

Structure and Expression of Large (+)RNA Genomes of Viruses of Higher Eukaryotes

Alexey A. Agranovsky

Faculty of Biology, Lomonosov Moscow State University, 119234 Moscow, Russia
e-mail: etdonas2@gmail.com

Received August 17, 2020

Revised September 21, 2020

Accepted September 29, 2020

Abstract—Viral positive-sense RNA genomes evolve rapidly due to the high mutation rates during replication and RNA recombination, which allowing the viruses to acquire and modify genes for their adaptation. The size of RNA genome is limited by several factors, including low fidelity of RNA polymerases and packaging constraints. However, the 12-kb size limit is exceeded in the two groups of eukaryotic (+)RNA viruses – animal nidoviruses and plant closteroviruses. These virus groups have several traits in common. Their genomes contain 5'-proximal genes that are expressed via ribosomal frameshifting and encode one or two papain-like protease domains, membrane-binding domain(s), methyltransferase, RNA helicase, and RNA polymerase. In addition, some nidoviruses (i.e., coronaviruses) contain replication-associated domains, such as proofreading exonuclease, putative primase, nucleotidyltransferase, and endonuclease. In both nidoviruses and closteroviruses, the 3'-terminal part of the genome contains genes for structural and accessory proteins expressed via a nested set of coterminal subgenomic RNAs. Coronaviruses and closteroviruses have evolved to form flexuous helically symmetrical nucleocapsids as a mean to resolve packaging constraints. Since phylogenetic reconstructions of the RNA polymerase domains indicate only a marginal relationship between the nidoviruses and closteroviruses, their similar properties likely have evolved convergently, along with the increase in the genome size.

DOI: 10.1134/S0006297921030020

Keywords: viral positive-sense RNA genomes, nidoviruses, SARS-CoV, closteroviruses, evolution, gene expression

INTRODUCTION

Rapid evolution of viral RNA genomes, which is due to the high mutation rates and gene shuffling during RNA replication, leads to the acquisition and modification of genes involved in virus adaptation. The majority of RNA viruses have compact genomes 4 to 12 kb in size [1]. The size of an RNA genome is limited by a number of factors, the major one being the low replication fidelity. Viral RNA-dependent RNA polymerases introduce approximately 10^{-4} errors per nucleotide, which is several orders of magnitude more than the error rate of DNA polymerases [2, 3]. It is believed that accumulation of unfavorable mutations during replication of large RNA

genomes may define a threshold for the reproduction of viable virus variants [4]. In addition, the size of viral RNA influences its stability and packaging [5].

In the course of evolution, RNA genomes of animal nidoviruses (families *Arteriviridae*, *Coronaviridae*, *Roniviridae*, and *Mesoniviridae*; order Nidovirales) and plant closteroviruses (fam. *Closteroviridae*) have exceeded the 12-kb limit. Coronaviruses have the largest undivided RNA genomes among all known (+)RNA viruses (up to 41 kb [6]). The size of closteroviral genomes ranges from 14.5 to 19 kb [7, 8]. This review is focused on the structure, encapsidation, replication, and expression of large RNA genomes. Comparison of nidoviruses and closteroviruses reveals paradoxical similarities between these evolutionary distant groups of animal and plant viruses.

CORONAVIRUSES

Members of the order Nidovirales – arteriviruses, coronaviruses, mesoniviruses, and roniviruses – show considerable variation in the genome size and structure.

Abbreviations: BYV, beet yellows closterovirus; DMV, double-membrane vesicle; MHV, murine hepatitis virus; MVC, multi-vesicular complex; nsp, nonstructural protein; ORF, open reading frame; pp, polyprotein; RFs, ribosomal frameshifting signal; SARS-CoV, severe acute respiratory syndrome coronavirus; sgRNA, subgenomic RNA; TRS, transcription regulatory signal.

