

Coronavirus Reverse Genetics and Development of Vectors for Gene Expression

L. Enjuanes (✉) · I. Sola · S. Alonso · D. Escors · S. Zúñiga

Department of Molecular and Cell Biology, Centro Nacional de Biotecnología, CSIC, Campus Universidad Autónoma, 28049 Cantoblanco, Madrid, Spain
L.Enjuanes @ cnb.uam.es

1	Introduction	162
2	Pathogenesis Induced by Group 1 Coronaviruses	164
3	Engineering Coronavirus Genome	165
4	Essential Genes Required for TGEV Replication	168
5	Transcription-Regulating Sequences	170
5.1	Control of Transcription in TGEV	173
5.2	Effect of TRS Copy Number on Transcription	177
6	Expression Systems Based on Group 1 Coronaviruses	177
6.1	Helper-Dependent Expression Systems	178
6.2	Single Genome Coronavirus Vectors	180
6.3	Replication-Competent, Propagation-Deficient Coronavirus-Derived Expression Systems.	182
7	Coronavirus Vector Cloning Capacity	183
8	Insertion Site, Stability, and Expression Levels	184
9	Molecular Basis of Group 1 Coronavirus Tropism	186
10	Modulation of Coronavirus Vector Virulence	187
11	Biosafety in Coronavirus-Derived Vectors.	188
12	Conclusions	189
	References.	190

Abstract Knowledge of coronavirus replication, transcription, and virus–host interaction has been recently improved by engineering of coronavirus infectious cDNAs. With the transmissible gastroenteritis virus (TGEV) genome the efficient ($>40 \mu\text{g}$ per 10^6 cells) and stable (>20 passages) expression of the foreign genes has been shown. Knowledge of the transcription mechanism in coronaviruses has been significantly increased, making possible the fine regulation of foreign gene expression. A new family of vectors based on single coronavirus genomes, in which essential genes

have been deleted, has emerged including replication-competent, propagation-deficient vectors. Vector biosafety is being increased by relocating the RNA packaging signal to the position previously occupied by deleted essential genes, to prevent the rescue of fully competent viruses that might arise from recombination events with wild-type field coronaviruses. The large cloning capacity of coronaviruses (>5 kb) and the possibility of engineering the tissue and species tropism to target expression to different organs and animal species, including humans, has increased the potential of coronaviruses as vectors for vaccine development and, possibly, gene therapy.

1 Introduction

Reverse genetics for coronaviruses has been initially achieved by targeted recombination (Masters 1999) (see the chapter by Masters and Rottier, this volume). Recently, the first coronavirus infectious cDNA clones have been constructed for transmissible gastroenteritis coronavirus (TGEV) (Almazán et al. 2000; Yount et al. 2000), human coronavirus (HCoV) 229E (Thiel et al. 2001a) (see the chapter by Thiel and Siddell, this volume), severe and acute respiratory syndrome coronavirus (SARS-CoV) (Yount et al. 2003), mouse hepatitis virus (MHV) (Yount et al. 2002) (see the chapter by Baric and Sims, this volume), and avian coronavirus (Casais et al. 2001).

The construction of virus vectors derived from RNA viruses is a comprehensive process that for optimum performance requires at least (1) the availability of an infectious cDNA clone; (2) knowledge of the virus transcription mechanism to optimize mRNA levels; (3) determination of the essential and nonessential genes to create room for heterologous genes; (4) understanding of the molecular basis of virus tropism, in order to control the species- and tissue specificity of the vector; (5) control of virus virulence, in order to generate attenuated vectors; and (6) design of a strategy for vector safety. In this chapter the progress on these aspects will be reviewed. The chapter will focus on the advances in the generation of virus vectors based on coronavirus genomes by reverse genetics, frequently using the TGEV genome as a model.

TGEV is an enveloped virus containing an internal core (Fig. 1A), formed by the genomic RNA, the N protein, and the M protein carboxy-terminus. Dissociation of the core by chaotropic agents leads to the release of a helical nucleoprotein composed of the genomic RNA and the N protein. All M protein molecules are embedded within the membrane. TGEV M protein presents two topologies. In one-third of the molecules both the amino and the carboxy terminus face the outside of the virion,

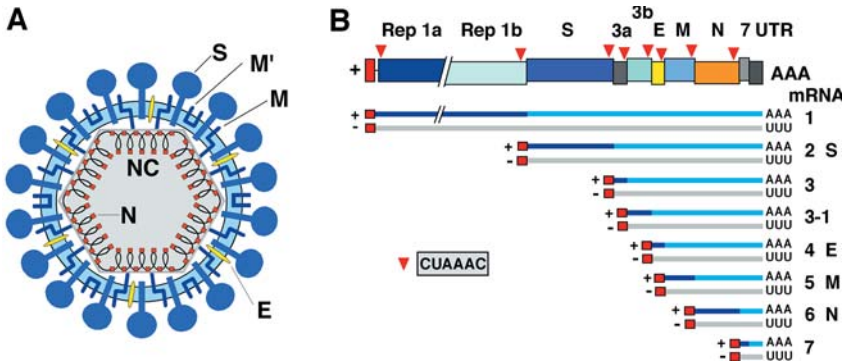


Fig. 1A, B. Coronavirus structure and genome organization. **A** Diagram of coronavirus structure using TGEV as a prototype. The scheme shows the envelope, the core, and the nucleocapsid structure. *S*, spike protein; *M* and *M'*, large membrane proteins with the amino terminus facing the external surface of the virion and the carboxy terminus toward the inside or the outside face of the virion, respectively; *E*, small envelope protein; *N*, nucleoprotein; *NC*, nucleocapsid. **B** Representation of a prototype TGEV genome and subgenomic RNAs. *Beneath the top bar* a set of positive- and negative-sense mRNA species synthesized in infected cells is shown. *Dark and semidark thin lines* (+), mRNA sequences translated and nontranslated into viral proteins, respectively. *Light lines* (-), RNAs complementary to the different mRNAs. Poly(A) and Poly(U) tails are indicated by AAA or UUU. *Rep 1a* and *Rep 1b*, replicase genes; other acronyms as in **A**

whereas in the other two-thirds the carboxy terminus is inside and is integrated within the core, being essential to maintain its structure (Escors et al. 2001a, b). In addition, the virus envelope contains two other proteins, *S* and *E* (Enjuanes et al. 2000a). The *S* protein is responsible for attachment and entry into cells and is the major inducer of TGEV-neutralizing antibodies (Suñé et al. 1990).

The TGEV genome is a single molecule of positive-sense, single-stranded RNA of 28.5 kb, which is infectious and contains eight functional genes, four of which, the spike (*S*), envelope (*E*), membrane (*M*), and nucleoprotein (*N*), encode structural proteins (Fig. 1B). The genes are arranged in the order 5'-*Rep1a-1b-S-3a-3b-E-M-N-7-3'*. TGEV mRNAs consist of seven to eight types of varying sizes, depending on the strain (Penzes et al. 2001).

2 Pathogenesis Induced by Group 1 Coronaviruses

Coronaviruses are classified in three groups according to genetic analysis (González et al. 2003; Siddell 1995). Group 1 includes coronaviruses infecting human, porcine, canine, and feline species, closely related in sequence and, in some cases, also antigenically (Sánchez et al. 1990). Coronaviruses are associated mainly with respiratory, enteric, hepatic, and central nervous system diseases. In humans and fowl, coronaviruses primarily cause upper respiratory tract infections, whereas porcine and bovine coronaviruses establish enteric infections that result in severe economic loss (USDA 2002). Human CoV are responsible for 10%–20% of all common colds and have been implicated in gastroenteritis, high- and low-respiratory tract infections, and rare cases of encephalitis (Denison 1999). HCoV have also been associated with infant necrotizing enterocolitis (Resta et al. 1985) and are tentative candidates for multiple sclerosis (Denison 1999). Recently, a new SARS-CoV has emerged, infecting more than 8,000 people and causing more than 800 deaths in 5 months (Drosten et al. 2003; Holmes and Enjuanes 2003; Marra et al. 2003; Snijder et al. 2003; Thiel et al. 2003a).

Epithelial cells are the main targets of porcine coronaviruses. Widely distributed cells such as macrophages are also infected. TGEV is an enteropathogenic coronavirus that replicates in both villous epithelial cells of the small intestine and in lung cells. A nonenteropathogenic virus related to TGEV, the porcine respiratory coronavirus (PRCV), appeared in Europe in the 1980s (Callebaut et al. 1988; Pensaert et al. 1986) and later on in North America (Vaughn and Paul 1993; Vaughn et al. 1995; Wesley et al. 1990b). This virus replicates to high titers in the respiratory tract and undergoes limited replication in submucosal cells of the small intestine (Cox et al. 1990a, b). In contrast to TGEV, PRCV infection of swine resulted no clinical signs of disease (Duret et al. 1988; Pensaert et al. 1986; Wesley et al. 1990b).

A TGEV-like disease was associated with porcine epidemic diarrhea coronavirus (PEDV) (Pensaert and De Bouck 1978). This virus, closely related to TGEV in sequence but antigenically distinct (Sánchez et al. 1990), also infects the enteric tract of swine. Nevertheless, in contrast to TGEV, PEDV does not infect the lungs. Probably, TGEV and PEDV use different receptors, because TGEV easily grows in porcine cells whereas PEDV does not and, in contrast, PEDV replicates in monkey (Vero) cells only permissive to certain strains of TGEV (J.M. Sanchez and L. Enjuanes, unpublished data).

Canine coronavirus (CCoV) usually produces a mild gastroenteritis, although some virus strains cause a more severe and sometimes fatal diarrhea. Feline coronavirus (FCoV) causes a disease involving an antibody-dependent enhancement of infection and immunocomplex-induced lesions. Two serotypes of feline enteric coronavirus have been identified that cause feline infectious peritonitis by infecting macrophages (Olsen 1993).

3 Engineering Coronavirus Genome

Two types of expression vectors have been developed for coronaviruses (Fig. 2). One requires two components (helper-dependent expression system) (Fig. 2A) and the other a single genome that is modified either by targeted recombination (Masters 1999) (Fig. 2B.1) or by engineering a cDNA encoding an infectious RNA. Infectious cDNA clones are available for porcine (Fig. 2B.2 and B.3), human (Fig. 2B.4), avian, and murine coronaviruses as indicated above. Infectious cDNA clones have also been constructed for the *Arteriviridae* family closely related to coronaviruses (de Vries et al. 2000; Meulenberg et al. 1998; van Dinten et al. 1997).

