CTMI (2005) 287:1–30

Coronavirus Genome Structure and Replication

D. A. Brian[1] (✉) · R. S. Baric[2, 3]

[1] Departments of Microbiology and Pathobiology, University of Tennessee, College of Veterinary Medicine, Knoxville, TN 37996-0845, USA
dbrian@utk.edu
[2] Department of Microbiology and Immunology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NV 27599-7400, USA
[3] Department of Epidemiology, Program of Infectious Diseases, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7400, USA

Abstract In addition to the SARS coronavirus (treated separately elsewhere in this volume), the complete genome sequences of six species in the coronavirus genus of the coronavirus family [avian infectious bronchitis virus-Beaudette strain (IBV-Beaudette), bovine coronavirus-ENT strain (BCoV-ENT), human coronavirus-229E strain (HCoV-229E), murine hepatitis virus-A59 strain (MHV-A59), porcine transmissible gastroenteritis-Purdue 115 strain (TGEV-Purdue 115), and porcine epidemic diarrhea virus-CV777 strain (PEDV-CV777)] have now been reported. Their lengths range from 27,317 nt for HCoV-229E to 31,357 nt for the murine hepatitis virus-A59, establishing the coronavirus genome as the largest known among RNA

viruses. The basic organization of the coronavirus genome is shared with other members of the Nidovirus order (the torovirus genus, also in the family *Coronaviridae*, and members of the family *Arteriviridae*) in that the nonstructural proteins involved in proteolytic processing, genome replication, and subgenomic mRNA synthesis (transcription) (an estimated 14–16 end products for coronaviruses) are encoded within the 5′-proximal two-thirds of the genome on gene 1 and the (mostly) structural proteins are encoded within the 3′-proximal one-third of the genome (8–9 genes for coronaviruses). Genes for the major structural proteins in all coronaviruses occur in the 5′ to 3′ order as S, E, M, and N. The precise strategy used by coronaviruses for genome replication is not yet known, but many features have been established. This chapter focuses on some of the known features and presents some current questions regarding genome replication strategy, the *cis*-acting elements necessary for genome replication [as inferred from defective interfering (DI) RNA molecules], the minimum sequence requirements for autonomous replication of an RNA replicon, and the importance of gene order in genome replication.

1 Introduction

Despite its unique property as the largest of the known plus-strand RNA genomes, the coronavirus genome shares with those of other plus-strand RNA viruses (excepting retroviruses) the properties of (1) infectiousness [and not using a packaged RNA-dependent RNA polymerase (RdRp)] (Brian et al. 1980; Schochetman et al. 1977) and (2) replication in the cytoplasm in close association with cellular membranes (Denison et al. 1999; Dennis and Brian 1982; Gosert et al. 2002; Sethna and Brian 1997; Shi et al. 1999; van der Meer et al. 1999). Many of the basic features of coronavirus genome structure and replication have been described in recent reviews (Cavanagh et al. 1997; Enjuanes et al. 2000a, 2000b; Lai and Cavanagh 1997; Lai and Holmes 2001; Luytjes 1995; van der Most and Spaan 1995). With the advent of reverse genetics enabling site-directed mutagenesis of any part of the genome (Almazan et al. 2000; Casais et al. 2001; Masters 1999; Thiel et al. 2001; Yount et al. 2000, 2002), many of the mechanistic features of coronavirus genome replication that could previously be learned only from direct manipulation of defective interfering (DI) RNA can now be examined in the context of the whole virus genome. In this chapter, we review the current knowledge of coronavirus genome structure and organization and the *cis*-acting elements in coronavirus replication and raise selected questions that we believe are important for approaching a better understanding of coronavirus genome replication.

2
Common Features of Genome Structure Among Coronaviruses

In addition to the SARS coronavirus (treated separately elsewhere in this volume), the genomes of six species of coronaviruses have now been fully sequenced and reported in GenBank (as of November 2002): IBV-Beaudette (NC 001451, Boursnell et al. 1987), BCoV-ENT (NC 003045, Chouljenko et al. 2001), MHV-A59 (NC 001846, Leparc-Goffart et al. 1997), HCoV-229E (NC 002645, Herold et al. 1993; Thiel et al. 2001), TGEV-Purdue (NC 002306, Almazan et al. 2000; Eleouet et al. 1995; Penzes et al. 2001), and PEDV-CV777 2001 (NC 003436, Kocherhans et al. 2001). These, representing all three coronavirus serogroups (Siddell 1995), are schematically depicted in Fig. 1. Additional strains of BCoV [BCoV-LUN (AF391542, Chouljenko et al. 2001)], BCoV-Mebus (U00735, Nixon and Brian, unpublished data) and BCoV-Quebec (AF220295, Yoo and Pei 2001), and MHV [MHV-2 (AF201929, Sarma et al. 1999)] have also been reported. The genome sizes range from 27,317 nt for HCoV-229E to 31,357 nt for MHV-A59, establishing them as the largest known among RNA viruses (Enjuanes et al. 2000a; Lai and Cavanagh 1997). The following similarities in genome structure among the six can be noted:

1. The 5′ UTRs ranging in length from 209 to 528 nt contain a similarly positioned short, AUG-initiated open reading frame (ORF) relative to the 5′ end [Table 1; a situation that, by current terminology, is problematic because the "untranslated region" now becomes in part potentially translatable and thus should preferably be called a "leader" (Morris and Geballe 2000). The term "leader," however, has an established meaning in the nidovirus lexicon (Lai and Cavanagh 1997; see subsequent chapters, this volume) of a 5'-terminal, genome-encoded sequence of 65–98 nt appearing on the 5′ terminus of each subgenomic mRNA species]. For purposes of this review, "5′ UTR" will refer to the sequence upstream of ORF 1 (gene 1) despite the internally positioned short ORF. The short AUG-initiated ORFs (except for HCoV-229E) begin in a suboptimal Kozak context for translation (Table 1) (Kozak 1991) and potentially encode peptides of 3–11 amino acids.
2. The 3′ UTRs range from 288 to 506 nt [although some strains of IBV have 3′ UTRs of greater length because of internal sequence duplications (Williams et al. 1993)], all possess an octameric sequence of GGAAGAGC beginning at base 73 to 80 upstream from the poly(A) tail, and all possess a 3′-terminal poly(A) tail (Table 1).

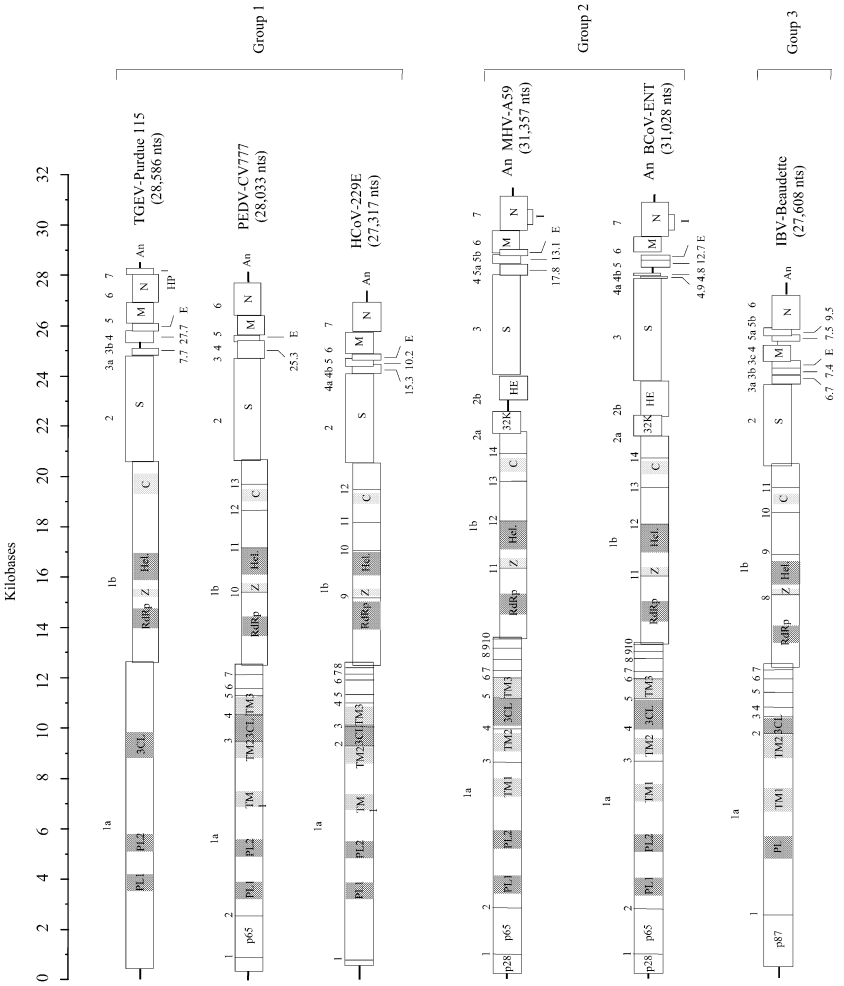


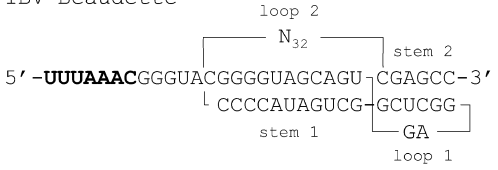
Fig. 1. Genomes of the six sequenced species of coronaviruses known prior to the discovery of the SARS coronavirus. Maps are drawn to approximate scale, and species are shown in decreasing order of size within each of the three groups. The representations are derived from data in the GenBank as of November 2002. For gene 1 (ORFs 1a and 1b) the predicted protease cleavage sites are indicated by numbers and domains of known or predicted function are shaded and identified (*PL*, papain-like protease; *3CL*, poliovirus 3C-like protease; *TM*, transmembrane domain; *RdRp*, RNA-dependent RNA polymerase; *Z*, zinc finger (metal-binding) domain; *Hel*, helicase domain; *C*, conserved sequence domain). Genes 2–8 (or 9) are identified by their transcript name (1a, 1b, etc.) or their abbreviated name of the protein product (*S*, spike; *E*, envelope; *M*, membrane; *N*, nucleocapsid; *HP*, hydrophobic protein; *HE*, hemagglutinin-esterase; *I*, internal). Literature references are described with the GenBank information (see text)