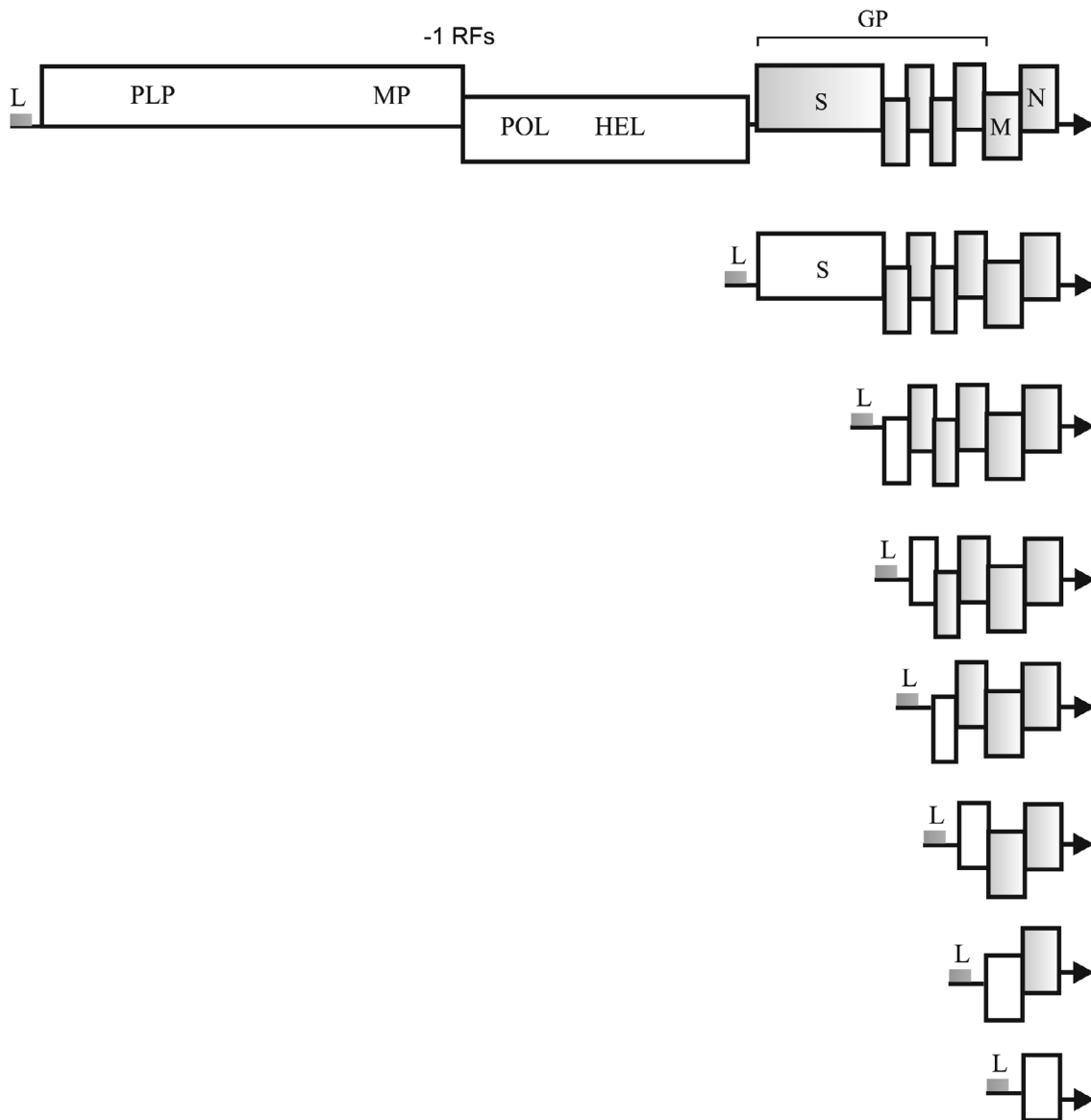


Fig. 1. Genomic RNA and sgRNAs of SARS-CoVs: L, leader sequence; -1 RFs, -1 ribosomal frameshifting signal. Encoded proteins and protein domains: PLP, papain-like cysteine protease; MP, (main) serine protease; POL, RNA polymerase; HEL, RNA helicase; S, spike glycoprotein; GP, accessory proteins and outer membrane glycoprotein; M, matrix protein; N, nucleocapsid protein. Genes unavailable for translation in each type of mRNA are shown as shaded boxes. Arrows indicate 3'-ends of RNAs. Drawn approximately to scale.

Comparison of amino acid sequences of the most conserved virus enzyme, RNA polymerase, indicates that nidoviruses form a compact phylogenetic cluster [9] in the picorna-like phylum [10]. This review is limited to the discussions of *Coronaviridae* viruses only as the most studied and epidemiologically important family of the Nidovirales order.

Genome structure and expression. Coronaviruses have helically symmetrical nucleocapsids surrounded by the lipoprotein membrane containing the spike (S) glycoprotein and other viral glycoproteins [9]. The genomic RNA has the 5'-terminal cap and the 3'-terminal poly(A) tract. ORFs 1a and 1b, coding for the replicase compo-

nents, occupy the 5'-proximal portion of the genome (Fig. 1). Translation of these genes via the -1 ribosomal frameshifting yields 1a and 1ab polyproteins (pps) of ~4,000 and ~7,000 aa, respectively, at an approximate ratio of 4 : 1. The processing of these polyproteins by the viral proteases yields 16 nonstructural proteins (nsps). The genes for the structural proteins – membrane glycoproteins, matrix protein (M), and nucleocapsid protein (N) – map to the 3'-terminal genome portion and are expressed via a nested set of 3'-coterminal subgenomic RNAs (sgRNAs) (Fig. 1). The genomic RNA and sgRNAs have identical 5'-terminal leader sequences (L) of 60-90 nt.

The 3'-terminal genes for the accessory and structural proteins are expressed via sgRNAs. In each sgRNA, only the most 5'-terminal gene is available for the ribosomes (Fig. 2). The sgRNAs are synthesized by the unique discontinuous transcription mechanism [11, 12]. Upstream of each viral ORF (except ORF 1b) in the genomic RNA, there is a conserved transcriptional regulatory signal (TRS) 6 to 8 nt in length: L-TRS (leader TRS) or B-TRS (body TRS) (Fig. 2). The synthesis of antigenomic (–)RNA might stall at the B-TRS in the template genomic (Fig. 2). Then the (–)RNA strand “jumps” to the 5'-terminal leader of the template strand or, more likely, the leader and the (–)RNA are brought in proximity to each other due to the looping of the template (+)RNA. This is followed by annealing of the anti-B-TRS and L-TRS and completion of the (–)RNA strand on the (+)L template (Fig. 2a). The resulting anti-sgRNAs, containing common anti-L sequence, serve as templates for the generation of sgRNAs. Phosphorylated free nucleocapsid protein N⁰ binds to the B-TRS and recruits cell helicase DDX1. These interactions at the late stage of the infection cycle allow the replicative complex to bypass the TRSs in the template genomic RNA and to synthesize the full-length (–)RNA serving as a template

for the generation of progeny (+)strands (Fig. 2b) [13, 14].

Replication-associated proteins. The replicative complex of coronaviruses consists of 16 nonstructural proteins produced by the processing of pp1a and pp1ab by the viral proteases (Fig. 3). Closely related coronaviruses may have one or two PLP domains [9]. In mouse hepatitis virus (MHV), the autocatalytic release of nsp1, nsp2, nsp3, and nsp4 is carried out by two PLP domains, whereas in SARS-CoV-1, SARS-CoV-2, MERS (Middle East respiratory syndrome)-CoV, and infectious bronchitis virus (IBV), these cleavages are mediated by a single PLP (Figs. 1 and 3). It is possible that the PLP domains have originated by duplication and then diverged in the course of coronavirus evolution. The majority of the cleaved bonds in the coronaviral polyproteins are hydrolyzed *in trans* by the main chymotrypsin-like proteinase (MP) [9]. Specific functions of coronaviral nsp1 and nsp2 are apparently associated with the degradation of cell mRNAs and inhibition of their translation [15, 16] and with the maturation of viral particles [17]. Point mutations blocking the nsp1/nsp2 cleavage site and even in-frame deletions of nsp1 and nsp2 have only a limited effect on the MHV replication in the cell culture [18].