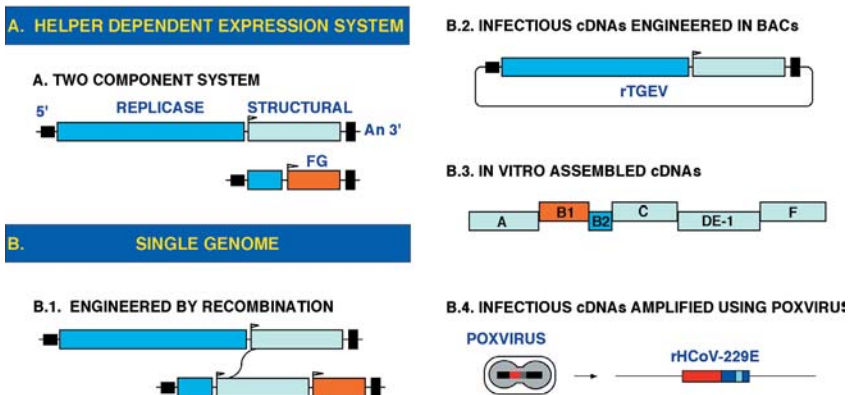


Fig. 2A, B. Coronavirus derived expression systems. A Helper-dependent expression system based on two components, the helper virus and a minigenome carrying the foreign gene (FG). An, poly A. B Single genome engineered by targeted recombination (B.1), by assembling an infectious cDNA clone derived from TGEV genome in BACs (B.2), by the in vitro ligation of six cDNA fragments (B.3), or by using poxviruses as the cloning vehicle (B.4)

An infectious coronavirus cDNA clone was first obtained for TGEV (Almazán et al. 2000; González et al. 2002). The strategy used to clone this cDNA was based on three points: (1) The construction was started from a defective minigenome (DI) that was stably and efficiently replicated by helper viruses (Izeta et al. 1999). During the filling in of minigenome deletions, a cDNA fragment that was toxic to the bacterial host was identified. This fragment was reintroduced into the cDNA in the last cloning step. (2) Transcription of the long coronavirus RNA genome, including a 5' cap, in the nucleus is essential for its infectivity. The RNA was expressed in a process mediated by the recognition of the cytomegalovirus (CMV) promoter by the cellular polymerase II. This process was followed by a second amplification in the cytoplasm driven by the viral polymerase and (3) increase of viral cDNA stability within bacteria by cloning the cDNA as a bacterial artificial chromosome (BAC) producing a maximum of two plasmid copies per cell. Following this procedure, an infectious TGEV cDNA clone that produces a virulent virus infecting both the enteric and the respiratory tract of swine was engineered (Almazán et al. 2000). The stable propagation of a TGEV full-length cDNA in bacteria as a BAC has been improved by the insertion of an intron to disrupt the toxic region identified in the viral genome (Fig. 3) (González et al. 2002). The viral RNA was transcribed in the cell nucleus under the control of the CMV promoter, and the intron was efficiently removed during translocation of this RNA to the cytoplasm. Intron insertion in two different positions allowed stable plasmid amplification for at least 200 generations. Infectious TGEV was efficiently recovered from cells transfected with the modified cDNAs. The great advantage of this system is that coronavirus reverse genetics only involves standard recombinant DNA technologies performed within bacteria. The experience in our laboratory over more than 5 years has proven that this approach leads to the efficient rescue of mutants in all viral genes.

A second procedure to assemble a full-length infectious construct of TGEV was based on the *in vitro* ligation of six adjoining cDNA subclones that spanned the entire TGEV genome. Each clone was engineered with unique flanking interconnecting junctions that determine a precise assembly with only the adjacent cDNA subclones, resulting in a full-length TGEV cDNA. *In vitro* transcripts derived from the full-length TGEV construct were infectious (Yount et al. 2000).

An infectious cDNA clone has also been constructed for HCoV-229E (Thiel et al. 2001a), another member of the Group 1 coronaviruses, MHV (Yount et al. 2002) (see the chapter by Baric and Sims, this volume) and SARS-CoV (Yount et al. 2003), which are Group 2 coronavirus

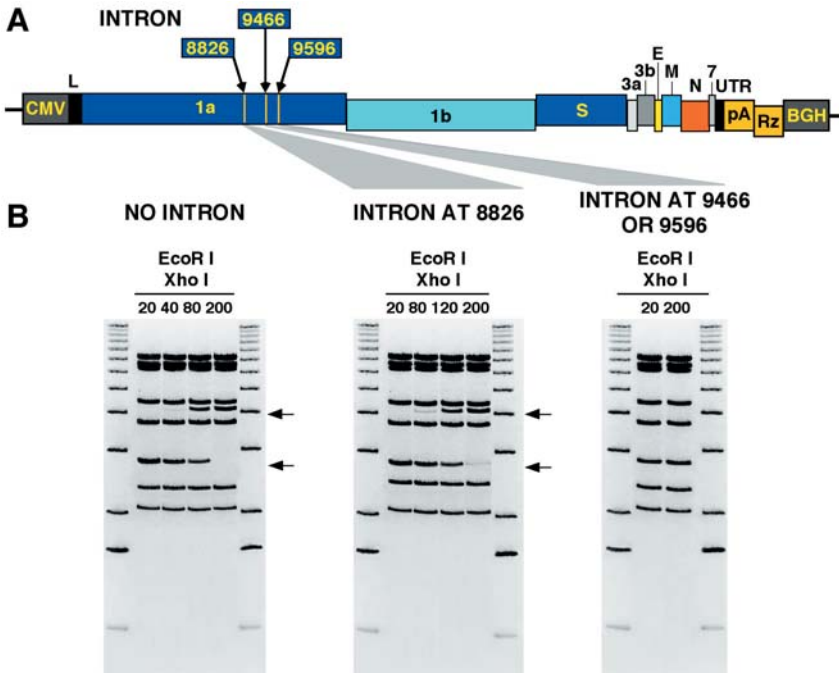


Fig. 3A, B. Intron insertion to stabilize TGEV full-length cDNA. **A** Strategy for the insertion of the 133-nt intron at the indicated positions of the TGEV sequence. **B** Analysis of the three intron-containing TGEV full-length cDNAs in *E. coli* cells. The *EcoRI-XhoI* restriction patterns of the three plasmids extracted from *E. coli* cells grown for the indicated number of generations are shown. Arrows indicate disappearance or appearance of a band. *M*, molecular mass markers

members, and IBV (Casais et al. 2001), a member of coronavirus Group 3. cDNA copies of the HCoV-229E and the MHV genomes have been cloned and propagated in vaccinia virus. Briefly, the full-length genomic cDNA clone of HCoV-229E is assembled by in vitro ligation and then cloned into the vaccinia virus DNA under the control of the T7 promoter. Recombinant vaccinia viruses containing the HCoV-229E genome are recovered after transfection of the vaccinia virus DNA into cells infected with fowlpox virus. In a second phase, the vaccinia virus DNA is purified and used as a template for in vitro transcription of HCoV-229E genomic RNA, which is transfected into susceptible cells for the recovery of infectious recombinant coronavirus. An IBV cDNA clone was assembled using the same strategy reported for HCoV-229E with some modifications. Similarly to HCoV-229E, the IBV genomic cDNA is assembled down-

stream of the T7 promoter by in vitro ligation and cloned into the vaccinia virus DNA. However, recovery of recombinant IBV is done after in situ synthesis of infectious IBV RNA by transfection of restricted recombinant vaccinia virus DNA (containing the IBV genome) into primary chicken cells previously infected with a recombinant fowlpox expressing T7 RNA polymerase.

A replicon has been constructed with the HCoV-229E genome (Thiel et al. 2001b). This replicon included the 5' and 3' ends of the HCoV-229E genome, the replicase gene of this virus, and a reporter gene coding for green fluorescent protein (GFP). RNA transcribed from this cDNA and transfected into BHK-21 cells led to only 0.1% of the cells showing strong fluorescence. This indicated that the coronavirus replicase gene products suffice for discontinuous subgenomic mRNA transcription, in agreement with the requirements for the arterivirus replicase (Molenkamp et al. 2000). Nevertheless, coexpression of N protein seems to increase rescue efficiency of infectious virus from cDNAs (Almazán et al. 2004; Thiel et al. 2003b; Yount et al. 2000), although this issue has not been systematically addressed.

A collection of replicons derived from the TGEV genome has also been constructed (Almazán et al. 2004). These replicons were launched from the cell nucleus with the CMV promoter, were efficiently rescued in the presence of N protein, expressed a heterologous gene in more than 83% of transfected cells, and had low or no cytopathogenicity for human cells.

4

Essential Genes Required for TGEV Replication

One of the most distinguishing features of the nidovirus genome is the conservation of the domain organization in the polyproteins pp1a and pp1ab, involved in genome replication, which are expressed by ribosomal frameshifting and polyprotein cleavage by viral proteases. A comparative analysis of replicative polyproteins of coronaviruses and arteriviruses identified the most variable regions in the N-terminal half of pp1a (Bonilla et al. 1994; Lee et al. 1991; Nelsen et al. 1999). Further insight into the *Coronaviridae* family showed some conserved domains in the pp1a/pp1b polyproteins between coronaviruses and toroviruses. The observation that some domains are not conserved in all coronaviruses and toroviruses indicated that these might be nonessential for the viral life cycle. In addition, all conserved arterivirus domains were found to be

smaller than their coronavirus counterparts, indicating that some sequences in the coronavirus genome could be dispensable for virus replicative functions (Gorbalenya 2001). Nevertheless, attempts to delete the replicase nonconserved domains to determine whether they are essential have not yet been made.

ORFs 3a, 3b, and 7 of TGEV encode nonstructural proteins (Enjuanes and Van der Zeijst 1995; Ortego et al. 2003; Sola et al. 2003) that are nonessential for virus replication in cell culture. Although in the enteric and virulent Miller strain of TGEV expression of mRNAs from ORFs 3a and 3b has been observed, another virulent TGEV isolate (McGoldrick et al. 1999) includes a large deletion in ORF 3a, suggesting that this ORF is not required for replication in the enteric tract or to lead to virulent isolates. In other attenuated TGEV strains with a growth essentially limited to the respiratory tract, including the Purdue strain, the subgenomic mRNA corresponding to ORF 3b is not transcribed because this gene is preceded by a noncanonical transcription-regulating sequence (TRS) (O'Connor and Brian 1999, 2000; Wesley et al. 1989). These PRCV variants have deletions of varying sizes within ORF 3a and 3b (Vaughn et al. 1995; Wesley et al. 1991). The lack of enteric tropism and attenuation of these TGEV strains has been associated with a deletion of around 670 nt located at the 5' end of the S gene and not with deletion of genes 3a and 3b (Sánchez et al. 1992, 1999).

The engineering of a TGEV genome with all the genes separated by unique restriction endonuclease sites (Ortego et al. 2003) allowed the systematic deletion of each ORF to analyze whether they are essential or dispensable for virus growth (Fig. 4). TGEV with deleted 3a and 3b genes (rTGEV- Δ 3) showed growth kinetics and mRNAs levels similar to those of the parental virus in cell cultures, demonstrating that the deletion of ORFs 3a and 3b did not affect either viral replication or transcription. In *in vivo* experiments, rTGEV- Δ 3 virus kept the replication efficiency and tropism of the wild-type virus with a very small reduction in virulence, confirming that these properties were not significantly influenced by genes 3a and 3b (Sola et al. 2003).

TGEV ORF 7 is located downstream of the essential N gene. With the TGEV infectious cDNA clone including unique engineered restriction endonuclease sites, a deletion mutant was generated (rTGEV- Δ 7) in which gene 7 expression has been abrogated (Ortego et al. 2003). The rTGEV- Δ 7 contained a deletion spanning 21 nt upstream of the ORF 7 AUG and the first nucleotides of this ORF. Recombinant rTGEV- Δ 7 showed standard kinetics in cell culture, indicating that the protein encoded by gene 7 was not essential for TGEV replication in tissue culture

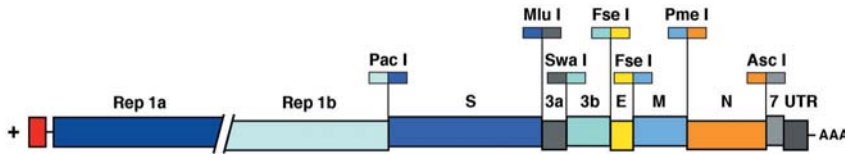


Fig. 4. Essential genes required for TGEV replication. The diagram shows a scheme of the TGEV genome in which the partial overlapping between genes has been resolved by duplicating sequences at the 5' end of each gene, starting at gene S, in order to be able to delete a gene without affecting the expression of the flanking ones. The duplicated sequences include termination codons of the preceding gene and TRSs including the CS (5'-CUAAAC-3'). In addition, the indicated unique restriction endonuclease sites (*top of the bars*) were introduced to create insertion sites for heterologous genes. The name of the gene is indicated *on top of each bar*

(Ortego et al. 2003; Sola et al. 2003). Interestingly, *in vivo* infection with rTGEV- Δ 7 showed an additional reduction in virus replication in the lung and gut compared with the parental virus and in virulence, indicating that TGEV gene 7 influences virus pathogenesis (Ortego et al. 2003).

5 Transcription-Regulating Sequences

To optimize virus vector expression levels it is essential to improve (1) replication levels without increasing virulence, (2) the accumulation levels of total mRNA, and (3) translation from mRNA. These objectives can only be achieved by knowing the mechanisms involved in these processes. To accomplish this goal, a brief review of mRNA transcription in coronavirus (particularly TGEV) is provided.