Table 1. Properties of the coronavirus 5′ UTR, intra-5′ UTR short ORF, and 3′ UTR

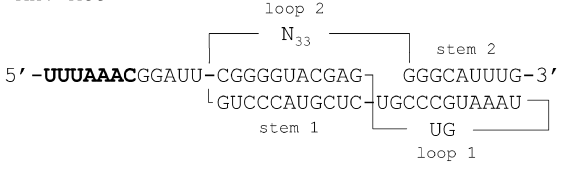
Coronavirus	Length of 5′ UTR (number of nt upstream of gene 1)	Position and Kozak context[a] of the intra-5′ UTR short ORF start codon	Number of amino acids encoded by the 5′ UTR short ORF	Amino acid sequence of the 5′ UTR short ORF product	Length of 3′ UTR (number of nt)	Position of the first nt in the octamer GGAAGAGC upstream from the 3′ poly(A) tail
TGEV-Purdue	314	117UCUaugA	3	MKS	279	76
PEDV-CV777	296	105GUUaugC	10	MLLEAGVEFH	334	73
HCoV-229E	292	86GCUaugG	11	MAGIFDAGVVV[b]	462	74
MHV-A59	209	99UCCaugC	8	MPAGLVLS	324	81
BCoV-ENT	210	100UCUaugC	8	MPVGVDFS[c]	288	78
IBV-Beaudette	528	131UGGaugG	11	MAPGHLSGFCY	506	80

[a] The optimal Kozak context for translation initiation is GCCaugG (Kozak 1991).
[b] A second ORF beginning 16 nt downstream from the first and in the plus 1 reading frame relative to the first encodes the amino acids MLES.
[c] The second amino acid in the BCoV-Mebus strain is L.

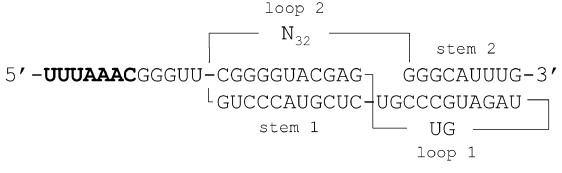
IBV-Beaudette



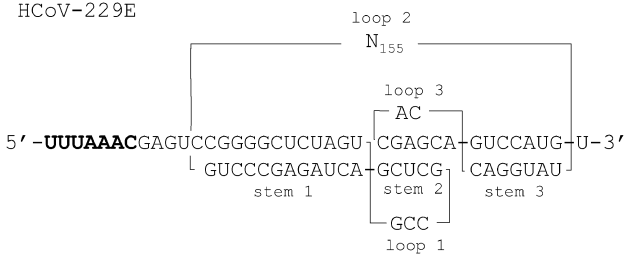
MHV-A59



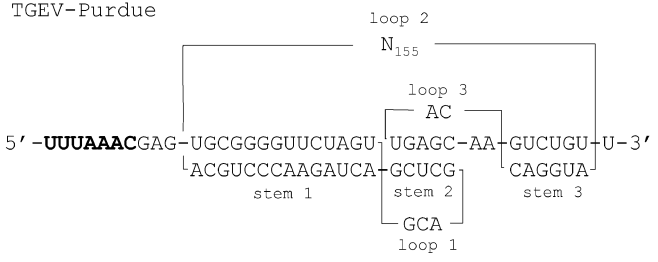
BCoV-ENT



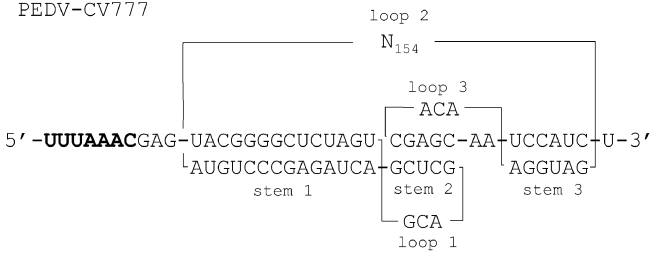
HCoV-229E



TGEV-Purdue



PEDV-CV777



3. All have an extremely large gene 1 (separated into ORFs 1a and 1b and extending over approximately two-thirds of the genome) encoding nonstructural proteins involved in proteolytic processing of the gene 1 polyprotein products, virus genome replication, and sgmRNA synthesis (transcription). In each, gene 1 is translated as ORFs 1a and 1ab, with 1ab resulting from a pseudoknot-induced −1 ribosomal frame shifting event at a slippery sequence of UUUAAAC at the ORF 1a/1b junction (Fig. 2) (Brown and Brierley 1995).
4. All encode the structural spike (S) glycoprotein, small envelope (E) protein, membrane (M) glycoprotein, and nucleocapsid (N) protein, in that order, 5′→3′ within the 3′-proximal one-third of the genome. A variable number of other ORFs appearing to be virus- or group-specific, many apparently encoding nonstructural proteins, are also found here. These (and their potential products) include ORF 3a (7.7-kDa protein), ORF 3b (27.7-kDa protein), and ORF 7 [0.7-kDa hydrophobic protein (HP)] in TGEV; ORF 3 (25.3-kDa protein) in PEDV; ORF 4a (15.3-kDa protein) and ORF 4b (10.2-kDa protein) in HCoV-229E; ORF 2a (32-kDa protein), ORF 2b [65-kDa complete or 34.6-kDa truncated hemagglutinin-esterase (HE) protein, depending on the strain], ORF 4 (17.8-kDa protein), ORF 5a (13.1-kDa protein), and an ORF internal to gene 7 [23-kDa internal (I) protein] in MHV; ORF 2a (32-kDa protein), ORF 2b (65-kDa HE protein), ORF 4a (4.9-kDa protein), ORF 4b (4.8-kDa protein), ORF 5 (12.7-kDa protein), and an ORF internal to gene 7 (23-kDa I protein) in BCoV; and ORF 3a (6.7-kDa protein), ORF 3b (7.4-kDa protein), ORF 5a (7.5-kDa protein), and ORF 5b (9.5-kDa protein) in IBV (Fig. 1; Brown and Brierly 1995, and references listed in the GenBank information noted above). Some of these, such as ORFs 3a and 3b in TGEV (McGoldrick 1999; Wesley et al. 1991) and ORFs 2a

Fig. 2. Pseudoknotted structures and slippery sequences responsible for highly efficient (25%–30%) –1 ribosomal frameshifting at the ORF 1a and 1b junction in gene 1 of the six coronaviruses shown in Fig. 1. The slippery sequence UUUAAAC, identified in bold, is the same in all sequenced genomes. The IBV pseudoknot-induced frameshifting was the first nonretroviral example of ribosomal frameshifting in higher eukaryotes (Brierley et al. 1987, 1989). The pseudoknots in MHV (Bredenbeek et al. 1990) and BCoV (Yoo and Pei 2001) are nearly identical and are similar to the structure in IBV. In HCoV-229E an elaborated pseudoknot with three stems was shown by mutation analysis to be the functional frameshifting structure (Harold and Siddell 1993). In TGEV (Eleouet et al. 1995) and in PEDV (Kocherhans et al. 2001) an elaborated pseudoknot was also predicted based on similarities to HCoV-229E

(Schwarz et al. 1990), 2b (HE) (Luytjes et al. 1988), 4 (Weiss et al. 1993; Yokomori and Lai 1991), 5a (Yokomori and Lai 1991), and I (Fischer et al. 1997) in MHV, have been shown to be nonessential for replication in cell culture, and their function in virus replication remains undetermined (de Haan et al. 2002).

Presumably all coronavirus genomes are capped with a 5′ methylated nucleotide, but so far this has been demonstrated only in MHV (Lai et al. 1982).

3
Cis-Acting RNA Elements in Coronavirus Genome Replication

3.1
The 5′ UTR and the Translation Step(s) Preceding Genome Replication

As with all nonretroviral plus-strand RNA viruses, a necessary early step in genome replication is translation of the genome for production of the RdRp and other proteins required for viral genome replication. The presence of a 5′ terminal methylated cap on MHV genomic and subgenomic mRNAs (Lai et al. 1982) would suggest that coronaviruses use a cap-mediated ribosomal entry mechanism for translation. Mutation analyses of the 5′ UTR of BCoV indicate that a scanning mechanism is used for entry of ribosomes onto ORF 1 (Senanayake and Brian 1999). Curiously in light of these results, a methylated cap on DI RNA transcripts is not required for initiation of replication of BCoV DI RNA, which contains a genomic 5′ UTR. This molecule has a *cis*-acting dependence on translation for replication (Chang et al. 1994; Chang and Brian 1996). It remains to be determined whether capping is required for translation and replication of the intact viral genome. It remains to be determined what enzyme functions to cap the viral RNAs (Ziebuhr et al. 2000).

In MHV it has been demonstrated that the viral nucleocapsid protein N binds tightly (K_d=14 nM) to the UCUAAAC intergenic region (also named transcription-regulating sequence, TRS) of the genomic leader and consequently may influence translation rate (Nelson et al. 2000; Tahara et al. 1998). Is this property of N common to all coronaviruses? If so, what role does it play in the regulation of genome replication?