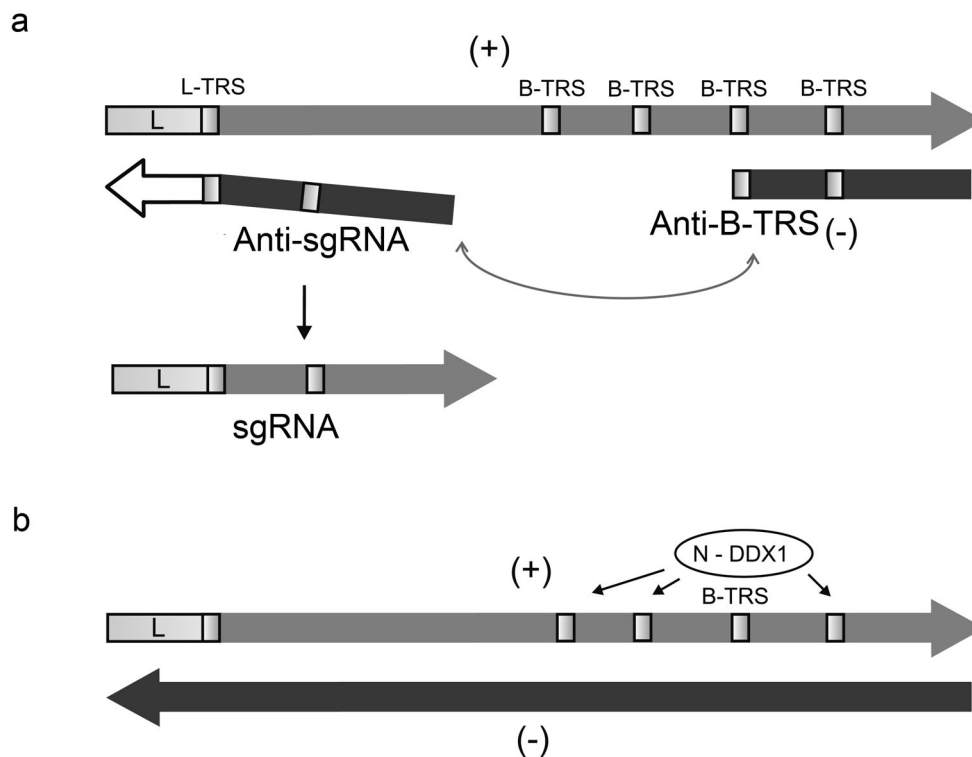


Fig. 2. Transcription and replication of coronavirus RNA. a) Discontinuous transcription. (–)RNA synthesis on the genomic (+)RNA may stop on the transcription regulatory signal (B-TRS, light box), with the following transition of the nascent (–)strand to the 5'-leader, annealing of B-TRS to L-TRS, and copying of the leader sequence. Anti-sgRNAs serve as templates to produce sgRNAs. b) Continuous replication. Free viral nucleocapsid protein N⁰ and cell helicase DDX1 attach to B-TRSs, allowing replicase to ignore the stop signals and to synthesize the full-length (–)RNA, which serves as a template for producing progeny (+)RNAs. The RNA 3'-ends are shown by arrowheads. Coding and noncoding sequences are drawn not to scale.

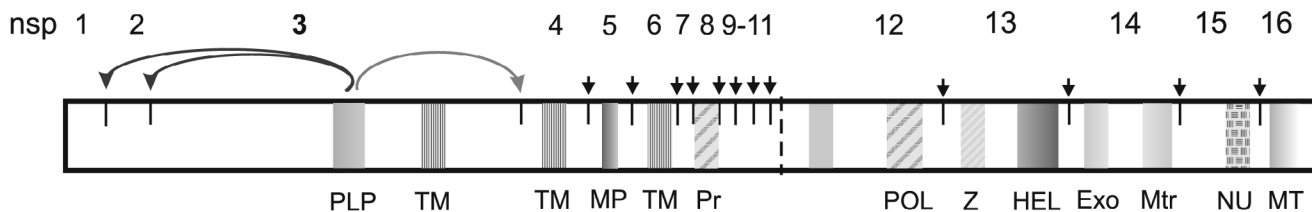


Fig. 3. Structure of the SARS-CoV-1 replicative pp1ab. Vertical dotted line indicates the boundary between 1a and 1b in the pp1ab protein. The domains for the nonstructural proteins nsp1-16 are shown as filled boxes. Designations: PLP, papain-like protease (*in cis* cleavage sites are indicated by curved arrows); MP, main chymotrypsin-like protease (*in trans* cleavage sites are shown by arrowheads); TM, transmembrane domain; Pr, putative primase; POL, RNA polymerase; Z, zinc-binding domain; HEL, RNA helicase; Exo, 3'-5' exonuclease; Mtr, N⁷-guanine methyltransferase; NU, nidoviral uridylylate-specific endoribonuclease; MT, 2'-O-ribose methyltransferase.

