CHAPTER 1

General introduction

VIRUS DIVERSITY AND THE ORDER *NIDOVIRALES*

Per definition viruses are inanimate organic entities that are capable to replicate themselves. However, in contrast to living beings their DNA or RNA genomes do not encode information for the expression of proteins involved in the synthesis of the four fundamental biological building blocks: amino acids, carbohydrates, lipids, and nucleoside triphosphates (NTPs). Because of this limitation, they are obligate intracellular parasites that strictly depend on the metabolism of a host cell. Additionally, host proteins may play essential or supporting roles during specific steps in the viral replication cycle – the most obvious being cellular receptors used for viral entry. In effect, host factors thus determine the spectrum of genetically related cellular species and cell types a virus can infect – in short the virus' host range.

At the moment the International Committee on Taxonomy of Viruses (ICTV) recognizes about 3000 different virus species (1). On the other hand, it has been estimated that about 15 million different cellular species (~9 million eukaryotes, ~6 million prokaryotes) live on this planet (2;3). If we assume that each of those is host to at least one virus species – likely a vast underestimation given that humans are host to 189 known viruses (4) – a lot remains to be discovered. To bring order into the known and anticipated virus diversity in terms of, for example, genome type and organization, or replication strategy, related viruses have been grouped into genera, (sub-)families, and orders (proceeding from lower to higher rank). However, due to the extreme divergence of viruses and fast evolution, the relationship between different ranks remains often obscure. In this respect, virus taxonomy stands in stark contrast to the Tree of Life that has been constructed for organisms to reflect the course of cellular evolution.

The viruses that are discussed in this thesis belong to the order *Nidovirales*. This name derives from the typical genome expression strategy of its members featuring a nested set of subgenomic (sg) mRNAs (in Latin, *nidus* means nest). At the moment four families with different host ranges are united in the order: *Arteriviridae* (vertebrate hosts), *Coronaviridae* (vertebrate hosts), *Mesoniviridae* (invertebrate hosts), and *Roniviridae* (invertebrate hosts) (5-8). With the exception of the *Mesoniviridae*, all families contain economically important pathogens infecting livestock, for example swine (arterivirus porcine reproductive and respiratory syndrome virus, coronaviruses porcine epidemic diarrhea virus and transmissible gastroenteritis virus), cattle (bovine coronavirus), poultry (coronavirus infectious bronchitis virus, IBV), and prawn (ronivirus yellow head virus), and hence cause severe losses to the respective industries (9-14). Additionally, established human coronaviruses may cause mild respiratory symptoms. Combined these are the second leading cause for common cold after rhinoviruses (*Picornaviridae*) (15).

Recent years also saw the emergence of two previously unknown and highly pathogenic zoonotic coronaviruses in the human population: severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002 and Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012 (16;17). In contrast to the established human coronaviruses, which are constantly circulating in the human population, these viruses were initially directly transmitted from an animal reservoir to humans. In the case of SARS-CoV it is now believed that this reservoir may be one of the numerous bat species (18;19). From these animals the virus spread to humans and caused the first pandemic of the 21st century with major outbreaks in China and Southeast Asia but also Canada (20). Despite concerns that SARS-CoV might mutate to permanently establish itself within the human population, the virus disappeared – thanks to the imposed control measures like strict quarantine protocols – from circulation in humans in 2003 after causing about 8500 cases, including 812 deaths (21). The second newly-emerged coronavirus, MERS-CoV, which might be transmitted by camels (22), appears to be even more lethal with a case fatality rate of above 30%. However, thus far the case numbers have remained low, with about 1000 cases between April 2012 and November 2014 (23). Still, the threat to global public health and economy, exemplified by the SARS and MERS outbreaks, but also the combined economic damage caused by the veterinary nidoviruses call for a more thorough understanding of nidovirus biology. Ultimately these efforts might contribute to the development of countermeasures to keep future outbreaks in check.