Does the intra-5′ UTR short ORF play a role in translation (or in subsequent replication) of the genome? With reverse genetics, disruption of

an analogous ORF in equine arterivirus had no apparent effect on virus replication in cell culture (Molenkamp et al. 2000), but the ORFs may not have homologous function in the two virus groups. Certainly, short upstream ORFs can have profound enhancing or suppressing effects on the translation of a downstream ORF (Morris and Geballe 2000), and their universal existence in coronavirus 5′ UTRs, albeit with little or no conservation in size or amino acid sequence (Table 1), would suggest that they function in the regulation of replication or gene expression. One possibility is that the intra-5′ UTR short ORF or some other 5′ UTR element, such as the binding site for N described above, is responsible for the repression of translation from the ORF 1 start codon in virus-infected cells (Senanayake and Brian 1999).

Some observed phenomena in coronavirus genome and DI RNA replication hint that the 5′ UTR might be bypassed altogether in order to meet the translation requirements for genome replication. One set of observations relates to a possible role for N in genome replication (Baric et al. 1988; Compton et al. 1987; Kim K and Makino 1995; Laude and Masters 1995; Nelson et al. 2000; Stohlman et al. 1988), a role that would set coronaviruses apart from arteriviruses in this regard because only gene 1 products have been shown to be sufficient for arterivirus genome replication (Molenkamp et al. 2000). N protein, for example, binds leader sequence with high affinity (Nelson et al. 2000), is present in a subpopulation of coronavirus RNA replication complexes (Sethna and Brian 1997; Sims et al. 2000), and is essential for infectivity of recombinant IBV full-length transcripts (Casais et al. 2001). If N is required, then some mechanism for the translation of N from the polycistronic genome, such as an internal entry of ribosomes onto genomic RNA or formation of an early subgenomic mRNA transcript, would be needed, at least when infection is initiated by the genome alone (as in transfection experiments). Some evidence for internal ribosomal entry has been demonstrated for IBV mRNA 3 (Liu and Inglis 1992), MHV mRNA 5 (Thiel and Siddell 1994; Jengrach et al. 1999), and TGEV mRNA 3 (O'Connor and Brian 2000), making it prudent to consider an internal entry at these or other sites on the genome for protein synthesis. Another set of observations relates to a requirement for translation in *cis* of the DI RNA molecule to be replicated. Although some DI RNAs with a single ORF do not appear to require translation *in cis* for replication (Liao and Lai 1995), others do (Chang and Brian 1996; De Groot et al. 1992; Van der Most et al. 1995). Might a *cis*-acting requirement for DI RNA translation reflect a similar *cis*-translation-dependent mechanism for genome replication as described for picornaviruses (Egger et al.

2000; Gamarnik and Andino 1998; Novak and Kirkegaard 1994) and flaviviruses (Khromykh et al. 1999)? If so, then perhaps an internal ribosomal entry for translation onto the 3′ proximal region of the genome might be needed for coronavirus genome replication.

3.2 The Pseudoknot and Slippery Sequence Involved in the – 1 Ribosomal Frameshifting at the ORF 1a/1b Junction

Ribosomal frameshifting in coronaviruses was the first described non-retroviral example of ribosomal frameshifting in higher eukaryotes (Brierly 1987), and the earliest described higher-order RNA structure recognized as a *cis*-acting element in coronavirus genome replication was the pseudoknot located immediately downstream of the UUUAAAC slippery sequence in the IBV genome (Brierly et al. 1987, 1989; Brown and Brierly 1995) (Fig. 2). The pseudoknot in IBV was described as a hairpin-type and was shown by mutation analyses to be responsible for the highly efficient (25%–30%) frameshifting. Subsequently, a pseudoknot with similar properties was found in gene 1 of MHV (Bredenbeek et al. 1990) and BCoV (Yoo and Pei 2001). Interestingly, the pseudoknot found in gene 1 of HCoV-229E was found to be quite different in structure, possessing an extremely large loop 2 and a stem 3 (Fig. 2). This structure was termed an "elaborated" pseudoknot and was shown to function as such in in vitro measurements of frameshifting (Herold and Siddell 1993). The predicted pseudoknots in TGEV and PEDV gene 1 appear to be quite similar to that in HCoV-229E (Eleouet et al. 1995; Kocherhans et al. 2001). The pseudoknot-associated slippery sequence is UUUAAAC in all sequenced coronaviruses described to date.

3.3 *Cis*-Acting Elements Required for Membrane Association of the RNA with the Replication Complex

Once made, or possibly concurrent with synthesis, viral proteins and (possibly) associated cellular proteins function to form the membrane-associated RNA replication complexes. Membrane association is a hallmark of replication complexes of plus-strand RNA viruses, but the origin of the membrane and the anatomy of the replication complexes appear to differ among virus families. A preliminary understanding of the coronavirus replication complex has come primarily from studies with MHV and partly from studies with TGEV. The following features have been observed:

1. The membrane in the MHV replication complex has shown markers for the endoplasmic reticulum and Golgi (Shi et al. 1999; Gosert at al. 2002) and, alternatively, the late endosomes (van der Meer et al. 1999; Sims et al. 2000).
2. The replication complex is intimately associated with double membrane structures, and the anchored proteins are the hydrophobic sequence-containing intermediate cleavage products p290 and p150, and p210 and p44, of ORF 1a (Gossert et al. 2002).
3. There appear to be two populations of membrane-associated replication complexes separable by isopycnic sedimentation (Sethna and Brian 1997; Sims et al. 2000). In MHV the less dense fraction (1.05–1.09 g/ml) was found to contain p65 and p1a-22, products of ORF 1a, whereas the denser fraction (1.12–1.25 g/ml) contained p28 and helicase from ORF 1b, and N (Sims et al. 2000).

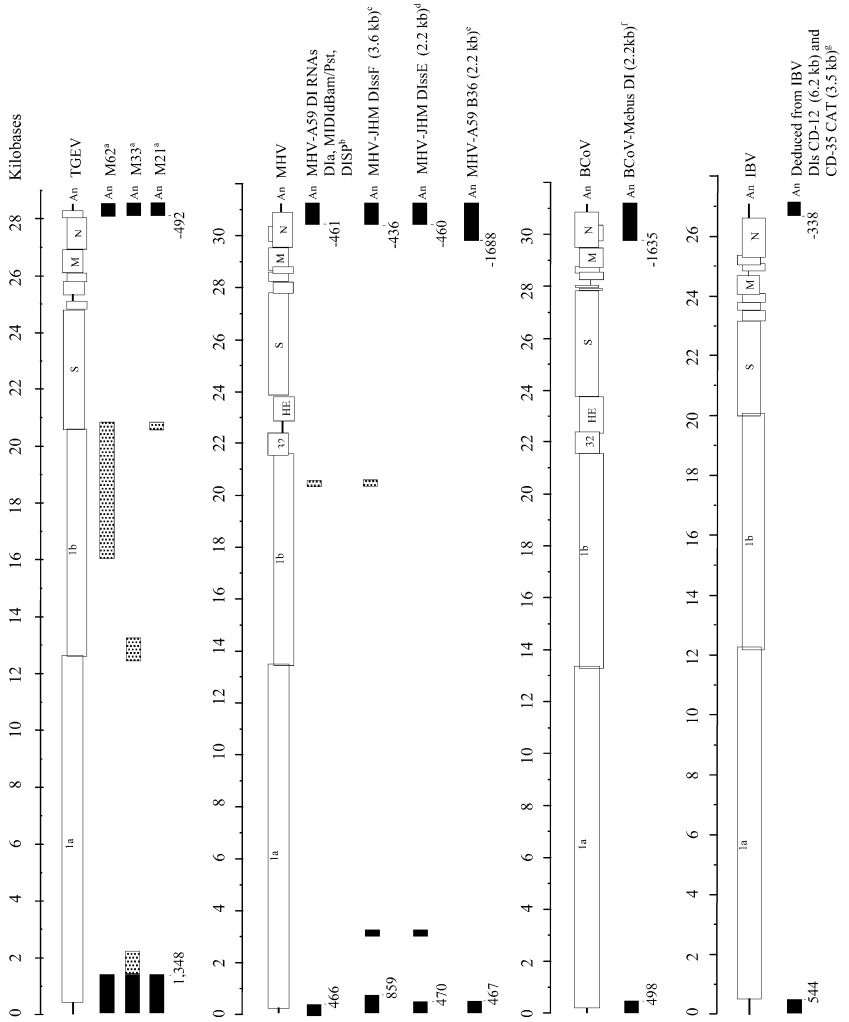
In TGEV two buoyant density populations (1.15–1.17 g/ml and 1.20–1.24 g/ml) were also found, and both had associated with them genome- and subgenome-length plus- and minus-strand RNAs (Sethna and Brian 1997). Some S, M, and N proteins were associated with the denser population. The TGEV membrane replication complexes, furthermore, appeared to have an unusual impermeability to micrococcal nuclease. It remains to be determined precisely what proteins, viral and cellular, function together to make up the coronavirus replication complexes and how they might be associated with the membranes and with one another. How might they differ between the processes of minus- and plus-strand synthesis? Between replication and transcription? Which proteins bind the RNA, both genomic and subgenomic, both plus and minus strands, within the complex? What is the stoichiometry of the components in the various complexes? What is the relationship between the RNA replication complex and the site of virus assembly at the Golgi and intermediate Golgi membranes? How is the genome selected and transported from the replication complex to the site of virus assembly? Does the evidence of resistance of coronaviral RNAs to ribonuclease suggest existence of a compartmentalized replication complex and have implications for resistance to RNA silencing (Ahlquist 2002) and long-term persistent coronaviral infections (Adami et al. 1995; Baric et al. 1999; Okumura et al. 1996; Stohlman et al. 1999)?