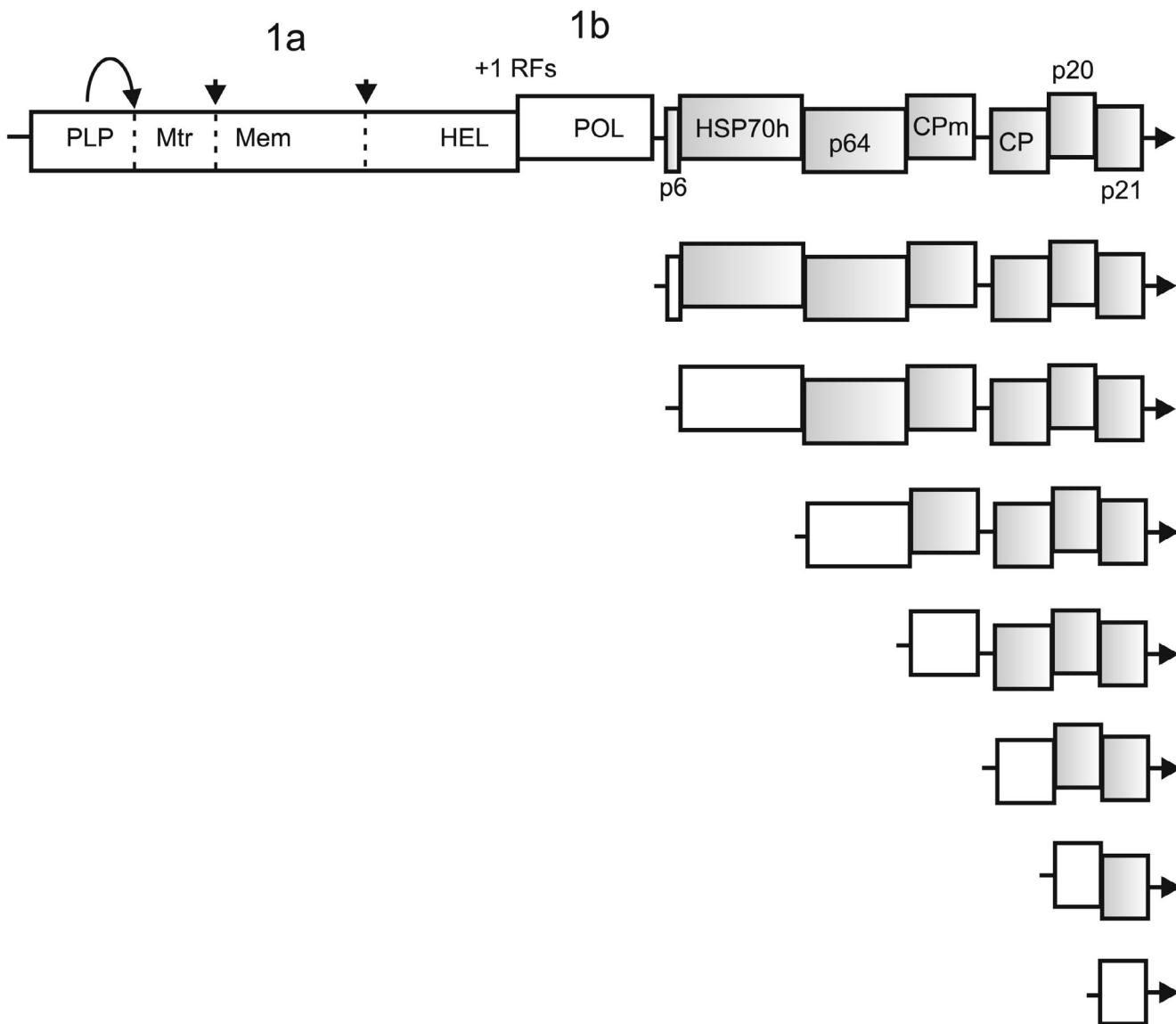


Fig. 4. Genomic RNA and sgRNAs of beet yellows closterovirus (BYV). Genes unavailable for translation in each type of mRNA are shown as shaded boxes; +1 RFs, +1 ribosomal frameshifting signal for the translation of ORFs 1a and 1b; arrows and dotted lines indicate cleavage sites in pp1a; Mem, membrane-binding domain; p6, small hydrophobic protein; HSP70h, homolog of HSP70 family heat-shock proteins; p64, 64-kDa protein; CPm, minor capsid protein; CP, major capsid protein.

The multifunctional nsp3 contains PLP and transmembrane (TM) domains [19]. The TM proteins nsp3, nsp4, and nsp6 induce reorganization of cell membranes and formation of replication compartments [20–22]. The nsp5 protease (MP) performs the majority of cleavages in the 1a and 1ab polyproteins (Fig. 4). Mature nsp5 is a part of the replicative complex. Small proteins nsp7–11 are involved in the RNA synthesis; nsp7 and nsp8 form a cylindrical heterooctameric complex, in which positively charged amino acid residues are exposed to the central lumen. A unique feature of coronaviruses, first demonstrated for SARS-CoV-1, is the presence of accessory RNA polymerase nsp8 (hypothetical primase, Pr; Fig. 3) [23]; nsp8 uses the consensus RNA sequence 5'-(G/U)CC as a template and synthesizes short complementary RNAs (up to 6 nt). In addition to poor processivity, nsp8 has the lowest fidelity among all known RNA polymerases (one misincorporation per 10 nt). It was suggested that nsp8 may act as a primase by synthesizing short RNA primers, or as a cofactor that increases the processivity of the RNA polymerase complex [23, 24].

The core RNA-synthesizing enzyme of coronaviruses nsp12 contains nucleotidyltransferase and RNA polymerase domains [24]; nsp13 contains the zinc-binding and helicase domains (Fig. 4).

The nsp14 protein includes the N-terminal 3'-5' exonuclease domain (Exo) and the C-terminal N⁷-guanine methyltransferase (Mtr) domain (Fig. 3). The activity of these domains, first predicted by computer methods, was later confirmed experimentally [25–27]. Cell exonucleases of the DEDD family, which are related to the coronaviral Exo, participate in the proofreading, repair, and recombination of nucleic acids. The nsp14 protein acts as a proofreading enzyme in the replication of coronaviral RNA and increases the fidelity of RNA copying; this function is unique for large nidoviruses [25, 26].

The nsp15 contains uridylylate-specific endoribonuclease (NU) domain (Fig. 3) [28], which is necessary for the replication and transcription and plays a role of interferon antagonist. The nsp16 contains 2'-O-ribose methyltransferase (MT) domain that methylates residue adjacent to the 5'-cap (m⁷GpppAm) in viral mRNA [29]. 2'-O-methylation allows the cell to recognize foreign mRNAs with the help of interferon-induced IFIT proteins; it is possible that 2'-O-methylation helps coronaviruses to evade this restriction mechanism [30].

Cytology of replication. After entering the cell, the coronavirus nucleocapsid is transported to the endosomes, where genomic (+)RNA is released from the N protein. Free genomic (+)RNA enters the cytoplasm and is directed to the ribosomes [31] for the synthesis of pp1a and pp1b, which are then cleaved by viral proteases to yield replication-associated proteins. Hydrophobic proteins nsp3, nsp4, and nsp6 bind to the Golgi and ER membranes and, apparently, with the aid of cell protein

partners, generate replication compartments, such as double-membrane vesicles (DMVs) and multivesicular complexes (MVCs) [20, 21]. Coronavirus nonstructural proteins and RNA bound to the DMV and MVC membranes form the “viral factories” that serve as sites for the synthesis of coronaviral antigenomic RNAs, progeny genomic RNAs, and sgRNAs [22]. Genomic and subgenomic RNAs migrate from the viral factories to the cytoplasm, where they are used as templates for translation and replication. The replication compartments induced by MHV and SARS CoV-2 have channels formed by nsp3 that open to the cytoplasm and serve for the import of substrate ribonucleoside triphosphates (rNTPs) and export of viral mRNAs [32].

CLOSTEROVIRUSES

The *Closteroviridae* family includes about 40 plant (+)RNA viruses [33, 34] characterized by several traits, such as (i) semi-persistent mode of insect transmission (virus inhabits its vector for a few hours); (ii) unique structure of filamentous particles with the spiral symmetry built of several proteins; (iii) large RNA genomes (up to 20 kb); (iv) the presence of a gene for a homolog of HSP70 chaperones; (v) the presence of duplicated genes or gene fragments (e.g., genes for the major and minor coat proteins).

Genome structure and expression. The RNA genome of beet yellows virus (BYV) was the first closteroviral genome sequenced. It consists of 14,480 nt and contains the 5'-cap, but lacks the 3'-poly(A) tail [7, 33, 35]. The 5'-terminal portion of BYV genome and genomes of other closteroviruses contain overlapping ORFs 1a and 1b coding for the replication-associated proteins (Fig. 4) [7, 33, 34]. Translation of these genes involves +1 ribosomal frameshifting and results in pp1a and pp1ab. It should be noted that the +1 frameshifting mechanism is rare in the virus world, unlike the –1 frameshifting required for the expression of RNA polymerases of animal retroviruses, nidoviruses, astroviruses, and plant sobemo-like viruses [7, 36], as well as of some eukaryotic genes and transposons [37]. The BYV genes for structural and accessory proteins are located in the 3'-terminal portion of the genome and are expressed via a set of 3'-coterminal sgRNAs [34] (Fig. 4).

The BYV PLP autocatalytically releases the leader protein by cleaving the Gly588/Gly589 bond [7]. Genomes of some other closteroviruses code for duplicated PLP domains, so that two leader proteins are released after the cleavage [8]. The leader protein influences amplification of the BYV RNA, as well as affect the long-distance transmission of the viral infection through the plant conductive tissues [38, 39]. In addition to the cleavage by PLP, pp1a undergoes processing by a yet unknown proteolytic enzyme with the formation of replication-

associated proteins of 63 kDa (Mtr) and 100 kDa (HEL) [40].