Coronavirus RNA synthesis occurs in the cytoplasm via a negative-strand RNA intermediate that contains short stretches of oligo(U) at the 5' end. Both genome-size and subgenomic negative-strand RNAs, which correspond in number of species and size to those of virus-specific mRNAs, have been detected (Brian 2001; Sawicki et al. 2001). Coronavirus mRNAs have a leader sequence at their 5' ends. Preceding every transcription unit on the viral genomic RNA, there is a conserved *core sequence* (CS) that is identical to a sequence located at the 3' end of the leader sequence (CS-L). This sequence motif constitutes part of the signal for subgenomic mRNA transcription. The common 5' leader sequence is only found at the very 5' terminus of the genome, which implies that the synthesis of subgenomic mRNAs involves fusion of non-

contiguous sequences. The mechanism involved in this process is under debate; nevertheless, the *discontinuous transcription during negative-strand RNA synthesis* model is compatible with most of the experimental evidence (Pasternak et al. 2001; Sawicki et al. 2001; van Marle et al. 1999; Zúñiga et al. 2004) (see chapter by Sawicki and Sawicki, this volume). As the leader-mRNA junction occurs during synthesis of the negative strand, within the sequence complementary to the CS (cCS), the nature of the CS is considered crucial for mRNA synthesis.

The CS motif includes six nucleotides that are highly conserved in all the genes of the same coronavirus (Fig. 5A). In addition, the 5' and 3' flanking sequences are partially conserved in the different genes of related viruses and these flanking sequences influence the activity of the CS (Fig. 5A) (Alonso et al. 2002a). Therefore, we consider that the TRSs could be divided into three sequence blocks, the CS, and the 5' (5'-TRS) and 3' (3'-TRS) flanking sequences (Fig. 5B). The most frequently used CS of coronaviruses belonging to Group 1 (hexamer 5'-CUAAAC-3') and Group 2 (heptamer 5'-UCUAAAC-3') share homology (Fig. 5C), whereas the CS of coronaviruses belonging to Group 3, like that of IBV, has the most divergent sequence (5'-CUUAACAA-3') (Fig. 5B). Also, arterivirus CSs have a sequence (5'-UCAACU-3') that partially resembles that of IBV.

Transcription levels may be influenced by many factors. Three of these, probably the most relevant, are as follows. (1) Potential base-pairing between the leader 3' end and sequences complementary to the TRS preceding the “body” (coding sequence) (cTRS-B), which guide the fusion between the nascent negative strand and the leader TRS. A minimum complementarity is required between the TRS located at the end of the leader (TRS-L) and the cTRS-B of each gene. Although several studies on the effect of the extension of this complementarity on mRNA synthesis have been performed (Alonso et al. 2002a; Joo and Makino 1992; La Monica et al. 1992; Makino and Joo 1993; Makino et al. 1991; Shieh et al. 1987; van der Most et al. 1994) the length of optimum TRSs has not been accurately defined. (2) Proximity of a gene to the 3' end. Because the TRSs are considered slow down or stop signals for the replicase complex, the smaller mRNAs should be the most abundant. Although this has been shown to be the case in the *Mononegavirales* (Wertz et al. 1998), and in coronaviruses shorter mRNAs are generally more abundant, the relative abundance of coronavirus mRNAs is not strictly related to their proximity to the 3' end (Alonso et al. 2002a; Penzes et al. 2001). (3) Potential interaction of proteins with the TRS RNA and protein-protein interactions that could regulate transcription levels. The as-

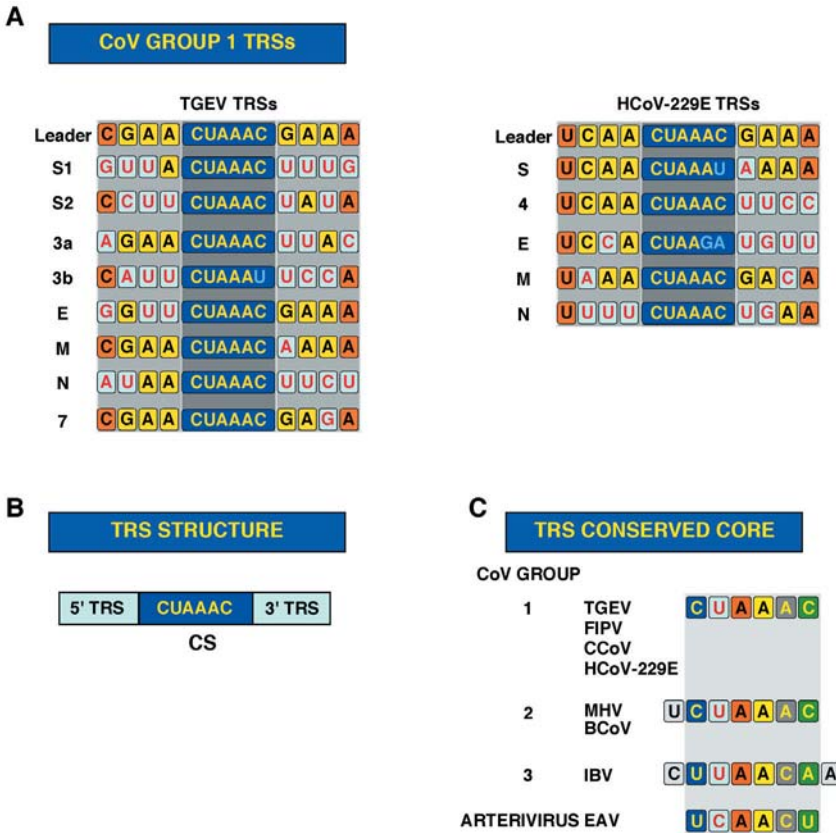


Fig. 5A–C. Coronavirus transcription-regulating sequences. **A** TRSs from two Group 1 CoVs (TGEV and HCoV-229E). CSs sequences are represented in *white letters* inside *dark boxes*. *Yellow boxes* highlight the identity among the sequences immediately flanking the CS at both 5' and 3' ends. **B** Group 1 coronavirus TRS sequence, including the highly conserved CS sequence (5'-CUAAAC-3') and the flanking 5' and 3' TRS. **C** Sequence of the most frequently used CS of coronaviruses and arteriviruses

sociation of the nascent RNA chain with the leader TRS is probably mediated by the approximation of the TRS-L to the TRS-B of each gene through RNA-protein and protein-protein interactions.

The influence of the CS in transcription has been analyzed in detail in the arteriviruses (Pasternak et al. 2001, 2003; van Marle et al. 1999) and in coronaviruses (Alonso et al. 2002a; Zúñiga et al. 2004). With infectious cDNA clones of EAV and TGEV it has been shown that mRNA syn-

thesis requires base-pairing interaction between the leader TRS and a cTRS in the nascent viral negative strand. The construction of double mutants in which a mutant leader CS was combined with the corresponding mutant of the body CS resulted in the restoration of the specific mRNA synthesis, initially suggesting that the sequence of the CS per se is not crucial, as long as the possibility for CS base-pairing is maintained. Nevertheless, it has been shown that other factors, besides leader-body base-pairing, also play a role in mRNA synthesis and that the primary sequence (or secondary structure) of TRSs may dictate strong base preferences at certain positions (Pasternak et al. 2001). Detailed analyses of the TRSs used in the arteriviruses (van Marle et al. 1999), MHV (Zhang and Liu 2000), BCoV (Ozdarendeli et al. 2001), and TGEV (Sola et al. 2003; Zúñiga et al. 2004) indicate that noncanonical CS sequences may also be used for strand transfer in the discontinuous mRNA synthesis in the *Nidovirales*.

5.1

Control of Transcription in TGEV

Template switching during synthesis of the negative RNA to join the leader is probably mediated by RNA-protein and protein-protein interactions (Fig. 6). In the RNA-protein interaction, primary or secondary RNA structure of sequences flanking the CS are probably involved. Therefore, identification of the nature and size of these sequences should help to understand the transcription mechanism. In TGEV, the TRSs have been characterized with a helper-dependent expression system based on coronavirus minigenomes, by studying the synthesis of subgenomic RNAs (sgmRNAs). TGEV TRSs include the CS (5'-CUAAAC-3'), which promotes from 100- to 1,000-fold increase of mRNA synthesis when this CS is located in the appropriate context (i.e., flanked by the appropriate 5' and 3' sequences) (Alonso et al. 2002a). The relevant sequences contributing to TRS activity have been studied by extending the CS 5' upstream and 3' downstream. Sequences from virus genes flanking the CS influenced transcription levels from moderate (10- to 20-fold variation) to complete mRNA synthesis silencing. For example, a canonical CS at nt 120 downstream of the initiation codon of the S gene did not lead to production of the corresponding mRNA because of the nature of the flanking sequences (Alonso et al. 2002a). The effect of 5' flanking sequences on transcription has also been studied using full-length genomes by inserting between genes N and 7 an expression cassette identical to those studied in the minigenome. The highest

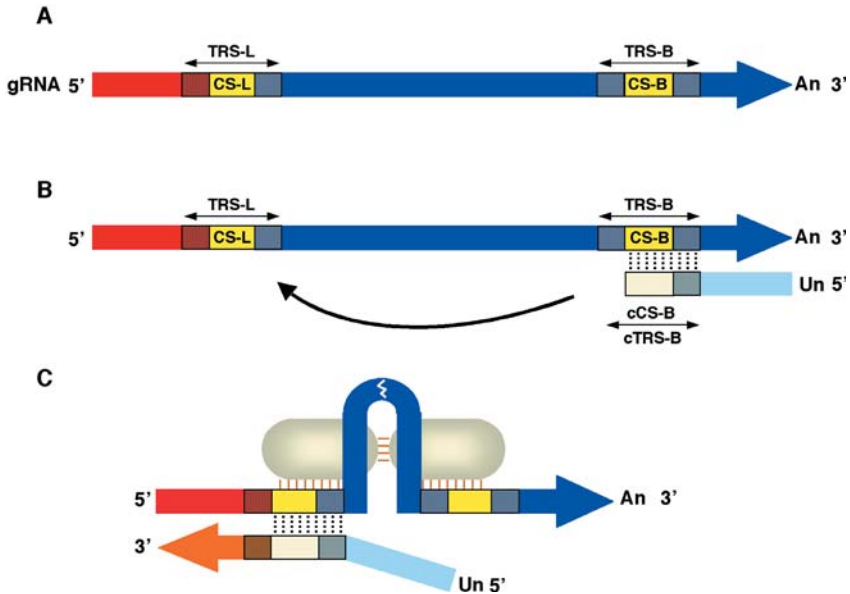


Fig. 6A–C. Diagram of the elements involved in coronavirus transcription. **A** Sequence elements probably involved in the discontinuous synthesis of the negative RNA strand. *CS-L* and *CS-B*, leader and body CSs; *TRS-L* and *TRS-B*, transcription-regulating sequences from leader and body; *An*, Poly(A). **B** Scheme of the discontinuous transcription during negative-strand synthesis and of the sequence elements involved. *cCS-B* and *cTRS-B* represent the *CS-B* and *TRS-B* complementary sequences, respectively; *Un*, Poly(U). **C** Leader and body sequences are probably located in close proximity in higher-order structures maintained by RNA-protein and protein-protein interactions

ratio of mRNA to genomic RNA was reached with 5' TRSs of about 88 nt, in agreement with the results obtained with the minigenomes and confirming the importance of sequences flanking the CS 5' end (Alonso et al. 2002a).

The influence of 3' TRSs on sgRNA synthesis has also been studied with helper-dependent expression systems (minigenomes) encoding GUS in which the CS 5' flanking sequences of the expression cassette were maintained constant. In seven constructs the CS was flanked at the 3' end by each of the sequences flanking the CS in the virus genes S, 3a, 3b, E, M, N, and 7. In addition, another construct was designed to extend in 12 nucleotides the potential base-pairing between the 3' end of the leader and cTRS. In this construct, the added nucleotides were identical to those present in the genomic RNA downstream of the CS present

at the 3' end of the leader. GUS expression by minigenomes with the CS 3' flanking sequences derived from the different virus genes differed by 10-fold, being maximum for minigenomes with a 3' TRS derived from M gene or with an extended complementarity to TRS-L. The results indicated that CS 3' flanking sequences have a large influence on sgRNA accumulation and on protein synthesis (Alonso et al. 2002a). With the helper-dependent expression system, an optimized TRS has been designed comprising 88 nt from the N gene 5' TRS, the CS, and 3 nt from the M gene 3' TRS.

