodate the protease, other than a portion of the nonstructural polyprotein composed of ≈300 amino acid residues and bounded by the C terminus of the Y domain and the N terminus of the "proline hinge." Detailed comparison of this sequence with various cellular and viral papain-like proteases revealed a region with moderate similarity to the putative protease of RubV (Fig. 5*C*). This alignment scored only 2.6 SD; however, the conservation of a short stretch of amino acid residues around the putative catalytic cysteine and of two

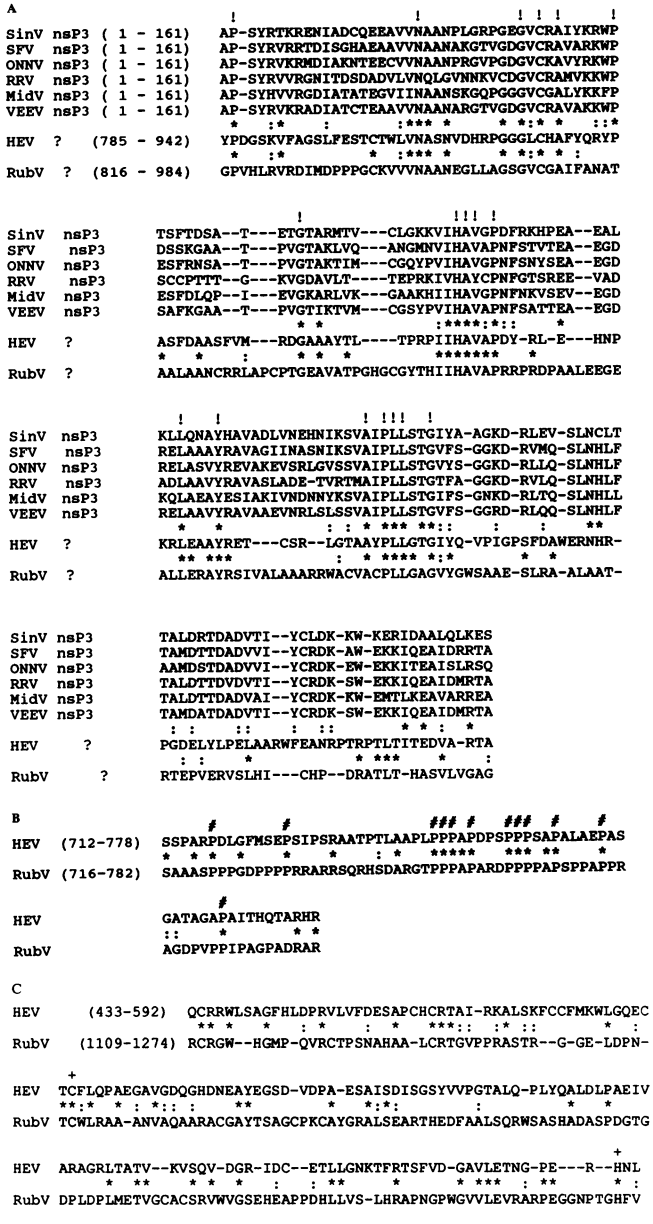

FIG. 5. The putative papain-like proteases, X domains, and proline-rich hinge domains of HEV and RubV. (*A*) Alignment of the X domains of HEV, RubV, and alphaviruses. Exclamation marks denote invariant residues, stars denote identities between HEV and RubV or between HEV and alphaviruses, and colons denote similar residues between HEV and RubV or between HEV and alphaviruses. In each case, an exception in one of the alphavirus sequences was allowed. Viruses: SinV, Sindbis; SFV, Semliki Forest; ONNV, O'Nyong-Nyong; RRV, Ross River; MidV, Middleburg; VEEV, Venezuelan equine encephalitis. (*B*) Alignment of the proline-rich hinge domains of HEV and RubV; proline residues are highlighted with a pound sign. Stars indicate identical residues, and colons indicate similar residues. (*C*) Alignment of the putative papain-like proteases of HEV and RubV. Plus signs indicate putative catalytic residues, stars indicate identical residues, and colons indicate similar residues.

additional cysteine residues was notable. No appreciable similarity could be found between this region of the HEV polyprotein and other viral or cellular papain-like proteases. It should be noted that the sequences of the other (putative) viral papain-like proteases also show only very modest conservation (28).