The major capsid protein (CP) coats ~95% genomic RNA, forming the “body” of the filamentous particle, while the minor capsid protein (CPm) forms a ‘tail’ that includes the 5'-terminal genome portion [41, 42] (Fig. 4). Formation of the BYV particles requires HSP70h and p64. Both proteins (one or several copies) are associated with the mature particles [43-45]. Closteroviral HSP70h has a conserved N-terminal ATPase domain (homologous to the equivalent domains in cell HSP70s) and a variable C-terminal domain [46]. The N-terminal domain of the BYV HSP70h displays the Mg-dependent ATPase activity *in vitro*, but, unlike its cellular orthologs, is unable to interact with unfolded proteins [47]. The BYV HSP70h also interacts with plasmodesmata of infected plant cells and plays a role in the cell-to-cell transmission of the viral infection [48, 49]. Virus transport also depends on p6, p64, CP, and CPm. Knocking out each of the respective genes blocks the cell-to-cell spreading of the BYV infection [50]. The products of the BYV 3'-terminal genes are involved in the long-distance transport of the virus (p20) and suppression of the post-transcriptional gene silencing (p21) [51] (Fig. 4).

Replication-associated proteins. Closteroviruses belong to the supergroup of alpha-like viruses that includes (+)RNA viruses of animals (alphaviruses, rubella virus, and hepatitis E virus) and plants (tobacco mosaic virus, brome mosaic virus, and others) (Fig. 1). Despite striking dissimilarities in the biological traits, virion morphology, and genome structure of alpha-like viruses, replicases of these viruses contain conserved Mtr, HEL and POL domains [52, 53]. The Mtr domain has the N⁷-guanine methyltransferase and guanylyltransferase activities and catalyzes the capping of viral RNA. The HEL domain unwinds RNA strands in replication, and the POL domain is responsible for the *de novo* synthesis of complementary RNA strands in a primer-independent fashion [54].

Comparisons of viral (+)RNA reveals a simple rule: the larger the size of genomic RNA, the larger the replicase gene. In other words, replication of large genomes requires more complex RNA replicative complexes [33]. Replicases of alpha-like virus have likely evolved due to the insertion of coding sequences between the fragments encoding the Mtr and HEL domains, whereas the length of the spacer between the HEL and POL domains remained almost unchanged [33]. The replicative complex of closteroviruses is more sophisticated, compared to those of closely related plant viruses, and includes at least five virus-specific proteins (PLP, Mtr, central 1a domain, HEL, and HEL-POL fusion) [34]. Closteroviral pp1a contains no domains equivalent to the nidoviral exonuclease and primase, thus leaving open the question as to whether closteroviruses possess enzymatic activities enhancing the processivity and fidelity of RNA synthesis

[51]. Replication of closteroviral genomes possibly follows the mechanism described for other alpha-like viruses: RNA replicase recognizes the 3'-terminal *cis*-signal on the (+)RNA and produces the antigenomic (–)RNA to be further used as a template for the synthesis of progeny genomic and subgenomic RNA strands. No experimental evidence has yet been reported in favor of either of two options of closterovirus transcription, namely, transcription from the subgenomic promoters on the antigenomic (–)RNA template or transcription on the antisubgenomic RNA templates [54]. Obviously, closteroviruses do not employ discontinuous transcription, since their sgRNAs do not have a common 5'-terminal leader sequence, and the subgenomic promoter regions contain no common elements resembling the TRSs of nidoviruses [55-58].

Cytology of replication. Closteroviral infection is accompanied by the induction of DMVs (~100 nm in diameter) and MVCs in the cells [59]. These ultrastructures resemble the replication factories of nidoviruses and flaviviruses [60]. The DMVs and MVCs of BYV are produced from the ER membranes. The BYV replicative proteins – PLP, Mtr, and HEL – are associated with the DMVs and MVCs, thus indicating the involvement of these structures in the RNA replication [61, 62]. In search of BYV proteins capable of membrane modification, fragments of the BYV 1a protein fused with the reporter GFP were transiently expressed in *Nicotiana benthamiana* plants [60, 63]. A 198-aa fragment (Mem; Fig. 4) containing the conserved hydrophobic domain with a predicted alpha-helix caused the remodeling of the perinuclear ER membranes and formation of ~2- μ m globules. Some globules were mobile and were associated with the actin filaments [63]. It was proposed that remodeling of the ER membranes by the hydrophobic Mem segment of the BYV 1a protein may be one of the steps in the induction of the closteroviral replication-associated ultrastructures in the cells [60, 63].

EVOLUTION OF LARGE RNA GENOMES

In the course of evolution, animal nidoviruses and plant closteroviruses have exceeded the 12-kb limit of the (+)RNA genome size. Several evolutionary “inventions” have allowed these viruses to solve the problems of packaging and replication of large RNAs. As suggested by Godeny et al. [64], the ancestor of extant Nidovirales most likely had an icosahedral nucleocapsid. In the course of subsequent divergence, “small” nidoviruses (13-16 kb RNA) have retained this core structure, whereas “large” nidoviruses (26-41 kb RNA) have acquired the N protein capable of forming helically symmetrical nucleocapsids that allow encapsidation of significantly larger RNAs. The alpha-like plant viruses closely related to *Closteroviridae* have icosahedral (bromoviruses) or rod-like virions (tobamoviruses). The size of RNA that can be

packaged into these particles is strictly limited, and it is possible that the evolution of the superflexible closteroviral particles built of several proteins has made it possible to resolve the problem of packaging of larger RNA (up to 19 kb [5]).

Mutations introduced by viral RNA polymerases [2] and, possibly, by cellular editing enzymes (e.g., deaminases) [65] are a key factor in the genetic variability of RNA viruses. Another driving force of virus evolution is RNA recombination [66]. Analysis of coronaviral and closteroviral genomes has revealed the obvious traces of recombination events, such as the capture of heterologous sequences and gene duplication. Thus, evolution of “large” nidoviruses has resulted in the acquisition of new enzymatic activities (nucleotidyltransferase, primase, 3'-5'-exonuclease, and endonuclease) [9], as well as of genes for the structural proteins (N protein and outer membrane protein related to the influenza virus hemagglutinin) [67]. Closteroviruses acquired the HSP70 gene (apparently, via recombination with a host mRNA) that has been adapted to perform specific functions in the viral cell-to-cell transmission and particle maturation [33, 51]. In some coronaviruses and closteroviruses, the sequence coding for the leader PLP proteinase has been duplicated. The capsid protein gene in closteroviral genomes has likely been duplicated several times, producing the extant genes for CP, CPm (some family members encode two minor CPs), and p64 (Fig. 4) [34, 51].