With full-length TGEV genomes, the insertion of a cassette for expressing heterologous genes between the N and 7 viral genes led to a new genetic organization of the 3' end of recombinant viruses. As a consequence, a major species of sgRNAs were generated from the non-canonical CS 5'-CUAAAA-3'. It was shown that extension of complementarity between CS flanking sequences and leader RNA was associated to transcription activation from a noncanonical CS. This observation suggests that additional base-pairing between the leader RNA and sequences flanking the CS motif could compensate for the absence of complete complementarity between leader and noncanonical CSs at the junction site (Sola et al. 2003; Zúñiga et al. 2004). Interestingly, other non-canonical CSs identical to the one described above (5'-CUAAAA-3'), present throughout the TGEV genome, did not promote transcription. These data confirm that in full-length coronavirus genomes the sequences flanking the CS also play a major role regulating sgRNA expression levels.

The role of base-pairing between the nascent minus-strand RNA synthesized during transcription and the TRS-L was studied with full-length TGEV genomes (Zúñiga et al. 2004). Each nucleotide of the leader CS and the CS located 5' upstream of the nonessential gene 3a were mutated to prevent base-pairing between the TRS-L and cTRS-B. In a parallel set of experiments, complementary mutations or mutations to preserve non-Watson-Crick base-pairing were also introduced. Interestingly, the relative amounts of mRNA transcribed were related to the free energy (ΔG) of TRS-L to cTRS-B duplex formation (Zúñiga et al. 2004).

Overall, these data indicated that complementarity between the TRS-L and the nascent negative RNA strand plays a major role in coronavirus transcription regulation. The experimental data are compatible with a working model for coronavirus transcription that includes three steps (Fig. 7) (Zúñiga et al. 2004): (1) interaction between the 5' and 3' ends of coronavirus genome to bring the TRS-L in close proximity to nascent minus RNA; (2) continuous base-pairing scanning between the

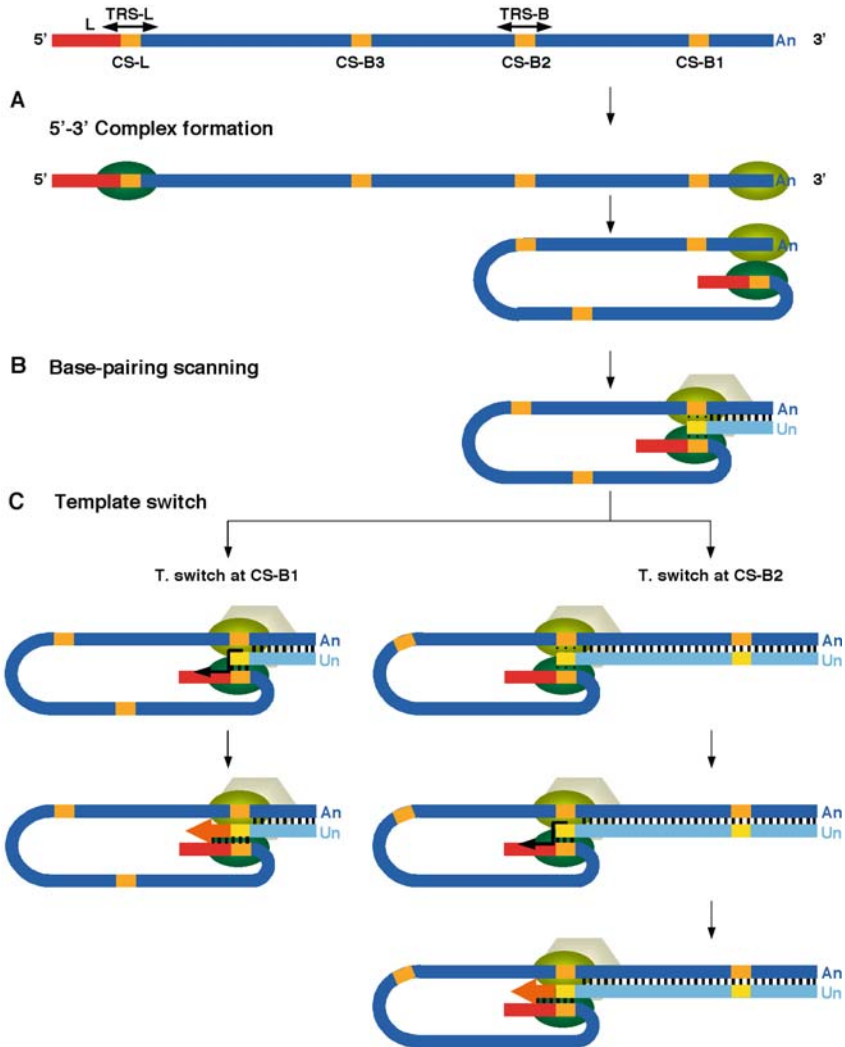


Fig. 7A–C. Three-step working model of coronavirus transcription. **A** 5'-3' complex formation step. Proteins binding the 5' and 3' end TGEV sequences are represented by the *green balls*. Leader sequence is colored in *red*, CS sequences are colored in *yellow*. *An*, poly(A) tail. **B** Base-pairing scanning step. Minus-strand RNA is in a lighter color compared with positive-strand RNA. The transcription complex is represented by the *hexagon*. *Vertical dotted bars* represent the base-pairing scanning by the TRS-L sequence in the transcription process. *Vertical solid bars* indicate complementarity between gRNA and the nascent minus strand. *Un*, poly(U) tail. **C** Template switch step. The *thin arrow* indicates the switch in the template made by the transcription complex to complete the synthesis of (–) sgRNA

TRS-L and the nascent minus RNA strand; and (3) decision between template switch to copy the leader or continuation of RNA synthesis using as template contiguous genome sequences. This model explains the formation of sgmRNAs of different length and the synthesis of alternative sgmRNAs in sequence domains with high sequence identity with the TRS-L, observed by sequencing more than 90 leader-to-body junctions of sgmRNA generated by introducing point mutations within CS-L and the CS-B preceding the gene 3a (Zúñiga et al. 2004).

5.2

Effect of TRS Copy Number on Transcription

Studies on coronavirus transcription were performed with more than one contiguous TRS in order to express the same mRNA. The accumulated sgmRNA remained nearly the same for constructs with one, two, or three TRSs, and transcription preferentially occurred in all cases at the 3'-most TRS (Joo and Makino 1995; Krishnan et al. 1996; Stirrups et al. 2000; van Marle et al. 1995). A similar result was obtained by using arterivirus recombinant viruses in which three TRSs were introduced (Pasternak et al. 2004).

6

Expression Systems Based on Group 1 Coronaviruses

Coronaviruses have several advantages as vectors over other viral expression systems: (1) Coronaviruses are single-stranded RNA viruses that replicate in the cytoplasm without a DNA intermediary, making integration of the virus genome into the host cell chromosome unlikely (Lai and Cavanagh 1997). (2) These viruses have the largest RNA virus genome and, in principle, have room for the insertion of large foreign genes (Enjuanes et al. 2000a; Masters 1999). (3) A pleiotropic secretory immune response is best induced by the stimulation of gut-associated lymphoid tissues. Because coronaviruses in general infect mucosal surfaces, both respiratory and enteric, they may be used to target the antigen to the enteric and respiratory areas to induce a strong secretory immune response. (4) The tropism of coronaviruses may be engineered by modifying the S gene (Ballesteros et al. 1997; Kuo et al. 2000; Leparco-Goffart et al. 1998; Sánchez et al. 1999). (5) Nonpathogenic coronavirus strains infecting most species of interest (human, porcine, bovine, canine, feline, and avian) are available and therefore suitable to develop

safe virus vectors. (6) Infectious coronavirus cDNA clones are available to design expression systems.

6.1

Helper-Dependent Expression Systems

Helper-dependent expression systems have been developed for coronaviruses from Groups 1, 2, and 3 (Enjuanes et al. 2001). Expression systems based on Group 1 coronaviruses have been developed for the porcine and human coronaviruses, because minigenomes are only available for these two coronavirus species. With TGEV-derived minigenomes an expression system has been assembled (Izeta et al. 1999). TGEV-derived RNA minigenomes were successfully transcribed in vitro with T7 polymerase and amplified after transfection into susceptible cells infected with a helper virus. TGEV-derived minigenomes of 3.3 (M33), 3.9 (M39), and 5.4 (M54) kb were efficiently used for the expression of heterologous genes (Alonso et al. 2002a, b). The smallest minigenome efficiently replicated and encapsidated by the helper virus was 3.3 kb in length (Izeta et al. 1999), although it has been shown that for replication and packaging of minigenomes including the β -glucuronidase (GUS) gene the most 5' 649 nt and the most 3' 278 nt may be sufficient (Escors et al. 2003).

With the M39 minigenome, a two-step amplification system was developed based on the cloning of a cDNA copy of the minigenome behind the CMV promoter (Izeta et al. 1999). Minigenome RNA is first transcribed in the nucleus by the cellular RNA *pol II*, and the RNA is then translocated into the cytoplasm, where it is amplified by the replicase of the helper virus. GUS and a surface glycoprotein (ORF5) that is the major protective antigen of the porcine respiratory and reproductive syndrome virus (PRRSV) have been expressed using this vector (Alonso et al. 2002b). GUS expression levels with an optimized TRS ranged between 2 and 8 μ g of protein per 10^6 cells. Protein levels were dependent on the extent of transcription and also on translation regulation because the presence of an appropriate Kozak context led to higher protein expression levels (Alonso et al. 2002a). Expression of GUS gene and PRRSV ORF5 with these minigenomes has been demonstrated in the epithelial cells of alveoli and in scattered pneumocytes of swine lungs, which led to the induction of a strong immune response to these antigens (Alonso et al. 2002b).

HCoV-229E, another member of coronavirus Group 1, has also been used to express new sgRNAs (Thiel et al. 1998). A synthetic RNA in-

cluding 646 nt from the 5' end plus 1,465 nt from the 3' end was amplified by the helper virus. Nevertheless, little information has been provided on the expression of heterologous genes with this system.

Most of the work in coronavirus Group 2 helper-dependent systems has been done with MHV defective RNAs (Liao et al. 1995; Lin and Lai 1993; Zhang et al. 1997). Three heterologous genes have been expressed with the MHV system, chloramphenicol acetyltransferase (CAT), hemagglutinin-esterase (HE), and interferon- γ (IFN- γ). Expression of CAT or HE was detected only in the first two passages because the minigenome used lacks the packaging signal (Liao and Lai 1995). When virus vectors expressing CAT and HE were inoculated intracerebrally into mice, HE- or CAT-specific subgenomic mRNAs were only detected in the brains at days 1 and 2, indicating that the genes in the DI vector were expressed only in the early stage of viral infection (Zhang et al. 1998). An MHV DI RNA was also developed as a vector for expressing IFN- γ . The murine IFN- γ was secreted into culture medium as early as 6 h after transfection and reached a peak level at 12 h. No inhibition of virus replication was detected with the IFN- γ produced by the DI RNA in cell cultures, but infection of susceptible mice with DI RNA producing IFN- γ caused significantly milder disease, accompanied by less virus replication than that caused by virus containing a control DI vector (Lai et al. 1997; Zhang et al. 1997).

Group 3 coronavirus-derived helper-dependent expression systems are based on IBV. A defective RNA (CD-61) derived from the Beaudette strain of the IBV virus was used as an RNA vector for the expression of two reporter genes, luciferase and CAT (Penzes et al. 1996; Stirrups et al. 2000). With IBV minigenomes, CAT expression levels between 1 and 2 μg per 10^6 cells have been described. Therefore, the highest expression levels have been obtained with a two-step amplification system based on TGEV-derived minigenomes with optimized TRSs (Alonso et al. 2002a; Izeta et al. 1999).