Despite the absence of a rigorous statistical argument, we believe that the region of the HEV polyprotein shown in Fig. 5*C* is probably the genuine protease of HEV. In addition to the similarity with the putative protease of RubV and the correlation between papain-like protease and X domains, there is an indication of more general supportive evidence for this hypothesis. All positive-strand RNA animal viruses studied in some detail to date encode proteases mediating viral polyprotein processing (28–31). We have extensively searched the HEV polyprotein sequence for segments resembling all known types of proteases but found no sequences with reasonable resemblance to known consensus patterns except those shown in Fig. 5*C*. Thus this seems to be the best possible identification of the HEV protease.

Comparison of Genome Organizations of HEV and Other Positive-Strand RNA Viruses: The Peculiarities of Virus Evolution. Comparative amino acid sequence analysis allowed us to assign a specific function or at least reveal sequence conservation with proteins of other viruses for about 1500 of 1693 residues of the nonstructural polyprotein of HEV. This highlights a qualitatively new situation in contemporary molecular virology when the sequence information contained in genomes of viruses belonging to additional groups can be deciphered to a significant extent by using the knowledge gained through computer-assisted comparative analysis of other virus genomes.

As indicated above, phylogenetic analysis of the polymerases and the helicases consistently grouped together HEV, RubV, and BNYVV. This suggested that these viruses could have evolved from a common ancestor virus and should be considered a distinct group within the alpha-like supergroup of positive-strand RNA viruses. This is compatible with the present identification of the Y domain, which seemed to be unique for this group. On the other hand, in at least three portions of the nonstructural polyprotein of HEV, the putative methyltransferase, the C-terminal part of the helicase, and the X domain, the highest similarity was observed with protein sequences of viruses that do not belong to this group. Thus in some cases the evolution of different domains within a single viral polyprotein might not be congruent.

Two groups of positive-strand RNA viruses that bring together plant and animal viruses and are characterized by conservation of arrays of related functional domains have been described. These include (*i*) animal picorna- and caliciviruses and plant como- and nepoviruses and (*ii*) animal alphaviruses and plant tobamo-, tobra-, and hordeiviruses (13, 14). The observations described here further support the notion that close evolutionary relationships between plant and animal viruses are typical of this virus class.

Comparison of the organization of the putative conserved functional domains in viral polyproteins showed that not only the nonstructural polyproteins of HEV and RubV but also those of phylogenetically distant alphaviruses are mainly colinear (Fig. 6). The principal differences lie in the mutual arrangement of individual domains in the cluster protease–X–helicase. By taking HEV as a standard, RubV shows a relocation of the protease domain, whereas in alphaviruses the helicase domain is relocated. On the other hand, the plant furovirus BNYVV lacked the protease–X block altogether (Fig. 6). This distinction may reflect a consistent difference in the expression strategies of plant and animal viruses.

The processing scheme of alphavirus nonstructural polyproteins has been described in detail, and it has been shown that the methyltransferase, X, and polymerase domains reside within distinct proteins (nsP1, nsP3, and nsP4, respectively), whereas the helicase and the protease are fused in nsP2 protein (Fig. 6; refs. 9 and 34). It remains unclear whether HEV and/or RubV follow analogous expression strategies. Extrapolating the information available for other positive-strand RNA animal viruses (13–16), it is reasonable to hypothesize that at least the helicase and the polymerase functions are probably contained in different proteins.