The copying of large RNA strands is mediated by more sophisticated replicases, and expansion of the RNA genome is due, for a large part, to the acquisitions in the replication-associated genes [33]. Coronaviruses possess enzymatic activities ensuring an improved fidelity and processivity of RNA replicase (primase and exonuclease). Closteroviral genomes do not encode equivalent enzymes, which might be manifested as a higher rate of point mutations in the citrus tristeza closterovirus (CTV) [68] compared to SARS-CoV and other coronaviruses [65]. Among the Nidovirales, coronaviruses and roniviruses (26-41 kb genome) code for the exonuclease, whereas mesoniviruses and arteriviruses (14-20 kb genome) lack the corresponding domain. Apparently, ~20 kb is a size limit of viral RNA, which requires no additional proofreading activity for its copying.

Comparison of the structure and expression of coronaviral and closteroviral genomes reveals a striking similarity between the two groups (Figs. 1 and 4). The genomes of these viruses contain overlapping 5'-terminal replicase genes expressed with via the ribosomal frameshifting to produce large polyproteins. The 1a and 1ab polyproteins are processed by one or two PLPs to release the leader proteins, as well as by the chymotrypsin-like MP (nidoviruses) or protease(s) of unknown origin (closteroviruses), yielding mature proteins with the methyltransferase, helicase, RNA polymerase, and membrane-binding domains. The genes for

the accessory and structural proteins of both coronaviruses and closteroviruses are translated using a set of 3'-coterminal sgRNAs (Figs. 1 and 4). These similarities do not imply that the two groups of viruses have originated from a common ancestor with similar gene set and expression strategies, as coronaviruses and closteroviruses belong to evolutionary remote lineages [1, 10] (Fig. 1). It is more probable that similar traits of genome organization and expression in coronaviruses and closteroviruses have emerged independently in the course of convergent evolution, along with the RNA genome expansion and acquisition of similar replication-associated functions and gene expression patterns [5, 7, 34, 51].

Acknowledgments. The author is grateful to A. G. Solovyev, S. Y. Morozov, and N. A. Gorgolyuk for discussions and critical reading of the manuscript.

Ethics declarations. The author declares no conflict of interest. This article does not contain description of studies with the involvement of humans or animal subjects.

Open access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

1. Koonin, E. V., and Dolja, V. V. (1993) Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences, *Crit. Rev. Biochem. Mol. Biol.*, **28**, 375-430, doi: 10.3109/10409239309078440.
2. Steinhauer, D. A., and Holland, J. (1987) Rapid evolution of RNA viruses, *Annu. Rev. Microbiol.*, **41**, 409-431, doi: 10.1146/annurev.mi.41.100187.002205.
3. Drake, J. W. (1993) Rates of spontaneous mutation among RNA viruses, *Proc. Natl. Acad. Sci. USA*, **90**, 4171-4175.
4. Holmes, E. C. (2003) Error thresholds and the constraints to RNA virus evolution, *Trends Microbiol.*, **11**, 543-546.
5. Dolja, V. V., Karasev, A. V., and Koonin, E. V. (1994) Molecular biology and evolution of closteroviruses: sophisticated build-up of large RNA genomes, *Annu. Rev. Phytopathol.*, **32**, 261-285.
6. Saberi, A., Gulyaeva, A. A., Brubacher, J. L., Newmark, P. A., and Gorbalenya, A. E. (2018) A planarian nidovirus expands the limits of RNA genome size, *PLoS Pathog.*, **14**, e1007314, doi: 10.1371/journal.ppat.1007314.
7. Agranovsky, A. A., Koonin, E. V., Boyko, V. P., Maiss, E., Froetschl, R., Lunina, N. A., and Atabekov, J. G. (1994) Beet yellows closterovirus: complete genome structure and identification of a leader papain-like thiol protease, *Virology*, **198**, 311-324, doi: 10.1006/viro.1994.1034.
8. Karasev, A. V., Boyko, V. P., Gowda, S., Nikolaeva, O. V., Hilf, M. E., et al. (1995) Complete sequence of the citrus