Expression systems based on minigenomes have the advantage of a large cloning capacity, but with the limitation of a reduced stability. For instance, TGEV-derived helper-dependent expression systems expressing GUS or PRRSV ORF5 synthesized the heterologous gene for more than five passages, but at this time smaller minigenomes were detected that finally displaced the larger ones (Alonso et al. 2002b).

6.2 Single Genome Coronavirus Vectors

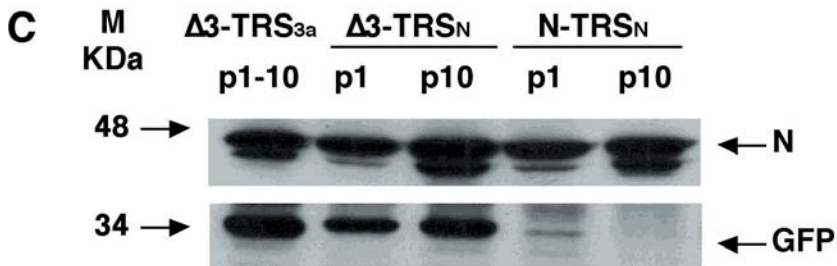
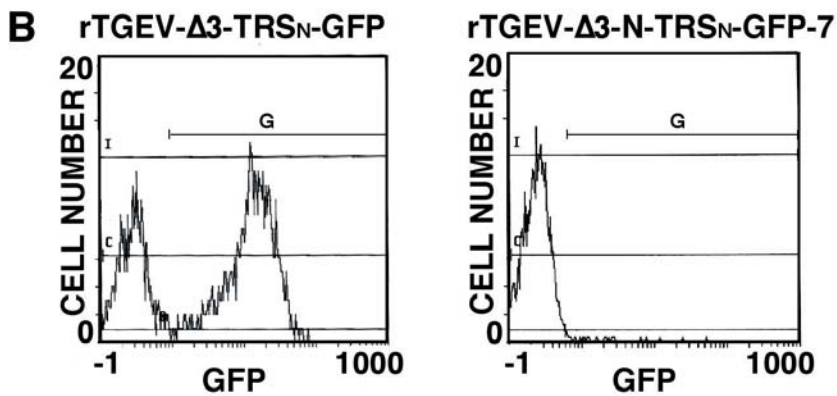
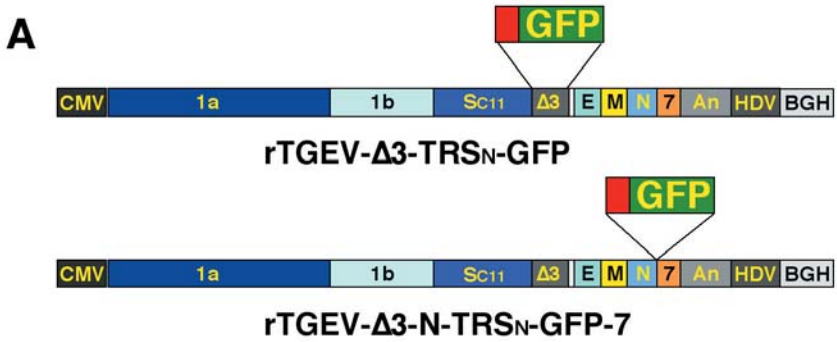
Coronavirus expression systems based on Group 1 viruses have been derived from TGEV and HCoV-229E. With a TGEV infectious cDNA, the GFP gene of 0.72 kb was cloned in two positions of the RNA genome, either by replacing the nonessential 3a and 3b genes or between genes N and 7. The engineered genome with the GFP gene at the position of ORFs 3a and 3b was very stable (>30 passages in cultured cells) and led to the production of high protein levels ($50 \mu\text{g}/10^6$ cells) (Fig. 8) (Sola et al. 2003). Therefore, expression levels with coronavirus-based vectors are similar to those described for vectors derived from other positive-strand RNA viruses such as Sindbis virus (Agapov et al. 1998; Frolov et al. 1996). The stability of viruses with expression cassettes at different positions was variable. For instance, GFP or GUS expression units inserted between genes N and 7 led to unstable viruses (Alonso et al. 2004; Sola et al. 2003). With TGEV-derived vectors expressing GFP, the acquisition of immunity by newborn piglets breast-fed by immunized sows (i.e., lactogenic immunity) was demonstrated (Sola et al. 2003).

Recombinant TGEVs have also been assembled by *in vitro* junction of six cDNA fragments encoding a full-length genome, in which GFP gene has replaced ORF3a, leading to the production of a TGEV that grew to titers of 10^8 pfu/ml and expressed GFP in a high proportion of cells (Curtis et al. 2002).

HCoV-229E has also been used as the base for expression systems using either the full-length infectious cDNA clone (Thiel et al. 2001a) or an autonomous replicating subgenomic RNA (replicon) (Thiel et al. 2003b) (see chapter by Thiel and Siddell, this volume). In each case, it has been shown that it is possible to insert transcriptional cassettes.

An infectious cDNA clone for Group 2 coronaviruses has been constructed for MHV (Yount et al. 2002). This system will provide significant advantages in the analysis of coronavirus replication and transcription by complementing a large collection of temperature-sensitive mutants of MHV. Reverse genetics in coronaviruses of Group 2 has also

Fig. 8A–C. GFP expression levels with coronavirus based vectors. **A** Schemes of BACs containing the TGEV genome in which an expression cassette with the GFP gene has been inserted replacing genes 3a and 3b (*top bar*) or between genes N and 7 (*lower bar*), respectively. **B** The relative proportion of cells expressing GFP was evaluated by



cytofluorometry. C The amount of protein synthesized by the virus vectors with expression modules inserted at genes 3a and 3b (Δ 3) or between genes N and 7 (N), using TRSs from gene 3a (*TRS_{3a}*) or from gene N (*TRS_N*) was analyzed by Western blot with GFP-specific monoclonal antibodies (α GFP). The amount of viral nucleoprotein was evaluated in the same samples with N-specific monoclonal antibodies (α N) as an internal control

been efficiently performed by targeted recombination between a helper virus and either nonreplicative or replicative coronavirus-derived RNAs and was also used to express heterologous genes. For instance, the gene encoding GFP was inserted into MHV between the S and E genes, resulting in the creation of the largest known RNA viral genome (Fischer et al. 1997). Coronavirus expression systems derived from Group 3 coronavirus are based on infectious IBV cDNA clones assembled with a strategy similar to that reported for HCoV-229E (Casais et al. 2001) (see chapter by Baric and Sims, this volume).

6.3

Replication-Competent, Propagation-Deficient Coronavirus-Derived Expression Systems

Replication-competent, propagation-deficient virus vectors, based on TGEV genomes deficient in the essential gene E obtained with E⁺ packaging cell lines have been developed (Ortego et al. 2002). Two types of cell lines expressing TGEV E protein have been selected (Fig. 9). One cell line transiently expresses E protein using the noncytopathic Sindbis virus replicon pSINrep21 (Frolov et al. 1999). Another cell line was ob-

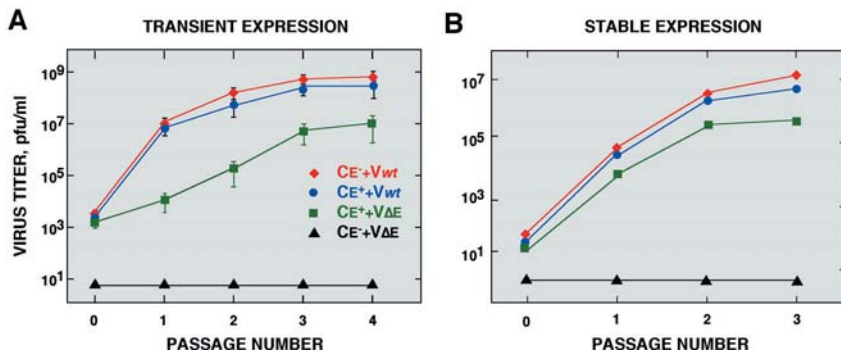


Fig. 9A, B. Rescue of recombinant TGEV- Δ E from cDNA in cells transiently or stably expressing E protein. **A** Titer of recombinant TGEVs rescued from cDNA in BHK cells expressing porcine aminopeptidase N (pAPN) (CE⁻) or the same cells transiently expressing TGEV E protein (CE⁺) with a Sindbis virus replicon. Cells were transfected either with rTGEV-*wt* (Vwt) or with rTGEV- Δ E virus (V Δ E). **B** Titer of recombinant TGEVs rescued from cDNA in BHK cells expressing pAPN (CE⁻) or the same cells stably expressing TGEV E protein (CE⁺) under the control of the CMV promoter. Cells were transfected either with rTGEV-*wt* (Vwt) or with rTGEV- Δ E virus (V Δ E). Error bars represent standard deviations of the mean from four experiments

tained in which the E gene is stably expressed under the control of CMV promoter. Recombinant TGEVs, deficient in the essential E gene, reached high titers (1×10^7 pfu/ml) in cells transiently expressing the TGEV E protein, whereas this titer was up to 5×10^5 pfu/ml in packaging cell lines stably expressing E protein. Virus titers were directly related to E protein expression levels (Ortego et al. 2002). Recovered virions showed the same morphology and stability at different pH and temperatures from the wild-type virus. The titers of the rescued viruses will most probably be increased by transforming new cell types with higher TGEV replication levels, leading to efficacious expression systems, or by increasing E protein accumulation.

A second strategy for the construction of replication-competent, propagation-deficient TGEV genomes expressing heterologous genes involves the assembly of an infectious cDNA from six cDNA fragments ligated *in vitro* (Curtis et al. 2002). The defective virus with the essential E gene deleted was complemented by the expression of E gene using the Venezuelan equine encephalitis (VEE) replicon. However, titers of recombinant virus expressing the GFP were at least 10- to 100-fold lower (around 10^4 pfu/ml) than with the system that used stably transformed cells or the Sindbis virus vector to complement E gene deletion (Ortego et al. 2002).

A multigene RNA replicon based on HCoV-229E has been developed (Thiel et al. 2003b) containing the 5' and 3' ends of this virus, the entire human coronavirus replicase gene, and three reporter genes [i.e., CAT gene, luciferase (LUC) gene, and the GFP gene]. Each reporter gene is located downstream of a human coronavirus TRS, which is required for the synthesis of individual mRNAs. The transfection of the vector and human coronavirus nucleocapsid protein mRNA into BKH-21 cells resulted in the expression of the CAT, LUC, and GFP reporter proteins. In addition, it has been shown that human coronavirus-based vector RNA can be packed into propagation-deficient pseudovirions that, in turn, can be used to transduce immature and mature human dendritic cells.

7

Coronavirus Vector Cloning Capacity

Coronavirus helper-dependent expression systems based on minigenomes have a theoretical cloning capacity close to 27 kb, because an RNA with a size of about 3 kb is efficiently amplified and packaged by the helper virus and the virus genome has about 30 kb. In contrast, the

theoretical cloning capacity for an expression system based on a single coronavirus genome, like TGEV, with the current available knowledge, is between 3 and 3.5 kb, taking into account that (1) the nonessential genes 3a (0.2 kb), 3b (0.73 kb), and 7 (0.24 kb) have been deleted, leading to a viable virus (Sola et al. 2003); (2) the standard S gene can be replaced by the S gene of a PRCV mutant with a deletion of 0.67 kb; and (3) both DNA and RNA viruses may accept genomes with sizes up to 105% of the wild-type genome.

In propagation-deficient coronavirus vectors, in which one or more of the essential genes S, E, M, and N would have been deleted, the cloning capacity could be increased over the former 3.5 kb by an additional 0.25–4.5 kb. This cloning capacity will probably be enlarged in the near future when nonessential domains of the replicase gene are deleted (see above).

8 Insertion Site, Stability, and Expression Levels

TGEV-derived helper-dependent expression systems have the advantage of a high cloning capacity. In contrast, they have a limited stability mainly derived from the presence of the foreign gene, because TGEV minigenomes of 9.7, 5.4, 3.9, and 3.3 kb, in the absence of the heterologous gene, were amplified, packaged, and efficiently propagated for at least 30 passages, without generating new dominant minigenome RNAs (Izeta et al. 1999; Méndez et al. 1996). In contrast, the insertion of genes such as GUS, TGEV S protein, and PRRSV ORF5 in the M39 minigenome led to the appearance of new smaller minigenomes that could easily be detected at passage 5. Nevertheless, with this minigenome heterologous gene expression was observed for about 10 passages. Minigenome stability is highly dependent on the nature of the foreign gene, as TGEV or IBV minigenomes expressing luciferase gene were lost in the first 2–3 passages (Stirrup et al. 2000).