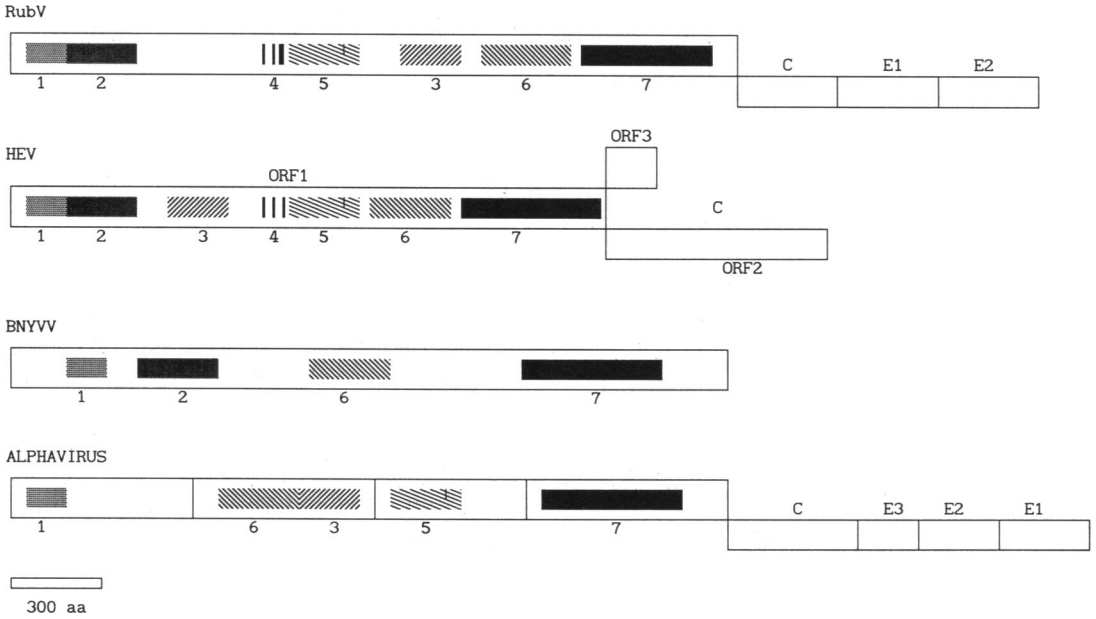

FIG. 6. Schematic representation of the domainal organization of proteins encoded by the genomes of HEV, RubV, BNYVV, and alphaviruses. Viral proteins are drawn to the scale shown in the bottom of the figure. The conserved domains are numbered as follows: 1, putative methyltransferase domain; 2, Y domain; 3, putative papain-like protease; 4, proline-rich hinge domain; 5, X domain linked to the viral papain-like proteases; 6, putative RNA helicase; 7, putative RNA polymerase; C, capsid protein; E1, E2, and E3, glycoproteins. Regions showing sequence conservation are highlighted by identical hatching. For BNYVV, only the scheme for the large nonstructural protein encoded in RNA 1 is shown. aa, amino acid. The coat protein and additional nonstructural proteins are encoded in RNA2 of this virus (32, 33).

The colinearity of genome organization of HEV, RubV, and alphaviruses extends, at least formally, into the 3′ ORF encoding the virion components. The major distinction is that HEV lacks the glycoprotein genes found both in RubV and in alphaviruses (Fig. 6). In this respect HEV resembles caliciviruses whose 3′ ORF encodes a single capsid protein species. Curiously, the size of the 3′ structural ORF is very similar in HEV and feline calicivirus (11). Comparison of the amino acid sequences of capsid proteins revealed only marginal similarities, at best, between HEV and either RubV or the calicivirus (not shown). Thus the possibility remains that HEV could evolve either by truncation of a RubV-like genome or by recombination with a calicivirus-like genome. Alternatively, HEV may be regarded as the ancestral form, whereas the glycoprotein genes could have been acquired by RubV via recombination. Accumulation of additional viral genome sequences and elucidation of virion crystal structures will be helpful in further clarifying the evolutionary relationships between HEV and other positive-strand RNA viruses.