3.4
5′ and 3′-Proximal RNA *Cis*-Acting Elements for DI RNA (and Presumably Genome) Replication

Since the first description of their cloning and replication in helper virus-infected cells, coronavirus DI RNAs have been used in attempts to define the minimal *cis*-acting sequence requirements for their replication (Brian and Spaan 1997; Makino et al. 1985, 1988a, 1988b; van der Most et al. 1991). Through deletion analyses the regions harboring minimal *cis*-acting sequences have been mapped for DI RNAs from TGEV, MHV, BCoV, and IBV (noted as filled regions in the DI RNA maps in Fig. 3). For most of the DI RNAs it can be seen that these sequences reside at the termini of the viral genomes for distances of 467–1,348 at the 5′ end and 338–1,635 at the 3′ end. Further reduction in the sizes of these regions may result from further deletion analyses. Requirements for internal genome sequence elements appear to be DI RNA specific but may reflect requirements of the intact genome (see below). What is the nature of the terminal *cis*-acting RNA elements? Is a specific sequence alone sufficient, or are higher-order structures required? So far, these questions have focused primarily on the small (2.2–2.3 kb) DI RNAs of the group 2 coronaviruses MHV and BCoV.

With regard to the 3′ UTR of MHV-A59 and BCoV-Mebus, common replication signals exist between the two viruses. This was demonstrated by experiments in which the entire 3′ UTR of the MHV genome was replaced with the equivalent region of the BCoV genome without loss of virus viability (Hsue and Masters 1997) and in a BCoV DI RNA chimera in which the BCoV 3′ UTR was replaced with the MHV 3′ UTR with no detectable loss of replicating ability (Ku, Williams, and Brian, unpublished data). More recently, BCoV DI RNA has been shown to replicate in the presence of MHV as helper virus (Wu et al. 2003). To date, three higher-order *cis*-acting elements mapping within the 3′ UTR have been characterized in MHV and BCoV (Fig 4):

Fig. 3. Map positions of minimal *cis*-acting sequences for RNA replication (*solid boxes*) and signals for packaging (*stippled boxes*) as determined from studies on DI RNAs and their derivatives. The schematic diagrams of the four coronaviruses studied in this manner are shown. (a) Izeta et al. 1999; deletion analyses were done on derivatives of TGEV DI RNA C (9.7 kb) (Mendez et al. 1996); M21 contains minimal sequence elements for replication and inefficient packaging; M33 and M62 contain



small nonoverlapping regions of ORFs 1a and 1b that contribute to packaging; (b) Luytjes et al. 1996; van der Most et al. 1991, 1995; deletion analyses were done on derivatives of MHV-A59 DIa RNA (5.5 kb); (c) Lin and Lai 1993; Makino et al. 1990; deletion analyses were done on DIssF; (d) Fosmire et al. 1992; Kim et al. 1993; Kim and Makino 1995; deletion analyses were done on DIssE; (e) Masters et al. 1994; DI B36 is synthetic and was designed after the BCoV-Mebus DI RNA; (f) Chang et al. 1994; deletion analyses were done on reporter-containing DI Drep1; (g) Dalton et al. 2001; deletion analyses were done on derivatives of 9.1-kb IBV DI RNA CD-91 (Penzes et al. 1994); unknown regions within the UTRs suffice for packaging of DI RNA, but packaging is inefficient

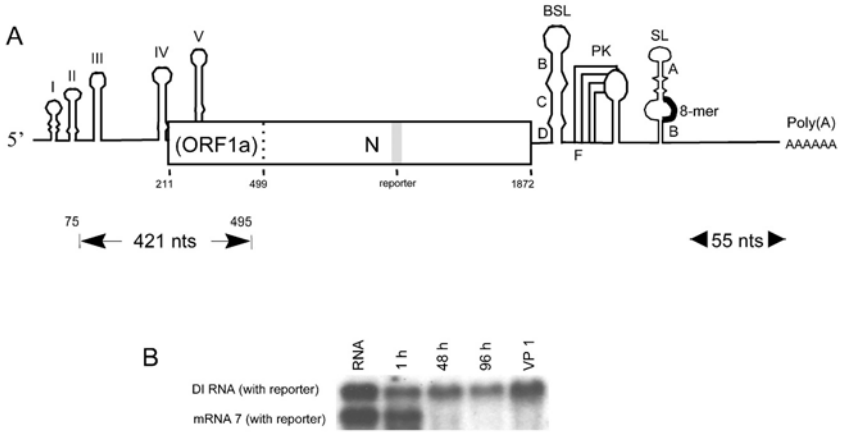


Fig. 4A, B. Terminal *cis*-acting replication sequences and higher-order structures identified to date in the smallest of the MHV and BCoV DI (group 2) RNAs. **A** The DI RNA illustrated is that for BCoV, but the structures drawn (with the exception of the 5′-proximal stem-loops I and II and the upper portion of the 3′-proximal octamer-associated stem loop) are phylogenetically conserved between MHV and BCoV. The *open rectangle* represents an open reading frame formed by the fusion of the first part of ORF1a and the entire N gene. The 3′ higher-order structures are a 61-nt bulged stem-loop (Hsue et al. 2000), a hairpin-type pseudoknot (Williams et al. 1999), a helix formed at the base of a long stem-loop and adjacent to the phylogenetically conserved octameric sequence (Liu et al. 2001). The poly(A) tail is required for replication (Lin and Lai 1993;, Spagnolo and Hogue 2000), and the 5′-terminal 55 nt are the minimal sequence requirements for minus-strand RNA synthesis in MHV (Lin et al. 1994). The 5′ higher-order structures are a stem-loop III and stem-loop IV within the 5′ UTR (Raman et al. 2002) and stem V within the partial ORF 1a sequence (Brown et al. 2002). **B** Experimental evidence for replication (accumulation) of reporter-containing DI RNA but not mRNA7 containing the same reporter after transfection into helper virus-infected cells (Chang et al. 1994). The only difference between the two molecules is a sequence of 421 nt mapping between nt 74 and 497 in the BCoV DI RNA

1. A 68-nt bulged stem-loop beginning immediately downstream of the N stop codon consists in MHV of four stems (B, C, D, and F) and a 14-nt terminal loop (Hsue and Masters 1997; Hsue et al. 2000). Stems C, D, and F have been shown to be required for replication of both the DI RNA and virus genome.
2. A 54-nt hairpin-type pseudoknot beginning 60 nt downstream of the bulged stem-loop (Williams et al. 1999). Both stems of the pseudoknot have been shown to be required for replication. The pseudoknot sequence overlaps the downstream arm of stem F in the bulged stem-loop

such that the two structures cannot exist simultaneously. This led Hsue et al. (2000) to suggest a possible interaction between the two elements, with the alternative conformations acting as a possible "switching" mechanism. This switch has now been confirmed experimentally (Goebel et al. 2004).The pseudoknot appears phylogenetically conserved to some degree in all coronaviruses.

3. A 74-nt bulged stem-loop mapping from nt 68 to 142 from the 3′ terminus in MHV contains two stems that demonstrated importance as *cis*-acting replication structures (referred to as stems A and B in Fig. 4) (Liu et al. 2001). Stem B, which shows greater importance in DI RNA replication, is phylogenetically conserved in structure between MHV and BCoV. Stem B is immediately adjacent downstream to the phylogenetically conserved 3′ UTR octamer GGAAGAGC (Liu et al. 2001). Unidentified cellular proteins of 120, 55, 40, and 25 kDa molecular mass bind to nt 130–142 which is the upstream half of the internal loop in stem B (Liu et al. 1997; Yu and Liebowitz 1995).

Proteins identified to date that bind within the 3′ region (or the minus-strand counterpart of this region) include the poly(A) binding protein (Spagnolo and Hogue 2000), mitochondrial aconitase, which binds within the 42-nt 3′-terminal region in MHV (Nanda and Leibowitz 2001), and the polypyrimidine tract-binding protein, which binds to minus-strand sequence complementary to nt 53–149 (strongly) and 270–307 (weakly) in MHV (Huang and Lai 1999). What roles the 3′ UTR higher-order structures play in RNA replication are not known. Because the 3′-terminal 55 nt were shown to be a minimal sequence requirement for minus-strand synthesis in MHV (Lin et al. 1994), the higher-order structures mapping upstream of the 55-nt sequence possibly play no role in minus-strand synthesis. Do they play a role in initiating or regulating plus-strand synthesis? Precedents in picornaviruses (Barton et al. 2001; Herold and Andino 2001), alphaviruses (Frolov et al. 2001),and flaviruses (You et al. 2001) would suggest they might. Certainly the poly(A) tail through the poly(A)-binding protein is a candidate for such a process, perhaps through genome circularization (Spagnolo and Hogue 2000).

With regard to the 5′ UTR it is known that the 5′-terminal sequence is required for DI RNA replication (Chang et al. 1994; Kim et al. 1993) and at least two stem-loops (stem-loops III and IV in Fig. 4) function as higher-order *cis*-acting signaling elements (Raman et al. 2003; Raman and Brian, unpublished data). A higher-order *cis*-acting structure mapping within the first 290 nt of ORF1 (stem-loop V in Fig. 4) has also been

found (Brown, Nixon, Senanayake, and Brian, unpublished data). Proteins shown to bind within the 5′ UTR include the viral N protein, which binds in and around the leader-adjacent intergenic sequence motif UCUAAAC (Nelson et al. 2000), the polypyrimidine tract binding protein, which also binds near the leader-adjacent UCUAAAC sequence motif (Li et al. 1999), and hnRNP A1, which binds the minus-strand complement of the leader-adjacent UCUAAAC sequence motif (Li et al. 1997). None of these has been reported to bind regions covered by stem-loops III, IV, or V depicted in Fig. 4. Might there be a process of leader priming of genome replication (Zhang and Lai 1996), as suggested by the phenomenon of high-frequency leader switching on DI RNAs during DI RNA replication (Chang et al. 1996; Makino and Lai 1989; Stirrups et al. 2000)?