The stability of the expression systems is also conditioned by the type of polymerases involved in minigenome amplification and mRNA transcription (Agapov et al. 1998). The TGEV-derived vector is based on expression of the minigenome using the CMV promoter. In this case, eukaryotic RNA polymerase II expresses the minigenome with an estimated error frequency of 5×10^{-6} (de Mercoyrol et al. 1992), which is lower than the error accumulation frequency of 10^{-4} to 10^{-5} during in vitro transcription of minigenome RNAs with T7 DNA-dependent RNA-poly-

merase (Boyer et al. 1992; Sooknanan et al. 1994). In addition, the eukaryotic RNA polymerase II has additional mechanisms to ensure even more accurate transcription (Thomas et al. 1998). After transfection of in vitro produced RNA, synthesis of DI-RNA and mRNAs by the viral RNA-dependent RNA-polymerase should have an accumulation of mutations with a relatively higher frequency of 10^{-3} to 10^{-4} (de Mercoyrol et al. 1992; Ward et al. 1988). Therefore, an improvement in expression stability should be observed by using expression systems initiated by DNA transfection.

To study the effect of TRS position within the minigenome on expression levels, a cassette encoding the GUS gene was inserted at different nucleotide distances (0.9, 1.6, 2.8, and 3.3 kb) from the 5' end of the TGEV minigenome M39 (Alonso et al. 2002a). The mRNA levels were high in the two insertion sites closer to the 3' end of the minigenome and slightly increased in the most 3' site. In contrast, in a minigenome derived from the HCoV-229E, an expression cassette was cloned into three different positions (at 1.1, 1.3, and 1.8 kb from the 5' end of a minigenome of about 2.1 kb) and the mRNA levels decreased for inserts located closer to the 3' end (Thiel et al. 1998). The experiments performed with TGEV and HCoV-229E apparently led to different results. With TGEV, and possibly with HCoV-229, some insertion sites were too close to the ends of the minigenomes and may have affected essential primary or secondary structures required for their replication. It is possible that in coronavirus the variation of expression levels with insertion site is mainly influenced by the sequences flanking the TRS in each position and that the relative position itself plays a less prominent role. In fact, in a systematic study using MHV, a 0.4-kb region including a TRS of 12 nt flanked upstream and downstream by 0.2-kb fragments was inserted at seven different positions of a 2.2-kb minigenome (Jeong et al. 1996). The 12-nt TRS core was flanked by upstream and downstream sequences in order to prevent the influence of variable flanking sequences within the different insertion sites. In all insertion sites, the level both of the minigenome and of the mRNA produced were similar, suggesting that the position of the insert along the minigenome had little influence on the mRNA expression level.

In full-length genomes, a correlation between the proximity to the 3' end of the genome and the relative efficiency of mRNA synthesis from a given TRS has been observed in several viral systems, including coronaviruses such as MHV and TGEV (Alonso et al. 2002a; de Haan et al. 2003; Hiscox et al. 1995; Sola et al. 2003; van Marle et al. 1995), arteriviruses such as EAV (Pasternak et al. 2003), and the *Mononegavirales*

(Iverson and Rose 1981; Wertz et al. 1998). In the TGEV single genome vector, in order to increase heterologous gene expression levels, an expression cassette encoding the GFP gene was inserted at the 3' end of the genome. Insertion of the expression cassette between TGEV genes N and 7 resulted in an unstable virus, leading to the complete deletion of the additional transcriptional unit. In contrast, insertion of similar expression cassettes replacing ORFs 3a and 3b led to stable expression of GFP (Sola et al. 2003). Therefore, the location of the insertion, and not the nature of the gene, was most likely responsible for the instability. The origin of TGEV recombinant virus instability was mediated either by homologous recombination promoted by the presence of duplicated viral sequences or by nonhomologous recombination yielding a virus that had lost the GFP gene and also the 5' end of gene 7. Therefore, in addition to similarity-essential recombination, similarity-nonessential recombination (Nagy and Simon 1997) may also lead to the instability of these viruses. The instability of expression cassettes inserted at the 3' end of the genome seems a general phenomenon because, in addition, it has been shown that several expression modules encoding the GUS gene were also unstable at this position of the genome but not at the ORF 3a site (Alonso et al. 2004). Furthermore, insertion of other sequence fragments (i.e., 3' end 141 nt of N gene or 717 nt of GFP) between the N gene and the 3' UTR of MHV also produced genomic instability (Hsue and Masters 1999).

9

Molecular Basis of Group 1 Coronavirus Tropism

Group 1 coronaviruses attach to host cells through the S glycoprotein by recognition of pAPN, the cell receptor (Delmas et al. 1992; Yeager et al. 1992). Group 2 coronaviruses use the carcinoembryonic antigen-related cell adhesion molecules (CEACAM) as receptors, or the angiotensin-converting enzyme 2 in the case of SARS-CoV (Li et al. 2003). Engineering the S gene led to changes both in tissue- and species specificity (Ballesteros et al. 1997; Kuo et al. 2000; Leparc-Goffart et al. 1998; Sánchez et al. 1999). Driving vector expression to different tissues may preferentially induce a specific type of immune response, that is, mucosal immunity by targeting the expression to gut-associated lymph nodes. Both tissue- and species specificity have been modified by engineering coronavirus genomes.

Enteric infection by TGEV requires virus binding to pAPN, mediated by an S protein domain encoded by nt 1,518–2,184, and also to a second factor (possibly binding to a coreceptor) mediated by an S protein domain encoded by nt 217–665. In fact, changes of two amino acids included in this S protein domain suffice to restore the enteric infection by a respiratory virus strain (Sánchez et al. 2004). Therefore, the S protein domain encoded by nt 217–665 dictates TGEV enteric tropism (Sánchez et al. 1999, 2004), whereas binding to APN alone is not sufficient to infect the enteric tract. Interestingly, MHV tropism is also influenced by the binding of a domain located at the N-terminus of murine coronavirus spike protein, in positions equivalent to those required for the potential coreceptor binding in TGEV (Kubo et al. 1994).

TGEV species specificity has been extended to infect canine and human cells by replacing the S gene of TGEV with S genes from canine (Riquelme et al. 2004) and human (Ortego et al. 2004) coronaviruses. In this case, a replication-competent, propagation-deficient TGEV has been used for safety because these viruses can only grow in packaging cell lines; therefore, they cannot be propagated from cell to cell in the host.

Animal model systems based on transgenic mice expressing the human APN (hAPN) are being developed (Lassnig et al. 2004; Wentworth and Holmes 2001; Wentworth et al. 2001) to study the molecular basis of human coronavirus, vector-host interaction (i.e., potential side-effects), and immune responses elicited by the virus vector. The transgenic mice express high amounts of hAPN in lungs, gut, spleen, liver, and brain. Cells derived from these mice replicate HCoV-229E (Lassnig et al. 2004; Wentworth and Holmes 2001; Wentworth et al. 2001), and infection of transgenic mice has been shown in one of the systems (Lassnig et al. 2004). The possibility of engineering coronavirus vectors with defined tissue- or species specificity and the development of laboratory animal model systems increases the potential use of this novel vector family.

10

Modulation of Coronavirus Vector Virulence

Tropism changes in general lead to a change in virulence. Certainly this is the case in porcine coronaviruses with a virulence directly related to their ability to grow in the enteric tract (Sánchez et al. 1999). Porcine coronaviruses exclusively growing in the lungs with titers higher than 1×10^6 pfu/g of tissue lead to no obvious clinical symptoms. Changes in the S gene modify virus tropism and also virulence (das Sarma et al.

2000; Gallagher and Buchmeier 2001; Leparç-Goffart et al. 1998; Navas et al. 2001; Phillips et al. 1999; Sánchez et al. 1999; Taguchi et al. 1995). Similarly, cell-to-cell virus spreading is influenced by S protein-dependent fusion activity, also affecting coronavirus pathogenicity (Gallagher and Buchmeier 2001; Luo et al. 1999).

Most coronavirus genes partially overlap, and TGEV is not an exception. To study the effect of nonessential gene deletion on virulence, unique restriction endonuclease sites and sequence duplications were introduced at the 5' end of each TGEV gene as shown above (Fig. 4). Because the TRS is also located at the 5' end of each gene, the insertion of duplicated sequences and of restriction endonuclease sites could alter virus pathogenicity. In fact, with the use of TGEV it has been shown that the introduction of one restriction endonuclease site between each pair of adjacent genes leads to a decrease in enteric tract virus growth and virulence by more than 10-fold and 5-fold, respectively (Ortego et al. 2003). The simultaneous modification of the 5' end of all essential genes (S, E, M, and N) led to a virus growth and virulence reduction of 100-fold and 12-fold, respectively. Therefore, this approach can be used to control virus vector virulence.

Gene expression among the nonsegmented negative-stranded RNA viruses is controlled by the highly conserved order of genes relative to the single transcriptional promoter. Rearrangement of the genes of vesicular stomatitis virus eliminates clinical disease in the natural host and is considered a new strategy for vaccine development (Flanagan et al. 2001). In coronavirus, a change in gene order led to virus attenuation in FIPV (see chapter by Masters and Rottier, this volume) (de Haan et al. 2002).

Deletion of nonessential genes 3a and 3b led to variable results in Group 1 coronaviruses. Although deletion of these genes reduced TGEV virulence very little (Sola et al. 2003), FIPV was clearly attenuated (Haijema et al. 2003). In contrast, deletion of the nonessential gene 7 attenuated both TGEV and FIPV. Therefore, modification of certain nonessential genes may be used as an efficient approach to reduce the virulence of coronavirus vectors.

11 Biosafety in Coronavirus-Derived Vectors

Application of virus vectors to humans requires a reduction of the risk to levels below those of conventional medical interventions (i.e., admin-

istration of a safe vaccine). Coronavirus vectors based on the TGEV genome have been engineered to infect human cells by replacing the porcine coronavirus S gene with that of human coronaviruses (Ortego et al. 2004). To increase the safety of the human-adapted vector, a replication-competent, propagation-deficient virus, in which two essential genes (E and N) were deleted, is being modified by introducing mutations that abrogate the activity of the RNA packaging signal (Ψ) and relocating an active Ψ between the two deleted genes (Escors et al. 2003). A recombination event leading to the recovery of the essential genes will most likely lead to loss of the packaging signal, generating a nonviable virus.

12 Conclusions

Both helper-dependent expression systems, based on two components, and single genome vectors constructed by targeted recombination, or by using infectious cDNAs, have been developed for coronaviruses. The sequences that regulate transcription have been characterized with helper-dependent expression systems and full-length infectious cDNA clones. Minigenome-based expression systems have the advantage of their large cloning capacity, in principle higher than 27 kb, produce reasonable amounts of heterologous antigens ($2\text{--}8\ \mu\text{g}/10^6$ cells), show a limited stability (synthesis of heterologous gene is maintained for around 10 passages), and elicit strong immune responses. In contrast, coronavirus vectors based on single genomes have at present a limited cloning capacity (around 5 kb) and expression levels of heterologous genes are 10-fold higher than those of helper-dependent systems ($>50\ \mu\text{g}/10^6$ cells) and are very stable (>30 passages). The possibility of expressing different genes under the control of TRSs with programmable strength and engineering tissue and species tropism indicates that coronavirus vectors are very flexible. High expression levels have been obtained with replication-competent, propagation-deficient vectors based on coronavirus genomes. Thus coronavirus-based vectors are emerging with high potential for vaccine development and, possibly, for gene therapy.

Acknowledgements This work has been supported by grants from the Comisión Interministerial de Ciencia y Tecnología (CICYT), La Consejería de Educación y Cultura de la Comunidad de Madrid, and Fort Dodge Veterinaria from Spain and the by European Communities (Life Sciences Program, Key Action 2: Infectious Diseases). DE, IS, SZ, and SA received contracts from the EU projects QLRT-1999-30739, QLRT-2000-00874, and QLRT-1999-00002, respectively).