1. Bradley, D. W. & Maynard, J. E. (1986) *Semin. Liver Dis.* **6**, 56–66.
2. Purcell, R. H. & Ticehurst, J. R. (1988) in *Viral Hepatitis and Liver Disease*, ed. Zuckerman, H. A. (Liss, New York), pp. 131–137.
3. Bradley, D. W., Krawczynski, K., Cook, E. H., Jr., McCaustland, K. A., Humphrey, C. D., Spelbring, J. E., Myint, H. & Maynard, J. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6277–6281.
4. Krawczynski, K. & Bradley, D. W. (1989) *J. Infect. Dis.* **159**, 1042–1049.
5. Reyes, G. R., Purdy, M. A., Kim, J. P., Luk, K.-C., Young, L., Fry, K. E. & Bradley, D. W. (1990) *Science* **247**, 1335–1339.
6. Reyes, G., Purdy, M. A., Kim, J. P., Young, L. M., Fry, K. E., Luk, K.-C. & Bradley, D. W. (1990) in *Excerpta Medica*, eds. Sung, J. L. & Chen, D.-S. (Elsevier, New York), pp. 249–255.
7. Tam, A. W., Smith, M. W., Guerra, M. E., Huang, Ch.-Ch., Bradley, D. W. & Reyes, G. R. (1991) *Virology* **185**, 120–131.
8. Fry, K. E., Tam, A. W., Smith, M. M., Kim, J. P., Luk, K.-C., Young, L. M., Paatak, M., Feldman, R. A., Yun, K. Y., Purdy, M. A., McCaustland, K., Bradley, W. & Reyes, G. R. (1991) *Virus Genes*, **6**, 173–185.
9. Strauss, E. G., Rice, C. M. & Strauss, J. H. (1984) *Virology*, **133**, 92–110.
10. Dominguez, G., Wang, C.-Y. & Frey, T. K. (1990) *Virology* **177**, 225–238.
11. Neill, J. D. (1990) *Virus Res.* **17**, 145–160.
12. Yarbough, P. O., Tam, A., Krawczynski, K., Fry, K. E., McCaustland, K. A., Miller, A., Fernandez, J., Huang, C.-C., Bradley, D. W. & Reyes, G. R. (1991) *J. Virol.* **65**, 5790–5797.
13. Zimmern, D. (1988) in *RNA Genetics*, eds. Holland, J. J., Domingo, E. R. & Ahlquist, P. (CRC, Boca Raton, FL), pp. 211–240.
14. Goldbach, R. & Wellink, J. (1988) *Intervirology* **29**, 260–267.
15. Morozov, S. Yu., Dolja, V. V. & Atabekov, J. G. (1989) *J. Mol. Evol.* **29**, 52–62.
16. Gorbalenya, A. E., Blinov, V. M., Donchenko, A. P. & Koonin, E. V. (1989) *J. Mol. Evol.* **28**, 256–268.
17. Koonin, E. V. (1991) *J. Gen. Virol.* **72**, 2197–2206.
18. den Boon, J. A., Snijder, E. J., Chirnside, E. D., de Vries, A. A. F., Horzinek, M. C. & Spaan, W. J. (1991) *J. Virol.* **65**, 2910–2920.
19. Leontovich, A. M., Brodsky, L. I. & Gorbalenya, A. E. (1990) *Biopolimery i Kletka* **6** (6), 12–19 (in Russian).
20. Koonin, E. V., Chumakov, K. M. & Gorbalenya, A. E. (1990) *Biopolymery i Kletka* **6** (6), 43–48 (in Russian).
21. Brodsky, L. I., Drachev, A. L., Tatuzov, R. L. & Chumakov, K. M. (1991) *Biopolimery i Kletka* **7** (1) 10–14.
22. Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P. & Blinov, V. M. (1988) *FEBS Lett.* **239**, 16–24.
23. Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P. & Blinov, V. M. (1989) *Nucleic Acids Res.* **17**, 4713–4730.
24. Rozanov, M. N., Koonin, E. V. & Gorbalenya, A. E. (1992) *J. Gen. Virol.*, in press.
25. Mi, S., Durbin, R., Huang, H. V., Rice, C. M. & Stollar, V. (1989) *Virology* **170**, 385–391.
26. Scheidel, L. M. & Stollar, V. (1991) *Virology* **181**, 490–499.
27. Mi, S. & Stollar, V. (1991) *Virology* **184**, 423–427.
28. Gorbalenya, A. E., Koonin, E. V. & Lai, M. M. C. (1991) *FEBS Lett.* **288**, 201–205.
29. Krausslich, H.-G. & Wimmer, E. (1988) *Annu. Rev. Biochem.* **57**, 701–756.
30. Gorbalenya, A. E., Blinov, V. M., Donchenko, A. P. & Koonin, E. V. (1989) *FEBS Lett.* **243**, 103–114.
31. Gorbalenya, A. E., Donchenko, A. P., Koonin, E. V. & Blinov, V. M. (1989) *Nucleic Acids Res.* **17**, 3889–3897.
32. Bouzoubaa, S., Ziegler, V., Becj, D., Guilley, H., Richards, K. & Jonard, G. (1986) *J. Gen. Virol.* **67**, 1689–1700.
33. Bouzoubaa, S., Quillet, L., Guilley, H., Jonard, G. & Richards, K. (1987) *J. Gen. Virol.* **68**, 615–626.
34. Hardy, W. R. & Strauss, J. H. (1989) *J. Virol.* **63**, 4653–4664.