The question of what *cis*-acting sequences act in coronavirus RNA replication has relevance not only for genome replication but also for poorly understood features of sgmRNA behavior. It has been suggested that coronavirus sgmRNAs amplify by a replication mechanism (Brian et al. 1994; Hofmann et al. 1990; Sethna et al. 1989). This hypothesis made use of the argument that the termini on the sgmRNAs and genome, identical at the 5′ end for the length of the leader (65–98 nt, depending on the virus species) and at the 3′ end for greater than the length of the 3′ UTR (i.e., greater than 300 nt), are larger than the known promoters for a viral RdRp [replication promoters in influenza and Sindbis viruses are less than 20 nt in length (Levis et al. 1986; Li and Palese 1992)] and are therefore large enough to harbor promoters for replication. The hypothesis was also consistent with the observations that (1) the molar ratios of minus-strand to plus-strand RNA are equivalent for sgmRNA and genome (i.e., 1:100), (2) the rate of sgmRNA accumulation is inversely proportional to the length of the molecule, (3) the rate of sgmRNA minus strand disappearance parallels that of antigenome, and (4) sgmRNA minus strands possess 3′-terminal sequences complementary to the leader (Sethna et al. 1989). Furthermore, (5) double-stranded subgenomic mRNA-length RFs and RIs (Hofmann and Brian 1991; Hofmann et al. 1990; Sawicki and Sawicki 1990; Sethna et al. 1989) were shown to be active in subgenomic mRNA synthesis (Baric and Yont 2000; Sawicki and Sawicki 1995, 1998; Sawicki et al. 2001; Schaad and Baric 1994). If the 3′-terminal 55 nt are the only requirement for minus-strand RNA synthesis (Lin et al. 1994), the possibility is left open that the subgenomic mRNAs function as a templates for minus-strand synthesis. At no time, however, has it been directly demonstrated that sgmRNA transcripts, with or without a reporter, are replicated in

the presence of a helper virus after transfection into helper virus-infected cells (Fig. 4B) (Chang et al. 1994; Makino et al. 1991). Therefore, what features enable the replication of the DI RNAs but not sgmRNAs on transfection into helper virus-infected cells? The answer could lie in the function of the 5′-proximal stem-loops III, IV, and V residing within the 421-nt region found in BCoV DI RNA but not found in sgmRNAs (Fig. 4A) (Chang et al. 1994). Do these higher-order structures bind viral or cellular proteins? Might they be signals working through long-distance RNA-RNA or RNA-protein interactions?

3.5
Internal *Cis*-Acting Signals for DI RNA (and Possibly Also for Genome) Replication

Most DI RNAs described for coronaviruses are comprised of more than just the terminal genomic sequences. That is, they are mosaics of internal and terminal genome sequences. Replication of MHV-JHM DI RNAs has been found to be dependent on a 57-nt sequence mapping within ORF 1a (Kim and Makino 1995; Lin and Lai 1993). This sequence has been shown to form a secondary structure in the positive strand, and both the higher-order structure and its sequence are important for function as a replication signal (Repass and Makino 1998). Does this structure represent a *cis*-acting replication signal required for replication of the intact genome?

4
Packaging Signals

Perpetuation of coronavirus infection via cell-to-cell spread requires that the genome be packaged into virions via one or more *cis*-acting packaging signals. Inasmuch as several small DI RNAs containing only terminal sequences are packaged, some form of signal sufficient for incorporation into virions must reside in the termini. This idea is consistent with the observed packaging of subgenomic mRNAs in TGEV (Sethna et al. 1989), BCoV (Hofmann et al. 1990), and IBV (Zhao et al. 1993). However, these packaging signals may not be the ones used by the virus genome for packaging. A candidate 69-nt genome packaging signal has been identified in mosaic DI RNAs of MHV (Fosmire et al. 1992; Makino et al. 1990; van der Most et al. 1991) that maps to a region within ORF 1b, shows correlation of function with maintenance of secondary struc-

ture (Fosmire et al. 1992), and confers packaging on reporter RNA molecules (Bos et al. 1997; Woo et al. 1997). A homologous structure in BCoV ORF 1b also leads to packaging of nonviral RNAs (Cologna and Hogue 2000). Do these represent the bona fide packaging signals for the viral genome? Is there perhaps more than one packaging signal, as suggested by the ability of more than a single region of ORF 1b to contribute to packaging efficiency in large TGEV DI RNAs (Izeta et al. 1999)? Perhaps not since a recent study shows only a single packaging signal encoded within the 5'-terminal 649 nts of the TGEU genome is sufficient (Escors et al. 2003). In addition to the N protein (Laude and Masters 1995), might the packaging signals interact with other components of the virion? Perhaps so since in MHV the envelop (E) protein (Narayanan and Makino 2001) and M protein (Narayanan et al. 2003) have been shown to play roles in packaging.

5
Minimum Sequence Requirements for (Autonomous) Genome Replication

Although gene 1 products are the only ones required for arterivirus genome replication and sgmRNA synthesis (Molenkamp et al. 2000), the story might be different for coronaviruses. Gene 1 of HCoV-229E in the presence of the genomic 5′ and 3′ UTRs was shown to be sufficient for sgmRNA synthesis when the intergenic sequence for mRNA 7 (N mRNA) and an mRNA body (gene for the green fluorescence protein) were present just downstream of gene 1 (Thiel et al. 2001). The authors, however, were unable to conclude that these sequences alone were sufficient for RNA replication or to rule out a role for N as an enhancer for transcription. These results, therefore, leave open the possibility that another gene function is important for replication. Autonomous replicons of TGEV containing only genes 1, 2, part of 5, and all of 6 and 7 have been described (Curtis et al. 2002). Reverse genetics with these and other coronaviruses now make feasible the analysis of the minimal sequences required for genome replication and should lead to a definitive resolution of the question of the role of N protein in RNA replication.

6 Importance of Gene Order for Genome Replication

The gene order for coronaviruses, as for many positive- and negative-stranded RNA virus families, is highly conserved. In coronaviruses the essential genes pol, S, E, M, and N are invariably found in that order, 5′ to 3′, although they are sometimes interspersed with genes showing no essential function for virus growth in cell culture (discussed above). What is the significance of this gene order? If it is altered, what might the consequences be on virus growth? Might pathogenesis be altered such that the variants could be used as vaccines or vectors for other uses?

The presence of nonessential genes 3a and 3b in TGEV for cell culture growth has enabled development of TGEV as a heterologous expression vector (see the chapter by Enjuanes et al., this volume) and as a virus to study the effects of gene rearrangements. In initial studies on the effect of gene rearrangement, the N gene has been duplicated (producing the genotype SNEMN) and repositioned (producing the genotype SNEM) by making use of gene positions 3A and 3B (K. Curtis and R. Baric, unpublished data). The N gene was chosen for repositioning because it encodes the most abundantly expressed sgmRNA and is translated into the most abundant of the viral proteins. On the basis of general gene expression patterns relative to the 3′ end of the genome in coronaviruses it was anticipated that expression of E and M would increase relative to N in the rearranged SNEM construct. When tested by transfection, the TGEV mutants SNEMN and SNEM were found to be viable but to replicate at about 10-fold and 1,000-fold less than wild-type virus levels, respectively. These results indicated that a specified gene order per se is not essential for coronavirus replication in cell culture, but that order contributes in some way to a more robust virus yield. When TGEV SNEM was serially passaged 15 times, the mutant gene order SNEM was maintained, but, surprisingly, virus growth was restored to near wild-type levels. Restoration of TGEV SNEM fitness as defined by virus yield was associated with changes within the N-(partial) Δ3B-E junction region. These included removal of most of the residual (partial) ΔORF3B sequence, deletion of the wt E intergenic sequence element, and activation of a new, highly transcriptionally active E intergenic sequence element just downstream of the newly inserted N gene (Fig. 5B). These results indicate that high-frequency RNA recombination does not function to restore a specific coronavirus gene order, at least over the short term, because the new N gene position in SNEM was stable for many passages. Rather, the