- tristeza virus RNA genome, *Virology*, **208**, 511-520, doi: 10.1006/viro.1995.1182.
9. Enjuanes, L., Gorbalenya, A. E., de Groot, R. J., Cowley, J. A., Ziebuhr, J., and Snijder, E. J. (2008) Nidovirales, in *Encyclopedia of Virology* (Mahy, B. W. J., and Van Regenmortel, M. H. V., eds) Oxford, Elsevier, pp. 419-430.
 10. Koonin, E. V., Dolja, V. V., Krupovic, M., Varsani, A., Wolf, Y. I., et al. (2020) Global organization and proposed megataxonomy of the virus world, *Microbiol. Mol. Biol. Rev.*, **84**, e00061-19, doi: 10.1128/MMBR.00061-19.
 11. Sawicki, S. G., and Sawicki, D. L. (2005) Coronavirus transcription: a perspective, *Curr. Top. Microbiol. Immunol.*, **287**, 31-55, doi: 10.1007/3-540-26765-4_2.
 12. Sola, I., Almazán, F., Zúñiga, S., and Enjuanes, L. (2015) Continuous and discontinuous RNA synthesis in coronaviruses, *Annu. Rev. Virol.*, **2**, 265-288, doi: 10.1146/annurev-virology-100114-055218.
 13. Zúñiga, S., Cruz, J. L., Sola, I., Mateos-Gomez, P. A., Palacio, L., and Enjuanes, L. (2010) Coronavirus nucleocapsid protein facilitates template switching and is required for efficient transcription, *J. Virol.*, **84**, 2169-2175.
 14. Wu, C. H., Chen, P. J., and Yeh, S. H. (2014) Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to continuous transcription, *Cell Host Microbe*, **16**, 462-472.
 15. Terada, Y., Kawachi, K., Matsuura, Y., and Kamitani, W. (2017) MERS coronavirus nsp1 participates in an efficient propagation through a specific interaction with viral RNA, *Virology*, **511**, 95-105, doi: 10.1016/j.virol.2017.08.026.
 16. Lokugamage, K. G., Narayanan, K., Huang, C., and Makino, S. (2012) Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation, *J. Virol.*, **86**, 13598-13608, doi: 10.1128/JVI.01958-12.
 17. Nakagawa, K., Narayanan, K., Wada, M., Popov, V. L., Cajimat, M., Baric, R. S., and Makino, S. (2018) The endonucleolytic RNA cleavage function of nsp1 of middle east respiratory syndrome coronavirus promotes the production of infectious virus particles in specific human cell lines, *J. Virol.*, **92**, e01157-18, doi: 10.1128/JVI.01157-18.
 18. Graham, R. L., Sims, A. C., Brockway, S. M., Baric, R. S., and Denison, M. R. (2005) The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication, *J. Virol.*, **79**, 13399-13411, doi: 10.1128/JVI.79.21.13399-13411.2005.
 19. Sawicki, S. G., Sawicki, D. L., Younker, D., Meyer, Y., Thiel, V., Stokes, H., and Siddell, S. G. (2005) Functional and genetic analysis of coronavirus replicase-transcriptase proteins, *PLoS Pathog.*, **1**, e39, doi: 10.1371/journal.ppat.0010039.
 20. Angelini, M. M., Akhlaghpour, M., Neuman, B. W., and Buchmeier, M. J. (2013) Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles, *MBio*, **4**, e00524-e00513.
 21. de Wilde, A. H., Snijder, E. J., Kikkert, M., and van Hemert, M. J. (2018) Host factors in coronavirus replication, *Curr. Top. Microbiol. Immunol.*, **419**, 1-42, doi: 10.1007/82_2017_25.
 22. Knoops, K., Kikkert, M., Worm, S. H., Zevenhoven-Dobbe, J. C., van der Meer, Y., et al. (2008) SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum, *PLoS Biol.*, **6**, e226, doi: 10.1371/journal.pbio.0060226.
 23. Te Velthuis, A. J., van den Worm, S. H., and Snijder, E. J. (2012) The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension, *Nucleic Acids Res.*, **40**, 1737-1747, doi: 10.1093/nar/gkr893.
 24. Kirchdoerfer, R. N., and Ward, A. B. (2019) Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors, *Nat. Commun.*, **10**, 2342, doi: 10.1038/s41467-019-10280-3.
 25. Bouvet, M., Imbert, I., Subissi, L., Gluais, L., Canard, B., and Decroly, E. (2012) RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus non-structural protein nsp10/nsp14 exoribonuclease complex, *Proc. Natl. Acad. Sci. USA*, **109**, 9372-9377, doi: 10.1073/pnas.1201130109.
 26. Subissi, L., Posthuma, C. C., Collet, A., Zevenhoven-Dobbe, J. C., Gorbalenya, A. E., et al. (2014) One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities, *Proc. Natl. Acad. Sci. USA*, **111**, E3900-9, doi: 10.1073/pnas.1323705111.
 27. Jin, X., Chen, Y., Sun, Y., Zeng, C., Wang, Y., et al. (2013) Characterization of the guanine-N7 methyltransferase activity of coronavirus nsp14 on nucleotide GTP, *Virus Res.*, **176**, 45-52, doi: 10.1016/j.virusres.2013.05.001.
 28. Ivanov, K. A., Hertzog, T., Rozanov, M., Bayer, S., Thiel, V., Gorbalenya, A. E., and Ziebuhr, J. (2004) Major genetic marker of nidoviruses encodes a replicative endoribonuclease, *Proc. Natl. Acad. Sci. USA*, **101**, 12694-12699, doi: 10.1073/pnas.0403127101.
 29. Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., and Guo, D. (2009) Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase, *Proc. Natl. Acad. Sci. USA*, **106**, 3484-3489.
 30. Daffis, S., Szretter, K. J., Schriewer, J., Li, J., Youn, S., et al. (2010) 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members, *Nature*, **468**, 452-456, doi: 10.1038/nature09489.
 31. Kalicharran, K., Mohandas, D., Wilson, G., and Dales, S. (1996) Regulation of the initiation of coronavirus JHM infection in primary oligodendrocytes and L-2 fibroblasts, *Virology*, **225**, 33-43, doi: 10.1006/viro.1996.0572.
 32. Wolff, G., Ronald, W., Limpens, R., Zevenhoven-Dobbe, J. C., Laugks, U., et al. (2020) A molecular pore spans the double membrane of the coronavirus replication organelle, *Science*, **369**, 1395-1398, doi: 10.1126/science.abd3629.
 33. Agranovsky, A. A. (1996) Principles of molecular organization, expression and evolution of closteroviruses: over the barriers, *Adv. Virus Res.*, **47**, 119-158.
 34. Agranovsky, A. A. (2016) Closteroviruses: molecular biology, evolution and interactions with cells, in *Plant Viruses: Evolution and Management* (Gaur, R. K., Petrova, N., and Stoyanova, M. I., eds) Springer Science+Business Media, Singapore, Chapt. 14, pp. 231-252, doi: 10.1007/978-981-10-1406-2_14.
 35. Agranovsky, A. A., Boyko, V. P., Karasev, A. V., Lunina, N. A., Koonin, E. V., and Dolja, V. V. (1991) Nucleotide sequence of the 3'-terminal half of beet yellows closterovirus RNA genome: unique arrangement of eight virus genes, *J. Gen. Virol.*, **72**, 15-23.
 36. Firth, A. E., and Brierley, I. (2012) Non-canonical translation in RNA viruses, *J. Gen. Virol.*, **93**, 1385-1409, doi: 10.1099/vir.0.042499-0.