References

- Agapov EV, Frolov I, Lindenbach BD, Pragai BM, Schlesinger S, Rice CM (1998) Non-cytopathic Sindbis virus RNA vectors for heterologous gene expression. *Proc Natl Acad Sci USA* 95:12989–12994
- Almazán F, González JM, Péntzes Z, Izeta A, Calvo E, Plana-Durán J, Enjuanes L (2000) Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc Natl Acad Sci USA* 97:5516–5521
- Almazán F, Galán C, Enjuanes L (2004) The nucleoprotein is required for efficient coronavirus genome replication. In press
- Alonso S, Izeta A, Sola I, Enjuanes L (2002a) Transcription regulatory sequences and mRNA expression levels in the coronavirus transmissible gastroenteritis virus. *J Virol* 76:1293–1308
- Alonso S, Sola I, Teifke J, Reimann I, Izeta A, Balach M, Plana-Durán J, Moormann RJM, Enjuanes L (2002b) In vitro and in vivo expression of foreign genes by transmissible gastroenteritis coronavirus-derived minigenomes. *J Gen Virol* 83:567–579
- Alonso S, Sola I, Zúñiga S, Plana-Durán J, Enjuanes L (2004) Induction of neutralizing antibodies against porcine respiratory and reproductive syndrome virus antibodies (PRRSV) ORF 5 by coronavirus derived vectors. Submitted for publication
- Ballesteros ML, Sánchez CM, Enjuanes L (1997) Two amino acid changes at the N-terminus of transmissible gastroenteritis coronavirus spike protein result in the loss of enteric tropism. *Virology* 227:378–388
- Bonilla PJ, Gorbalenya AE, Weiss SR (1994) Mouse hepatitis virus strain A59 RNA polymerase gene ORF 1a: heterogeneity among MHV strains. *Virology* 198:736–740
- Boyer JC, Bebenek K, Kunkel TA (1992) Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates. *Proc Natl Acad Sci USA* 89:6919–6923
- Brian DA (2001) Nidovirus genome replication and subgenomic mRNA synthesis. Pathways followed and *cis*-acting elements required. In: Lavi E, Weiss S and Hingley ST (eds) *Nidoviruses*. Plenum Press, New York Adv. Exp. Med. Biol., vol 494, pp 415–428
- Callebaut P, Correa I, Pensaert M, Jiménez G, Enjuanes L (1988) Antigenic differentiation between transmissible gastroenteritis virus of swine and a related porcine respiratory coronavirus. *J Gen Virol* 69:1725–1730
- 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
- Cox E, Hooyberghs J, Pensaert MB (1990a) Sites of replication of a porcine respiratory coronavirus related to transmissible gastroenteritis virus. *Res Vet Sci* 48:165–169
- Cox E, Pensaert MB, Callebaut P, van Deun K (1990b) Intestinal replication of a porcine respiratory coronavirus closely related antigenically to the enteric transmissible gastroenteritis virus. *Vet Microbiol* 23:237–243

- Curtis KM, Yount B, Baric RS (2002) Heterologous gene expression from transmissible gastroenteritis virus replicon particles. *J Virol* 76:1422–1434
- das Sarma J, Fu L, Tsai JC, Weiss SR, Lavi E (2000) Demyelination determinants map to the spike glycoprotein gene of coronavirus mouse hepatitis virus. *J Virol* 74:9206–9213
- de Haan CAM, Volders H, Koetzner CA, Masters PS, Rottier PJM (2002) Coronavirus maintain viability despite dramatic rearrangements of the strictly conserved genome organization. *J Virol* 76:12491–12502
- de Haan CAM, van Genne L, Stoop JN, Volders H, Rottier JMP (2003) Coronaviruses as vectors: position dependence of foreign gene expression. *J Virol* 77:11312–11323
- Delmas B, Gelfi J, L'Haridon R, Vogel LK, Norén O, Laude H (1992) Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. *Nature* 357:417–420
- de Mercoyrol L, Corda Y, Job C, Job D (1992) Accuracy of wheat-germ RNA polymerase II. General enzymatic properties and effect of template conformational transition from right-handed B-DNA to left-handed Z-DNA. *Eur J Biochem* 206:49–58
- Denison MR (1999) The common cold. Rhinoviruses and coronaviruses. In: Dolin R and Wright PF (eds) *Viral infections of the respiratory tract*. Marcel Dekker, Inc., New York Lung Biology in Health and Disease, vol 127, pp 253–280
- de Vries AAF, Glaser AL, Raamsman MJB, de Haan CAM, Sarnataro S, Godeke GJ, Rottier PJM (2000) Genetic manipulation of equine arteritis virus using full-length cDNA clones: separation of overlapping genes and expression of a foreign epitope. *Virolog Coronaviridae*. In: van Regenmortel MHV, Fauquet CM, Bishop DHL, Carsten EB, Estes MK, Lemon SM, McGeoch DJ, Maniloff J, Mayo MA, Pringle CR and Wickner RB (eds) *Virus taxonomy. Classification and nomenclature of viruses*. Academic Press, San Diego, California, pp 835–849
- Enjuanes L, Sola I, Almazán F, Ortego J, Izeta A, González JM, Alonso S, Sánchez-Morgado JM, Escors D, Calvo E, Riquelme C, Sánchez CM (2001) Coronavirus derived expression systems. *J Biotech* 88:183–204
- Escors D, Ortego J, Laude H, Enjuanes L (2001a) The membrane M protein carboxy terminus binds to transmissible gastroenteritis coronavirus core and contributes to core stability. *J Virol* 75:1312–1324
- Escors D, Camafeita E, Ortego J, Laude H, Enjuanes L (2001b) Organization of two transmissible gastroenteritis coronavirus membrane protein topologies within the virion and core. *J Virol* 75:12228–12240
- Escors D, Izeta A, Capiscol MC, Enjuanes L (2003) Transmissible gastroenteritis coronavirus packaging signal is located at the 5' end of the virus genome. *J Virol* 77:7890–7892
- Fischer F, Stegen CF, Koetzner CA, Masters PS (1997) Analysis of a recombinant mouse hepatitis virus expressing a foreign gene reveals a novel aspect of coronavirus transcription. *J Virol* 71:5148–5160
- Flanagan EB, Zamparo JM, Ball LA, Rodriguez L, Wertz GW (2001) Rearrangement of the genes of vesicular stomatitis virus eliminates clinical disease in the natural host: new strategy for vaccine development. *J Virol* 75:6107–6114

- Frolov I, Hoffman TA, Prágai BM, Dryga SA, Huang HV, Schlesinger S, Rice CM (1996) Alphavirus-based expression vectors: strategies and applications. *Proc Natl Acad Sci USA* 93:11371–11377
- Frolov I, Agapov E, Hoffman TA, Prágai BM, Lippa M, Schlesinger S, Rice CM (1999) Selection of RNA replicons capable of persistent noncytopathic replication in mammalian cells. *J Virol* 73:3854–3865
- Gallagher T, Buchmeier MJ (2001) Coronavirus spike proteins in viral entry and pathogenesis. *Virology* 279:371–374
- González JM, Penzes Z, Almazán F, Calvo E, Enjuanes L (2002) Stabilization of a full-length infectious cDNA clone of transmissible gastroenteritis coronavirus by the insertion of an intron. *J Virol* 76:4655–4661
- González JM, Gomez-Puertas P, Cavanagh D, Gorbalenya AE, Enjuanes L (2003) A comparative sequence analysis to revise the current taxonomy of the family *Coronaviridae*. *Arch Virol* 148:2207–2235
- Gorbalenya AE (2001) Big nidovirus genome. When count and order of domains matter. In: Lavi E, Weiss S and Hingley ST (eds) *The Nidoviruses (Coronaviruses and Arteriviruses)*. Kluwer Academic/Plenum Publishers, New York, vol 494, pp 1–17
- Hajjema BJ, Volders H, Rottier PJM (2003) Switching species tropism: an effective way to manipulate the feline coronavirus genome. *J Virol* 77:4528–4538
- Hiscox JA, Mawditt KL, Cavanagh D, Britton P (1995) Investigation of the control of coronavirus subgenomic mRNA transcription by using T7-generated negative-sense RNA transcripts. *J Virol* 69:6219–6227
- Holmes KV, Enjuanes L (2003) The SARS coronavirus: a postgenomic era. *Science* 300:1377–1378
- Hsue B, Masters PS (1999) Insertion of a new transcriptional unit into the genome of mouse hepatitis virus. *J Virol* 73:6128–6135
- Iverson LE, Rose JK (1981) Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. *Cell* 23:477–484
- Izeta A, Smerdou C, Alonso S, Penzes Z, Méndez A, Plana-Durán J, Enjuanes L (1999) Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. *J Virol* 73:1535–1545
- Jeong YS, Repass JF, Kim Y-N, Hwang S-M, Makino S (1996) Coronavirus transcription mediated by sequences flanking the transcription consensus sequence. *Virology* 217:311–322
- Joo M, Makino S (1992) Mutagenic analysis of the coronavirus intergenic consensus sequence. *J Virol* 66:6330–6337
- Joo M, Makino S (1995) The effect of two closely inserted transcription consensus sequences on coronavirus transcription. *J Virol* 69:272–280
- Krishnan R, Chang RY, Brian DA (1996) Tandem placement of a coronavirus promoter results in enhanced mRNA synthesis from the downstream-most initiation site. *Virology* 218:400–405
- Kubo H, Yamada YK, Taguchi F (1994) Localization of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein. *J Virol* 68:5403–5410

- Kuo L, Godeke G-J, Raamsman MJB, Masters PS, Rottier PJM (2000) Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: crossing the host cell species barrier. *J Virol* 74:1393–1406
- Lai MMC, Cavanagh D (1997) The molecular biology of coronaviruses. *Adv Virus Res* 48:1–100
- Lai MMC, Zhang X, Hinton D, Stohlman S (1997) Modulation of mouse hepatitis virus infection by defective-interfering RNA-mediated expression of viral proteins and cytokines. *J Neurovirol* 3: S33–S34
- La Monica N, Yokomori K, Lai MMC (1992) Coronavirus mRNA synthesis: identification of novel transcription initiation signals which are differentially regulated by different leader sequences. *Virology* 188:402–407
- Lassnig C, Sánchez CM, Enjuanes L, Muller M (2004) Obtention of transgenic mice susceptible to human coronavirus infection. Submitted for publication
- Lee HJ, Shieh CK, Gorbalenya AE, Koonin EV, Lamonica N, Tuler J, Bagdzhadzhyan A, Lai MMC (1991) The complete sequence (22 kilobases) of murine coronavirus gene-1 encoding the putative –327
- Lin YJ, Lai MMC (1993) Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontinuous sequence for replication. *J Virol* 67:6110–6118
- Luo Z, Matthews AM, Weiss SR (1999) Amino acid substitutions within the leucine zipper domain of the murine coronavirus spike protein cause defects in oligomerization and the ability to induce cell-to-cell fusion. *J Virol* 73:8152–8159
- Makino S, Joo M, Makino JK (1991) A system for study of coronavirus messenger RNA synthesis: a regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion. *J Virol* 65:6031–6041
- Makino S, Joo M (1993) Effect of intergenic consensus sequence flanking sequences on coronavirus transcription. *J Virol* 67:3304–3311
- Marra MA, Jones SJM, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YSN, Khattra J, Asano JK, Barber SA, Chan SY, Cloutier A, Coughlin SM, Freeman D, Girn N, Griffith OL, Leach SR, Mayo M, McDonald H, Montgomery SB, Pandoh PK, Petrescu AS, Robertson AG, Schein JE, Siddiqui A, Smailus DE, Stott JM, Yang GS, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth TF, Bowness D, Czub M, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples GA, Tyler S, Vogrig R, Ward D, Watson B, Brunham RC, Kraiden M, Petric M, Skowronski DM, Upton C, Roper RL (2003) The genome sequence of the SARS-associated coronavirus. *Science* 300:1399–1404
- Masters PS (1999) Reverse genetics of the largest RNA viruses. *Adv Virus Res* 53:245–264
- McGoldrick A, Lowings JP, Paton DJ (1999) Characterisation of a recent virulent transmissible gastroenteritis virus from Britain with a deleted ORF 3a. *Arch Virol* 144:763–770
- Méndez 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