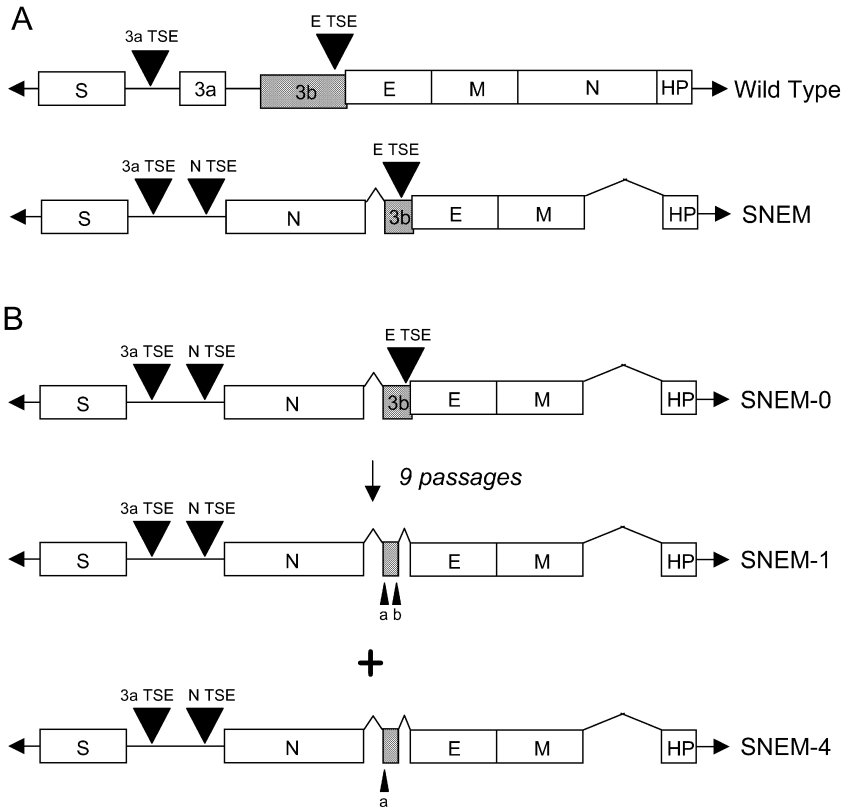


Fig. 5A, B. Effects of moving the N gene within the TGEV genome from its normal position to an upstream site. The N gene including its immediate transcription stimulating element (TSE)-containing upstream sequence of 24 nt was placed just downstream of the 3a TSE sequence in a TGEV genome from which the entire 3a and a portion of the 3b gene had been deleted (**A**). Transcripts of the recombinant TGEV genome, designated SNEM, were transfected into cells, and progeny viruses were studied (**B**). Immediately after transfection (passage 0) the titer of progeny was low ($<10^5$ PFU/ml) and the genome sequence was identical to the original construct. The progeny (SNEM-1 and SNEM-4) grew more efficiently (~5.0×10^6 PFU/ml) after 9 passages and reached wild-type levels (~1.0×10^8) after 24 passages. In all progeny the upstream 3a TSE sequence was used for leader fusion of the N transcript. For expression of the E gene, however, the story was different. At passage 0 (SNEM-0), transcripts of the E gene used the wt TSE as well as two additional sites, designated a and b within the ORF3b residual sequence, for leader fusion. In the SNEM-1 and SNEM-4 viruses the wt E TSE was deleted and transcripts of the E gene used the two new TSEs formed within the residual gene 3b sequence (a=4/5 clones, b=1/5 clones) in SNEM-1. In SNEM-4 only the a site was used for E gene expression. Thus the reordered TGEV genome was stable with regard to the new (upstream) position of N

coronavirus genome can rapidly develop compensatory changes to restore virus replication rate (fitness) while maintaining a new gene order. Mechanisms of fitness restoration appeared to include recombination events and point mutations (Baric et al., unpublished data). It is likely that gene order mutants will provide novel insights into the regulation of coronanvirus transcription and replication, identify protein-protein interactions that function cooperatively to maintain robust virus fitness and growth, and assist in the identification of core sequence elements that function in sgmRNA synthesis.

7 Future Directions

It is anticipated that reverse genetics, which now enables an alteration of any part of the coronavirus genome, will facilitate examination of the *cis*- and *trans*-acting elements in RNA replication and transcription within the context of the intact genome. These elements have until now been studied primarily in DI RNAs. In light of precedents established with many much smaller plus-strand RNA viruses of animals and plants, it would not be surprising to find novel long-distance RNA-RNA and protein-RNA interactions involving genome sequences not present in DI RNAs. Long-distance interactions are hinted at in comparative studies of DI RNAs (which replicate) and sgmRNAs (which do not replicate). What genes are important in regulation of replication and transcription, and how important is gene order in these processes? These questions can now be rigorously approached with reverse genetics. It is also anticipated that a greater understanding of the assembly, stoichiometry, and function of the RNA synthesizing complexes will be gained through similar rigorous analyses. It is anticipated that one practical outcome of reverse genetics will be the development of safe coronavirus-based replicon vectors, not necessarily only those that become packaged, for vaccine and other biomedical uses. Still in waiting is the development of an in vitro virus replication system such as that used for poliovirus (Molla

for over 24 passages, but in SNEM-1 and SNEM-4 additional mutations were selected upstream of the 3aTSE and in the M gene that greatly enhanced virus fitness and N gene expression. In SNEM the sequences of the TSEs are AACTAAACT for 3a, and ACAAAAC for E, TAACTAAACT for N, AACTAAAG for a, and AACACAAAAC for b

et al. 1991), in which complete virus replication can be accomplished in cell lysates. This approach would enable still more detailed analyses of the requirements for genome replication beginning with the infectious genome transcript. All in all, it is likely that the next decade will bring significant breakthroughs regarding our understanding of the mechanisms involved in coronavirus genome replication and transcription, the function of the replication complexes, and the development and application of coronavirus recombinant vectors for the treatment of animal and human diseases.

Acknowledgements We thank Cary Brown, Kimberley Nixon, Sharmila Raman, Gwyn Williams, and Hung-Yi Wu in the Brian laboratory and Kristopher Curtis and Boyd Yount in the Baric laboratory for invaluable discussions and experimentation. Work in D. Brian's laboratory is supported by grant AI-14367 from the National Institutes of Health and work in R. Baric's laboratory by grants AI-23946 and GM-63228 from the National Institutes of Health.

References

Adami C, Pooley J, Glomb J, Stecker E, Fazal F, Fleming JO, Baker SC (1995) Evolution of mouse hepatitis virus (MHV) during chronic infection: quasispecies nature of the persisting MHV RNA. Virology 209:337–346

Ahlquist P (2002) RNA-dependent RNA polymerases, viruses, and RNA silencing. Science 296:1270–1273

Almazan F, Gonzalez JM, Penzes Z, Izeta a, Calvo E, Plana-Duran J, Enjuanes L (2000) Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. Proc Natl Acad Sci USA 97:5516–5521

Baric RS, Nelson GW, Fleming JO, Deans RJ, Keck JG, Casteel N, Stohlman SA (1988) Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. J. Virol 62:4280–4287

Baric RS, Sullivan E, Hensley L, Yount B, Chen W (1999) Persistent infection promotes cross-species transmissibility of mouse hepatitis virus. J Virol 73:638–649

Baric RS, Yount B (2000) Subgenomic negative-strand RNAs function during mouse hepatitis virus infection. J Virol 74:4039–4046

Barton DJ, Donnell BJO, Flanegan JB (2001) 5′ Cloverleaf in poliovirus RNA is a *cis*-acting replication element required for negative-strand synthesis. EMBO J 20:1439–1448

Bos ECW, Dobbe JC, Luytjes W, Spaan WJM (1997) A subgenomic mRNA transcript of the coronavirus mouse hepatitis virus strain A59 defective interfering (DI) RNA is packaged when it contains the DI RNA packaging signal. J Virol 71:5684–5687

Boursnell MEG, Brown TDK, Foulds IJ, Green, PF, Tomley FM, Binns MM (1987) Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J Gen Virol 68:57–77

Bredenbeek PJ, Pachuk CJ, Noten AFH, Charite, J, Luytjes W, Weiss SR, Spaan WJM (1990) The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59: a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. Nucleic Acids Res 18:1825–1832

Brian DA, Chang RY, Hofmann MA, Sethna PB (1994) Role of subgenomic minus-strand RNA in coronavirus replication. Arch Virol 9 (Suppl): 173–180

Brian DA, Dennis DE, Guy JS (1980) Genome of porcine transmissible gastroenteritis virus. J Virol 34:410–415

Brian DA, Spaan WJM (1997) Recombination and coronavirus defective interfering RNAs. Semin Virol 8:101–111

Brierly I, Boursnell MEG, Binns MM, Bilimoria B, Blok VC, Brown TDK, Inglis SC (1987) An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J 6:3779–3785

Brierly I, Digard P, Inglis SC (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell 57:537–547

Brown TDK, Brierley I (1995) The coronavirus nonstructural proteins. In The Coronaviridae (S.G. Siddell, ed.), Plenum Press, New York and London, pp. 191–2171

Casais R, Thiel V, Siddell SG, Cavanagh D, Britton P (2001) Reverse genetics system for the avian coronavirus infectious bronchitis virus. J Virol 75:12359–12369

Cavanagh D, Brian DA, Brinton MA, Enjuanes L, Holmes KV, Horzinek MC, Lai MMC, Laude H, Plagemann PGW, Siddell, SG, Spaan W, Taguchi F, Talbot PJ (1997) Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch Virol. 142:629–633

Chang RY, Brian DA (1996) *Cis* requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. J Virol 70:2201–2207

Chang RY, Hofmann MA, Sethna PB, Brian BA (1994) A *cis*-acting function for the coronavirus leader in defective interfering RNA replication. J Virol 68:8223–8231

Chang RY, Krishnan R, Brian DA (1996) The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. J Virol 70:2720–2729

Chouljenko VN, Lin XQ, Storz J, Kousoulas KG, Gorbalenya AE (2001) Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia. J Gen Virol 82:2927–2933

Cologna R, Hogue BG (2000) Identification of a bovine coronavirus packaging signal. J Virol 74:580–583

Compton SR, Rogers DB, Holmes KV, Fertsch D, Remenick J, McGowan JJ (1987) In vitro replication of mouse hepatitis virus strain A59. J Virol 61:1814–1820

Curtis K, Yount B, Baric RS (2002) Heterologous gene expression from transmissible gastroenteritis virus replicon particles. J Virol 76:1422–1434

Dalton K, Casais R, Shaw K, Stirrups K, Evans S, Britton P, Brown TDK, Cavanagh, D (2001) *cis*-Acting sequences required for coronavirus infectious bronchitis virus defective-RNA replication and packaging. J Virol 75:125–133

de Groot RJ, van der Most RG, Spaan WJM (1992) The fitness of defective interfering murine coronavirus DI-a and its derivatives is decreased by nonsense and frameshift mutations. J Virol 66:5898–5905

de Haan CAM, Masters PS, Shen X, Weiss S, Rottier PJM (2002) The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. Virology 296:177–189