37. Giedroc, D. P., and Cornish, P. V. (2009) Frameshifting RNA pseudoknots: structure and mechanism, *Virus Res.*, **139**, 193-208.
38. Peremyslov, V. V., Hagiwara, Y., and Dolja, V. V. (1998) Genes required for replication of the 15.5-kilobase RNA genome of a plant closterovirus, *J. Virol.*, **72**, 5870-5876.
39. Peng, C. W., Napuli, A. J., and Dolja, V. V. (2003) Leader proteinases of beet yellows virus functions in long-distance transport, *J. Virol.*, **77**, 2843-2849.
40. Erokhina, T. N., Zinovkin, R. A., Vitushkina, M. V., Jelkmann, W., and Agranovsky, A. A. (2000) Detection of beet yellows closterovirus methyltransferase-like and helicase-like proteins *in vivo* using monoclonal antibodies, *J. Gen. Virol.*, **81**, 597-603.
41. Agranovsky, A. A., Lesemann, D. E., Maiss, E., Hull, R., and Atabekov, J. G. (1995) "Rattlesnake" structure of a filamentous plant RNA virus built of two capsid proteins, *Proc. Natl. Acad. Sci. USA*, **92**, 2470-2473.
42. Zinovkin, R. A., Jelkmann, W., and Agranovsky, A. A. (1999) The minor coat protein of beet yellows closterovirus encapsidates the 5'-terminus of RNA in virions, *J. Gen. Virol.*, **80**, 269-272.
43. Napuli, A. J., Alzhanova, D. V., Doneanu, C. E., Barofsky, D. F., Koonin, E. V., and Dolja, V. V. (2003) The 64-kDa capsid protein homolog of beet yellows virus is required for assembly of virion tails, *J. Virol.*, **77**, 2377-2384.
44. Napuli, A. J., Falk, B. W., and Dolja, V. V. (2000) Interaction between HSP70 homolog and filamentous virions of the beet yellows virus, *Virology*, **274**, 232-239.
45. Alzhanova, D. V., Napuli, A., Creamer, R., and Dolja, V. V. (2001) Cell-to-cell movement and assembly of a plant closterovirus: roles for the capsid proteins and Hsp70 homolog, *EMBO J.*, **20**, 6997-7007.
46. Agranovsky, A. A., Boyko, V. P., Karasev, A. V., Koonin, E. V., and Dolja, V. V. (1991) The putative 65K protein of beet yellows closterovirus is a homologue of HSP70 heat shock proteins, *J. Mol. Biol.*, **217**, 603-610.
47. Agranovsky, A. A., Folimonova, S. Y., Folimonov, A. S., Denisenko, O. N., and Zinovkin, R. A. (1997) The beet yellows closterovirus p65 homologue of HSP70 chaperones has ATPase activity associated with its conserved N-terminal domain but does not interact with unfolded protein chains, *J. Gen. Virol.*, **78**, 535-542.
48. Medina, V., Peremyslov, V. V., Hagiwara, Y., and Dolja, V. V. (1999) Subcellular localization of the HSP70-homolog encoded by beet yellows closterovirus, *Virology*, **260**, 173-181.
49. Agranovsky, A. A., Folimonov, A. S., Folimonova, S. Y., Morozov, S. Y., Schiemann, J., Lesemann, D. E., and Atabekov, J. G. (1998) Beet yellows closterovirus HSP70-like protein mediates the cell-to-cell movement of a potyvirus transport-deficient mutant and a hordeivirus-based chimeric virus, *J. Gen. Virol.*, **79**, 889-895.
50. Alzhanova, D. V., Hagiwara, Y., Peremyslov, V. V., and Dolja, V. V. (2000) Genetic analysis of the cell-to-cell movement of beet yellows closterovirus, *Virology*, **268**, 192-200.
51. Dolja, V. V., Kreuze, J. F., and Valkonen, J. P. (2006) Comparative and functional genomics of closteroviruses, *Virus Res.*, **117**, 38-51.
52. Goldbach, R., Le Gall, O., and Wellink, J. (1991) Alpha-like viruses of plants, *Semin. Virol.*, **2**, 19-25.
53. Buck, K. W. (1996) Comparison of the replication of positive-stranded RNA viruses of plants and animals, *Adv. Virus Res.*, **47**, 159-251, doi: 10.1016/s0065-3527(08)60736-8.
54. Miller, W. A., and Koev, G. (2000) Synthesis of subgenomic RNAs by positive strand RNA viruses, *Virology*, **273**, 1-8.
55. Agranovsky, A. A., Koenig, R., Maiss, E., Boyko, V. P., Casper, R., and Atabekov, J. G. (1994) Expression of the beet yellows closterovirus capsid protein and p24, a capsid protein homologue, *in vitro* and *in vivo*, *J. Gen. Virol.*, **75**, 1431-1439.
56. Karasev, A. V., Hilf, M. E., Garnsey, S. M., and Dawson, W. O. (1997) Transcriptional strategy of closteroviruses: mapping the 5' termini of the citrus tristeza virus subgenomic RNAs, *J. Virol.*, **71**, 6233-6236.
57. Peremyslov, V. V., and Dolja, V. V. (2002) Identification of the subgenomic mRNAs that encode 6-kDa movement protein and Hsp70 homolog of beet yellows virus, *Virology*, **295**, 299-306.
58. Vitushkina, M. V., Rogozin, I. B., Jelkmann, W., Koonin, E. V., and Agranovsky, A. A. (2007) Completion of the mapping of transcription start sites for the five-gene block subgenomic RNAs of beet yellows closterovirus and identification of putative subgenomic promoters, *Virus Res.*, **128**, 153-158.
59. Cronshaw, J., Hoefert, L., and Esau, K. (1966) Ultrastructural features of beta leaves infected with beet yellows virus, *J. Cell. Biol.*, **31**, 429-443.
60. Gushchin, V. A., Solovyev, A. G., Erokhina, T. N., Morozov, S. Y., and Agranovsky, A. A. (2013) Beet yellows virus replicase and replicative compartments: parallels with other RNA viruses, *Front. Microbiol.*, **4**, 38, doi: 10.3389/fmicb.2013.00038.
61. Erokhina, T. N., Vitushkina, M. V., Zinovkin, R. A., Lesemann, D. E., Jelkmann, W., Koonin, E. V., and Agranovsky, A. A. (2001) Ultrastructural localisation and epitope mapping of beet yellows closterovirus methyltransferase-like and helicase-like proteins, *J. Gen. Virol.*, **82**, 1983-1994.
62. Zinovkin, R. A., Erokhina, T. N., Lesemann, D. E., Jelkmann, W., and Agranovsky, A. A. (2003) Processing and subcellular localization of the leader papain-like proteinase of beet yellows closterovirus, *J. Gen. Virol.*, **84**, 2265-2270.
63. Gushchin, V. A., Karlin, D. G., Makhotenko, A. V., Khromov, A. V., Erokhina, T. N., et al. (2017) A conserved region in the Closterovirus 1a polyprotein drives extensive remodeling of endoplasmic reticulum membranes and induces motile globules in *Nicotiana benthamiana* cells, *Virology*, **506**, 106-113, doi: 10.1016/j.virol.2016.12.006.
64. Godeny, E. K., Chen, L., Kumar, S., Methven, S. L., Koonin, E. V., and Brinton, M. A. (1993) Complete genome sequence and phylogenetic analysis of the lactate dehydrogenase-elevating virus (LDV), *Virology*, **194**, 585-96.
65. Simmonds, P. (2020) Rampant C/U hypermutation in the genomes of SARS-CoV-2 and other coronaviruses: causes and consequences for their short- and long-term evolutionary trajectories, *mSphere*, **5**, e00408-20, doi: 10.1128/mSphere.00408-20.
66. Bentley, K., and Evans, D. J. (2018) Mechanisms and consequences of positive-strand RNA virus recombination, *J. Gen. Virol.*, **99**, 1345-1356, doi: 10.1099/jgv.0.001142.
67. Zeng, Q., Langereis, M. A., van Vliet, A. L., Huizinga, E. G., and de Groot, R. J. (2008) Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution, *Proc. Natl. Acad. Sci. USA*, **105**, 9065-9069, doi: 10.1073/pnas.0800502105.
68. Harper, S. J. (2013) Citrus tristeza virus: evolution of complex and varied genotypic groups, *Front. Microbiol.*, **4**, 9310, doi: 3389/fmicb.2013.00093.