- Meulenberg JJM, Bos-de Ruijter JNA, van de Graaf R, Wenswoort G, Moormann RJM (1998) Infectious transcripts from cloned genome-length cDNA of porcine reproductive and respiratory syndrome virus. *J Virol* 72:380–387
- Molenkamp R, van Tol H, Rozier BC, van der Meer Y, Spaan WJ, 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
- Nagy PD, Simon AE (1997) New insights into the mechanisms of RNA recombination. *Virology* 235:1–9
- Navas S, Seo SH, Chua MM, das Sarma J, Lavi E, Hingley ST, Weiss SR (2001) Murine coronavirus spike protein determines the ability of the virus to replicate in the liver and cause hepatitis. *J Virol* 75:2452–2457
- Nelsen CJ, Murtaugh MP, Faaberg KS (1999) Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. *J Virol* 73:270–280
- O'Connor BJ, Brian DA (1999) The major product of porcine transmissible gastroenteritis coronavirus gene 3b is an integral membrane glycoprotein of 31 kDa. *Virology* 256:152–161
- O'Connor JB, Brian DA (2000) Downstream ribosomal entry for translation of coronavirus TGEV gene 3b. *Virology* 269:172–182
- Olsen CW (1993) A review of feline infectious peritonitis virus: molecular biology, immunopathogenesis, clinical aspects, and vaccination. *Vet Microbiol* 36:1–37
- Ortego J, Escors D, Laude H, Enjuanes L (2002) Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome. *J Virol* 76:11518–11529
- Ortego J, Sola I, Almazan F, Ceriani JE, Riquelme C, Balasch M, Plana-Durán J, Enjuanes L (2003) Transmissible gastroenteritis coronavirus gene 7 is not essential but influences in vivo virus replication and virulence. *Virology* 308:13–22
- Ortego J, DeDiego ML, Enjuanes L (2004) Novel human vector based on coronavirus genomes. Submitted for publication
- Ozdarendeli A, Ku S, Rochat S, Senanayake SD, Brian DA (2001) Downstream sequences influence the choice between a naturally occurring noncanonical and closely positioned upstream canonical heptameric fusion motif during bovine coronavirus subgenomic mRNA synthesis. *J Virol* 75:7362–7374
- Pasternak AO, van den Born E, Spaan WJM, Snijder EJ (2001) Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J* 20:7220–7228
- Pasternak AO, van den Born E, Spaan WJM, Snijder EJ (2003) The stability of the duplex between sense and antisense transcription-regulating sequences is a crucial factor in arterivirus subgenomic mRNA synthesis. *J Virol* 77:1175–1183
- Pasternak AO, Spaan WJM, Snijder EJ (2004) Regulation of relative abundance of arterivirus subgenomic mRNAs. *J Virol*. In press
- Pensaert M, Callebaut P, Vergote J (1986) Isolation of a porcine respiratory, non-enteric coronavirus related to transmissible gastroenteritis. *Vet Quart* 8:257–260
- Pensaert MB, De Bouck P (1978) A new coronavirus-like particle associated with diarrhea in swine. *Arch Virol* 58:243–247

- 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
- Penzes Z, González JM, Calvo E, Izeta A, Smerdou C, Mendez A, Sánchez CM, Sola I, Almazán F, Enjuanes L (2001) Complete genome sequence of transmissible gastroenteritis coronavirus PUR46-MAD clone and evolution of the Purdue virus cluster. *Virus Genes* 23:105–118
- Phillips JJ, Chua MM, Lavi E, Weiss SR (1999) Pathogenesis of chimeric MHV4/MHV-A59 recombinant viruses: the murine coronavirus spike protein is a major determinant of neurovirulence. *J Virol* 73:7752–7760
- Resta S, Luby JP, Rosenfeld CD, Siegel JD (1985) Isolation and propagation of a human enteric coronavirus. *Science* 229:978–981
- Riquelme C, Ortego J, Izeta A, Plana-Durán J, Enjuanes L (2004) Engineering a recombinant canine coronavirus with reduced virulence using an infectious cDNA clone of transmissible gastroenteritis coronavirus. Submitted for publication
- Sánchez CM, Jiménez G, Laviada MD, Correa I, Suñé C, Bullido MJ, Gebauer F, Smerdou C, Callebaut P, Escribano JM, Enjuanes L (1990) Antigenic homology among coronaviruses related to transmissible gastroenteritis virus. *Virology* 174:410–417
- Sánchez CM, Gebauer F, Suñé C, Méndez A, Dopazo J, Enjuanes L (1992) Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. *Virology* 190:92–105
- Sánchez CM, Izeta A, Sánchez-Morgado JM, Alonso S, Sola I, Balasch M, Plana-Durán J, Enjuanes L (1999) Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. *J Virol* 73:7607–7618
- Sánchez CM, Sola I, Sánchez-Morgado JM, Enjuanes L (2004) The amino terminus of transmissible gastroenteritis coronavirus spike protein dictates the enteric tropism of the virus. In press
- 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* 82:386–396
- Shieh C-k, Soe LH, Makino S, Chang M-F, Stohlman SA, Lai MMC (1987) The 5'-end sequence of the murine coronavirus genome: implications for multiple fusion sites in leader-primed transcription. *Virology* 156:321–330
- Siddell SG (1995) *The Coronaviridae*. Plenum Press, New York
- Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LLM, Guan Y, Rozanov M, Spaan WJM, Gorbalenya AE (2003) Unique and conserved features of genome and proteome of SARS-coronavirus, and early split-off from the coronavirus group 2 lineage. *J Mol Biol* 331:991–1004
- Sola I, Alonso S, Zúñiga S, Balach M, Plana-Durán J, Enjuanes L (2003) Engineering transmissible gastroenteritis virus genome as an expression vector inducing lactogenic immunity. *J Virol* 77:4357–4369
- Sooknanan R, Howes M, Read L, Malek LT (1994) Fidelity of nucleic acid amplification with avian myeloblastosis virus reverse transcriptase and T7 RNA polymerase. *BioTechniques* 17:1077–1085

- Stirrupps K, Shaw K, Evans S, Dalton K, Casais R, Cavanagh D, Britton P (2000) Expression of reporter genes from the defective RNA CD-61 of the coronavirus infectious bronchitis virus. *J Gen Virol* 81:1687–1698
- Suñé C, Jiménez G, Correa I, Bullido MJ, Gebauer F, Smerdou C, Enjuanes L (1990) Mechanisms of transmissible gastroenteritis coronavirus neutralization. *Virology* 177:559–569
- Taguchi F, Kubo H, Takahashi H, Suzuki H (1995) Localization of neurovirulence determinant for rats on the S1 subunit of murine coronavirus JHMV. *Virology* 208:67–74
- Thiel V, Siddell SG, Herold J (1998) Replication and transcription of HCV 229E replicons. *Adv Exp Med Biol* 440:109–114
- Thiel V, Herold J, Schelle B, Siddell S (2001a) 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 (2001b) Viral replicase gene products suffice for coronavirus discontinuous transcription. *J Virol* 75:6676–6681
- Thiel V, Ivanov KA, Putics A, Hertzog T, Schelle B, Bayer S, Wessbrich B, Snijder EJ, Rabenau H, Doerr HW, Gorbalenya AE, Ziebuhr J (2003a) Mechanisms and enzymes involved in SARS coronavirus genome expression. *J Gen Virol* 84:2305–2315
- Thiel V, Karl N, Schelle B, Disterer P, Klagge I, Siddell SG (2003b) Multigene RNA vector based on coronavirus transcription. *J Virol* 77:9790–9798
- Thomas MJ, Platas AA, Hawley DK (1998) Transcriptional fidelity and proofreading by RNA polymerase II. *Cell* 93:627–637
- USDA (2002) Part II: reference of swine health and health management in the United States, 2000. National Animal Health Monitoring System
- van der Most RG, De Groot RJ, Spaan WJM (1994) Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus: a study of coronavirus transcription initiation. *J Virol* 68:3656–3666
- van Dinten LC, den Boon JA, Wassenaar ALM, Spaan WJM, Snijder EJ (1997) An infectious arterivirus cDNA clone: identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc Natl Acad Sci USA* 94:991–996
- van Marle G, Luytjes W, Van der Most RG, van der Straaten T, Spaan WJM (1995) Regulation of Coronavirus mRNA transcription. *J Virol* 69:7851–7856
- van Marle G, Dobbe JC, Gulyaev AP, Luytjes W, Spaan WJM, Snijder EJ (1999) Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc Natl Acad Sci USA* 96:12056–12061
- Vaughn EM, Paul PS (1993) Antigenic and biological diversity among transmissible gastroenteritis virus isolates of swine. *Vet Microbiol* 36:333–347
- Vaughn RM, Halbur PG, Paul PS (1995) Sequence comparison of porcine respiratory coronavirus isolates reveals heterogeneity in the S, 3, and 3-1 genes. *J Virol* 69:3176–3184
- Ward CD, Stokes MAM, Flanagan JB (1988) Direct measurement of the poliovirus RNA polymerase error frequency in vitro. *J Virol* 62:558–562

- Wentworth DE, Holmes KV (2001) Molecular determinants of species specificity in the coronavirus receptor aminopeptidase N (CD13): influence of N-linked glycosylation. *J Virol* 75:9741–9752
- Wentworth DE, Tresnan DB, Lerman I, Levis R, Shapiro LH, Holmes KV (2001) Subceptibility of transgenic mice expressing the receptor for human coronavirus-229E. In: ASV 20th Annual Meeting, University of Wisconsin-Madison, Madison, p 157
- Wertz GW, Perepelitsa VP, Ball LA (1998) Gene rearrangement attenuates expression and lethality of a nonsegmented negative strand RNA virus. *Proc Natl Acad Sci USA* 95:3501–3506
- Wesley RD, Cheung AK, Michael DM, Woods RD (1989) Nucleotide sequence of coronavirus TGEV genomic RNA: evidence of 3 mRNA species between the peplomer and matrix protein genes. *Virus Res* 13:87–100
- Wesley RD, Woods RD, Hill HT, Biwer JD (1990b) Evidence for a porcine respiratory coronavirus, antigenically similar to transmissible gastroenteritis virus, in the United States. *J Vet Diagn Invest* 2:312–317
- Wesley RD, Woods RD, Cheung AK (1991) Genetic analysis of porcine respiratory coronavirus, an attenuated variant of transmissible gastroenteritis virus. *J Virol* 65:3369–3373
- Yeager CL, Ashmun RA, Williams RK, Cardellichio CB, Shapiro LH, Look AT, Holmes KV (1992) Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* 357:420–422
- Yount B, Curtis KM, Baric RS (2000) Strategy for systematic assembly of large RNA and DNA genomes: the transmissible gastroenteritis virus model. *J Virol* 74:10600–10611
- 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
- Yount B, Curtis KM, Fritz EA, Hensley LE, Jahrling PB, Prentice E, Denison MR, Geisbert TW, Baric RS (2003) Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc Natl Acad Sci USA* 100:12995–13000
- Zhang X, Hinton DR, Cua DJ, Stohlman SA, Lai MMC (1997) Expression of interferon- γ by a coronavirus defective-interfering RNA vector and its effect on viral replication, spread, and pathogenicity. *Virology* 233:327–338
- Zhang X, Hinton DR, Park S, Parra B, Liao C-L, Lai MMC (1998) Expression of hemagglutinin/esterase by a mouse hepatitis virus coronavirus defective-interfering RNA alters viral pathogenesis. *Virology* 242:170–183
- Zhang X, Liu R (2000) Identification of a noncanonical signal for transcription of a novel subgenomic mRNA of mouse hepatitis virus: implication for the mechanism of coronavirus RNA transcription. *Virology* 278:75–85
- Zúñiga S, Sola I, Alonso S, Enjuanes L (2004) Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J Virol* 78:980–994