Denison MR, Spaan WJ, van der Meer Y, Gibson CA, Sims AC, Prentice E, Lu XT (1999) The putative helicase of the coronavirus mouse hepatitis virus is processed from the replicase gene poly protein and localizes in complexes that are active in viral RNA synthesis. J Virol 73:6862–6871

Dennis DE, Brian DA (1982) RNA-dependent RNA polymerase activity in coronavirus-infected cells. J Virol 42:153–164

de Vries AAF, Horzinek MC, Rottier PJM, de Groot RJ (1997) The genome organization of the Nidovirales: similarities and differences between arteri-, toro-, and coronaviruses. Semin Virol 8:33–47

Egger D, Teterina N, Ehrenfeld E, Bienz K (2000) Formation of the poliovirus replication complex requires coupled viral translation, vesicle production, and viral RNA synthesis. J Virol 74:6570–6580

Eleouet JF, Rasschaert D, Lambert P, Levy L, Vende P, Laude H (1995) Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. Virology 206:817–822

Enjuanes L, Brian D, Cavanagh D, Holmes K, Lai MMC, Laude H, Masters P, Rottier PJM, Siddell SG, Spaan WJM, Taguchi F, Talbot P (2000) Coronaviridae. In: Virus Taxonomy, Seventh Report of the International Committee on Taxonomy of Viruses (MHV van Regenmortel, CM Fauquet, DHL Bishop, EB Carstens, MK Estes, SM Lemon, J Maniloff, MA Mayo, DJ McGeoch, CR Pringle, RB Wickner, eds) Academic Press, San Diego. pp 835–849

Enjuanes L, Spaan WJM, Snijder E, Cavanagh D (2000) Nidovirales. In: Virus Taxonomy, Seventh Report of the International Committee on Taxonomy of Viruses (MHV van Regenmortel, CM Fauquet, DHL Bishop, EB Carstens, MK Estes, SM Lemon, J Maniloff, MA Mayo, DJ McGeoch, CR Pringle, RB Wickner, eds) Academic Press, San Diego. pp 827–834

Escors D, Izeta A, Capiscol C, Enjuanes L (2003) Transmissible gastroenteritis coronavirus packaging signal is located at the 5'end of the virus genome. J Virol 77:7890–7902

Fosmire JA, Hwang K, Makino S (1992) identification and characterization of a coronavirus packaging signal. J Virol 66:3522–3530

Frolov I, Hardy R, Rice CM (2001) *Cis*-acting RNA elements at the 5′ end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. RNA 7:1638–1651

Gamarnik AV, Andino R (1998) Switch from translation to RNA replication in a positive-stranded RNA virus. Genes Dev 12:2293–2304

Goebel SJ, Hsue B, Dombrowski TF, Masters PS (2004) Characterization of the RNA components of a putative molecular switch in the 3'untranstated region of the murine coronavirus. J Virol 78:669–682

Gosert R, Kanjanahaluethai A, Egger D, Bienz K, Baker SC (2002) RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. J. Virol 76:3697–3708

Herold J, Andino R (2001) Poliovirus RNA replication requires genome circularization through a protein-protein bridge. Mol Cell 7:581–591

Herold J, Raabe T, Schelle-Prinz B, Siddell SG (1993) Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. Virology 195:680–691

Herold, J, Siddell SG (1993) An "elaborated" pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. Nucleic Acids Res 21:5838–5842

Hofmann MA, Brian DA (1991) The 5′ end of coronavirus minus-strand RNAs contains a short poly(U) tract. J Virol 65:6331–6333

Hofmann MA, Sethna PB, Brian DA (1990) Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. J Virol 64:4108–4114

Hsue B, Hartshorne T, Masters PS (2000) Characterization of an essential RNA secondary structure in the 3′ untranslated region of murine coronavirus genome. J Virol 74:6911–6921

Hsue B, Masters PS (1997) A bulged stem-loop structure in the 3′ untranslated region of the coronavirus mouse hepatitis virus genome is essential for replication. J Virol 71:7567–7578

Huang P, Lai MMC (1999) Polypyrimidine tract-binding protein binds to the complementary strand of the mouse hepatitis virus 3′ untranslated region, thereby altering RNA conformation. J Virol 73:9110–9116

Izeta A, Smerdou C, Alonso S., Penzes Z, Mendez A, Plana-Duran J, Enjuanes, L (1999) Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. J Virol 73:1535–1545

Jengrach M, Thiel V, Siddell S (1999) Characterization of an internal ribosome entry site within mRNA 5 of murine hepatitis virus. Arch Virol 144:921–933

Khromykh AA, Sedlak PL, Westaway EG (1999) Trans-complementation analysis of the flavivirus Kunjin ns5 gene reveals an essential role for translation of its N-terminal half in RNA replication. J Virol 73:9247–9255

Kim KH, Makino S (1995) Two murine coronavirus genes suffice for viral RNA synthesis. J Virol 69:2313–2321

Kim Y, Makino S (1995) Characterization of a murine coronavirus defective interfering RNA internal *cis*-acting replication signal. J Virol 69:4963–4971

Kim YN, Jeong YS, Makino S (1993) Analysis of *cis*-acting sequences essential for coronavirus defective interfering RNA replication. Virology 197:53–63

Kim YN, Lai MMC, Makino S (1993) Generation and selection of coronavirus defective interfering RNA with large open reading frames by RNA recombination and possible editing. Virology 194:244–253

Kocherhans R. Bridgen A, Ackermann M, Tobler K (2001) Completion of the porcine epidemic diarrhoea coronavirus (PEDV) genome sequence. Virus Genes 23:137–144

Kozak M (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem 266:19867–19870

Lai MMC, Cavanagh D (1997) The molecular biology of coronaviruses. Adv Virus Res 48:1-100

Lai MMC, Holmes KC (2001) *Coronaviridae*: the viruses and their replication. In Fields Virology, Fourth Edition (Knipe DM, Howley PM, eds), Lippincott Williams and Wilkins, Philadelphia, pp. 1163–1185

Lai MMC, Patton CD, Stohlman SA (1982) Further characterization of mRNAs of mouse hepatitis virus: presence of common 5′-end nucleotides. J Virol 41:557–565

Laude H, Masters PS (1995) The coronavirus nucleocapsid protein. In The Coronaviridae (S.G. Siddell, ed.), Plenum Press, New York and London, pp. 141–163

Leparc-Goffart I, Hingley ST, Chua MM, Jiang X, Lavi E, Weiss SR (1997) Altered pathogenesis of a mutant of the murine coronavirus MHV-A59 is associated with a Q159L amino acid substitution in the spike protein. Virology 239:1-10

Levis R, Weiss BG, Tsiang M, Huang H, Schlesinger S (1986) Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. Cell 44:137–145

Li HP, Huang P, Park S, Lai MMC (1999) Polypyrimidine tract-binding protein binds to the leader RNA of mouse hepatitis virus and serves as a regulator of viral transcription. J Virol 73:772–777

Li HP, Zhang X, Duncan R, Comai L, Lai MMC (1997) Heterogeneous nuclear ribonucleoprotein A1 binds to the transcription-regulatory region of mouse hepatitis virus RNA. Proc Natl Acad Sci USA 94:9544–9549

Li X, Palese P (1992) Mutational analysis of the promoter required for influenza virus virion RNA synthesis. J Virol 66:4331–4338

Liao CL, Lai MMC (1995) A *cis*-acting viral protein is not required for the replication of a coronavirus defective interfering RNA. Virology 209:428–436

Lin Y, Lai MMC (1993) Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontiguous sequence for replication. J Virol 67:6110–6118

Lin YJ, Liao CL, Lai MMC (1994) Identification of the *cis*-acting signal for minus-strand RNA synthesis of a murine coronavirus: implications for the role of minus-strand RNA in RNA replication and transcription. J Virol 68:8131–8140

Liu DX, Inglis SC (1992) Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus. J Virol 66:6143–6154

Liu Q, Johnson RF, Leibowitz JL (2001) Secondary structural elements within the 3′ untranslated region of mouse hepatitis virus strain JHM genomic RNA. J Virol 75:12105–12113

Liu Q, Yu W, Leibowitz JL (1997) A specific host cellular protein binding element near the 3′ end of mouse hepatitis genomic RNA. Virology 232:74–85

Luytjes W (1995) Coronavirus gene expression. In The Coronaviridae (S.G. Siddell, ed.), Plenum Press, New York and London, pp. 33–54

Luytjes W, Gerritsma H, Spaan WJ (1996) Replication of synthetic defective interfering RNAs derived from coronavirus mouse hepatitis virus-A59. Virology 216:174–183

Makino S, Fujioka N, Fujiwara K (1985) Structure of the intracellular defective viral RNAs of defective interfering particles of mouse hepatitis virus. J Virol 54:329–336

Makino S, Joo M, Makino JK (1991) A system for study of coronavirus mRNA synthesis: a regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion. J Virol 65:6031–6041
Makino S, Lai MMC (1989) High-frequency leader sequence switching during coronavirus defective interfering RNA replication. J Virol 63:5285–5292
Makino S, Shieh CK, Keck JG, Lai MMC (1988) Defective interfering particles of murine coronavirus: mechanism of synthesis of defective viral RNAs. Virology 163:104–111
Makino S, Shieh CK, Soe LH, Baker SC, Lai MMC (1988) Primary structure and translation of a defective interfering RNA of murine coronavirus. Virology 166:1-11
Makino S, Yokomori K, Lai MMC (1990) Analysis of efficiently packaged defective interfering RNAs of murine coronavirus: localization of a possible RNA-packaging signal. J Virol 64:6045–6053
Masters PS (1999) Reverse genetics of the largest RNA viruses. Adv Virus Res 53:245–264
Masters PS, Koetzner CA, Kerr CA, Heo Y (1994) Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in the coronavirus mouse hepatitis virus. J Virol 68:328–337
Mendez A, Smerdou C, Izeta A, Gebauer F, Enjuanes L (1996) Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. Virology 217:495–507
Molenkamp R, van Tol H, Rozier BCD, van der Meer Y, Spaan WJM, Snijder EJ (2000) The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. J Gen Virol 81:2491–2496
Molla A, Paul AV, Wimmer E (1991) Cell-free, de novo synthesis of poliovirus. Science 254:1647–1651
Morris DR, Geballe AP (2000) Upstream open reading frames as regulators of mRNA translation. Mol Cell Biol 20:8635–8642
Nanda SK, Leibowitz JL (2001) Mitochondrial aconitase binds to the 3′ untranslated region of the mouse hepatitis virus genome. J Virol 75:3352–3362
Narayanan K, Chen C-J, Maeda J, Makino S (2003) Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal. J Virol 77:2922–2927
Narayanan K, Makino S (2001) Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA packaging. J Virol 75:9059–9067
Nelson GW, Stohlman SA, Tahara SM (2000) High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. J Gen Virol 81:181–188
Novak JE, Kirkegaard K (1991) Improved method for detecting poliovirus negative strands used to demonstrate specificity of positive-strand encapsidation and the ratio of positive to negative strands in infected cells. J Virol 65:3384–3387
Okumura A, Machii K, Azuma S, Toyoda Y, Kyuwa S (1996) Maintenance of pluripotency in mouse embryonic stem cells persistently infected with murine coronavirus. J Virol 70:4146–4149
Penzes Z, Gonzalez JM, Calvo E, Izeta A, Smerdou C, Mendez A, Sanchez CM, Sola I, Almazan F, Enjuanes L (2001) Complete genome sequence of transmissible gas-

troenteritis coronavirus PUR46-MAD clone and evolution of the Purdue virus cluster. Virus Genes 23:105–118
Penzes Z, Tibbles K, Shaw K, Britton P, Brown TDK, Cavanagh D (1994) Characterization of a replicating and packaged defective RNA of avian coronavirus infectious bronchitis virus. Virology 203:286–293
Penzes Z, Wroe C, Brown TDK, Britton P, Cavanagh D (1996) Replication and packaging of coronavirus infectious bronchitis virus defective RNAs lacking a long open reading frame. J Virol 70:8660–8668
Raman S, Bouma P, Williams GD, Brian DA (2003) Stem-loop III in the 5′ UTR is a *cis*-acting element in bovine coronavirus DI RNA replication. J Virol in press
Repass JF, Makino S (1998) Importance of the positive-strand RNA secondary structure of a murine coronavirus defective interfering RNA internal replication signal in positive-strand RNA synthesis. J Virol 72:7926–7933
Sarma JD, Hingley ST, Lai MMC, Weiss SR, Lavi E (1999) Direct submission to GenBank.
Sawicki SG, Sawicki DL (1990) Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in mRNA synthesis. J Virol 64:1050–1056
Sawicki SG, Sawicki DL (1995) Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. Adv Exp Med Biol 380:499–506
Sawicki SG, Sawicki DL (1998) A new model for coronavirus transcription. Adv Exp Med Biol 280:215–218.
Sawicki DL, Wang T, Sawicki SG (2001) The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. J Gen Virol 82:385–396
Schaad MC, Baric RS (1994) Genetics of mouse hepatitis virus transcription: evidence that subgenomic negative strands are functional templates. J Virol 68:8169–8179
Schochetman G, Stevens RH, Simpson RW (1977) Presence of infectious polyadenylated RNA in the coronavirus avian bronchitis virus. Virology 77:772–782
Senanayake SD, Brian DA (1999) Translation from the 5′ UTR of mRNA 1 is repressed, but that from the 5′ UTR of mRNA 7 is stimulated in coronavirus-infected cells. J Virol 73:8003–8009
Sethna PB, Brian DA (1997) Coronavirus subgenomic and genomic minus-strand RNAs copartition in membrane-protected replication complexes. J Virol 71:7744–7749
Sethna PB, Hofmann MA, Brian DA (1991) Minus-strand copies of replicating coronavirus mRNAs contain antileaders. J Virol 65:320–325
Sethna, PB, Hung SL, Brian DA (1989) Coronavirus subgenomic minus-strand RNA and the potential for mRNA replicons. Proc Natl Acad Sci USA 86:5626–5630
Shi ST, Schiller JJ, Kanjanahaluethai, A, Baker SC, Oh JW, Lai MMC (1999) Colocalization and membrane association of murine hepatitis virus gene 1 products and de novo-synthesized viral RNA in infected cells. J Virol 73:5957–5969
Siddell, SG (1995) The *Coronaviridae.* In The Coronaviridae (S.G. Siddell, ed.), Plenum Press, New York and London, pp. 1–10
Snijder EJ, Meulenberg JJM (1998) The molecular biology of arteriviruses. J Gen Virol 79:961–979

Spagnolo JF, Hogue BG (2000) Host protein interactions with the 3′ end of bovine coronavirus RNA and the requirement of the poly(A) tail for coronavirus defective genome replication. J Virol 74:5053–5065

Stirrups K, Shaw K, Evans S, Dalton K, Cananagh D, Britton P (2000) Leader switching occurs during the rescue of defective RNAs by heterologous strains of the coronavirus infectious bronchitis virus. J Gen Virol 81:791–801

Stohlman SA, Baric RS, Nelson GN, Soe LH, Welter LM, Deans RJ (1988) Specific interactions between coronavirus leader RNA and nucleocapsid protein. J Virol 62:4288–4295

Stohlamn SA, Bergmann CC, Perlman S (1999) Selected animal models of viral persistence: mouse hepatitis virus. In Persistent Viral Infections (Ahmed R, Chen ISY, eds) John Wiley and Sons, New York, pp. 537–557

Tahara SM, Dietlin TA, Nelson GW, Stohlman SA, Manno DJ (1998) Translation effector properties of mouse hepatitis virus nucleocapsid protein. Adv Exp Med Biol 440:313–318

Thiel V, Herold J, Schelle B, Siddell SG (2001) Infectious RNA transcribed in vitro from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. J Gen Virol 82:1273–1281

Thiel V, Herold J, Schelle B, Siddell SG (2001) Viral replicase gene products suffice for coronavirus discontinuous transcription. J Virol 75:6676–6681

Thiel V, Siddell SG (1994) Internal ribosomal entry in the coding region of murine hepatitis virus mRNA 5. J Gen Virol 75:3041–3046

van der Meer Y, Snijder EJ, Dobbe JC, Schleich S, Denison MR, Spaan WJ, Locker JK (1999) Localization of mouse hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. J Virol 73:7641–7657.

van der Most RG, Bredenbeek PJ, Spaan WJM (1991) A domain at the 3′ end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. J Virol 65:3219–3226

van der Most RG, Luytjes W, Rutjes S, Spaan WJM (1995) Translation but not the encoded sequence is essential for the efficient propagation of the defective interfering RNAs of the coronavirus mouse hepatitis virus. J Virol 69:3744–3751

van der Most RG, Spaan WJM (1995) Coronavirus replication, transcription, and RNA recombination. In The Coronaviridae (S.G. Siddell, ed.), Plenum Press, New York and London, pp. 11–31

van Marle G, van Dinten LC, Spaan WJM, Luytjes W, Snijder EJ (1999) Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. J Virol 73:5274–5281

Williams AK, Wang L, Sneed LW, Collisson EW (1993) Analysis of a hypervariable region in the 3′ non-coding end of the infectious bronchitis virus genome. Virus Res 28:19–27

Williams GD, Chang RY, Brian DA (1999) A phylogenetically conserved hairpin-type 3′ untranslated region pseudoknot functions in coronavirus RNA replication. J Virol 73:8349–8355

Woo K, Joo M, Narayanan K, Kim KH, Makino S (1997) Murine coronavirus packaging signal confers packaging to nonviral RNA. J Virol 71:824–827

Wu HY, Guy JS, Yoo D, Vlasak R, Urbach E, Brian DA (2003) Common RNA replication signals exist among group 2 coronaviruses: evidence for in vivo recombination between animal and human coronavirus molecules. Virology 315:174–183

Yoo D, Pei Y (2001) Full-length genomic sequence of bovine coronavirus (31 kb). Adv Exp Med Biol 494:73–76

You S, Falgout B, Markoff L, Padmanabhan R (2001) In vitro RNA synthesis from exogenous dengue viral RNA templates requires long range interactions between 5′- and 3′-terminal regions that influence RNA structure. J Biol Chem 276:15581–15591

Yount B, Curtis KM, Baric RS (2000) Strategy for systematic assembly of large RNA and DNA genomes: the transmissible gastroenteritis virus model. J Virol 74:10600–19611

Yount B, Denison MR, Weiss SR, Baric RS (2002) Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. J Virol 76:11065–11078

Yu W, Leibowitz JL (1995) Specific binding of host cellular proteins to multiple sites within the 3′ end of mouse hepatitis virus genomic RNA. J Virol 69:2016–2023

Zhang X, Lai MMC (1996) A 5′-proximal RNA sequence of murine coronavirus as a potential initiation site for genomic-length mRNA transcription. J Virol 70:705–711

Zhao S, Shaw K, Cavanagh D (1993) Presence of subgenomic mRNAs in virions of coronavirus IBV. Virology 196:172–178

Ziebuhr J, Snijder EJ, Gorbalenya AE (2000) Virus-encoded proteinases and proteolytic processing in the Nidovirales. J Gen Virol 81:853–879