THE NIDOVIRUS REPLICATION CYCLE

Nidoviruses enter a host cell by receptor-mediated endocytosis utilizing a variety of entry receptors (24). Afterwards the viral genome, which is a single RNA molecule of positive (mRNA) polarity carrying a type-1 cap structure (cap-1) ($^{m}GpppN_{m}$) and a polyadenylate (polyA) tail at its 5′ and 3′ end, respectively, is released into the cytoplasm. The genome is organized into multiple open reading frames (ORFs) (Figure 1), of which ORF1a and ORF1b encode all nonstructural proteins (nsps) separated by a ribosomal frameshift site, comprising a secondary structure element called RNA pseudoknot and a uridine-rich so-called "slippery sequence". It is estimated that in equine arteritis virus (EAV) about 15-20% and in the coronaviruses mouse hepatitis virus (MHV) and IBV up to 40% of the translating ribosomes perform the -1 frameshift and hence synthesize a large polyprotein called pp1ab (25-28). In the remainder of the cases the ribosome reaches a stop codon that is located just downstream of the frameshift signal. The resulting polyprotein is known as pp1a. Interestingly, all key enzymes for RNA synthesis and processing, for example the RNA-dependent RNA polymerase (RdRp), helicase, and – in the case of coronaviruses – also the proofreading exoribonuclease and capping enzymes,

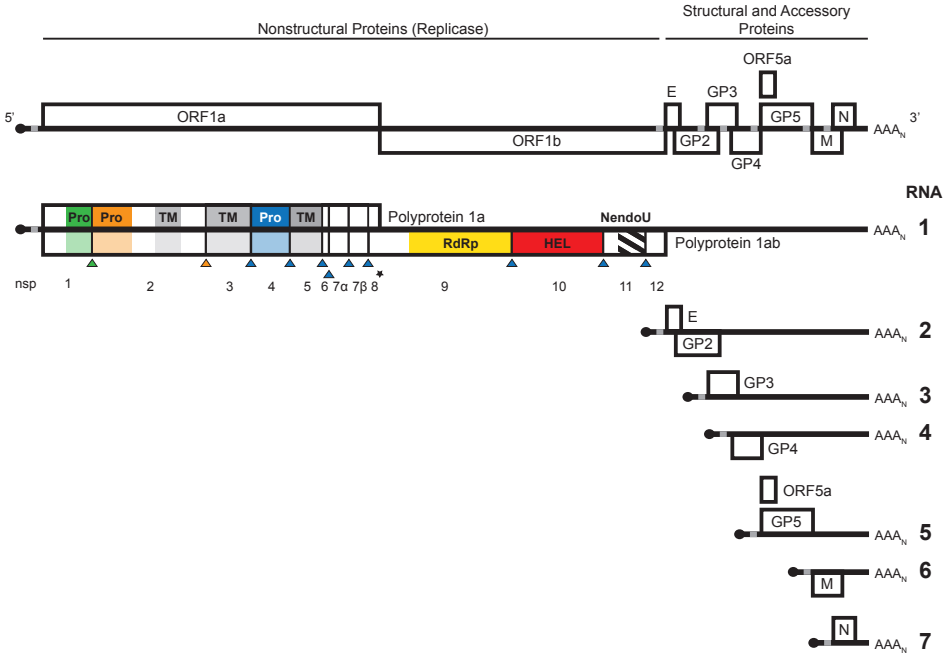

Figure 1: Typical nidovirus genome organization illustrated using equine arteritis virus (EAV). Open reading frames (ORFs) are indicated as boxes. Cleavage sites of replicase proteins in polyproteins 1a and 1ab are marked by triangles corresponding in color to the protease responsible for cleavage. Known transmembrane and enzymatic domains are indicated. Pro, protease; TM, transmembrane domain; RdRp, RNA-dependent RNA polymerase; HEL, helicase; NendoU, endoribonuclease; E, envelope protein; GP, glycoprotein; M, membrane protein; N, nucleocapsid protein. The ribosomal frameshift site leading to expression of polyprotein 1ab is labeled with a star. Transcription-regulating sequences are indicated as gray boxes. Presumed 5′ cap structures are depicted as black dots.

are encoded downstream of the frameshift while known co-factors of these enzymes, RNA-binding and membrane-anchoring proteins, as well as proteolytic enzymes are so far exclusively mapped to pp1a (29-32). The frameshift is thus an elegant way to regulate the relative abundance of these key enzymes compared to other proteins controlling genome replication and expression. Surprisingly, a second frameshift site was recently discovered in all arteriviruses except EAV. This site, located further upstream roughly in the middle of ORF1a, is able to direct a -2 as well as a -1 frameshift and thus gives rise to two additional variants of the membrane-bound nonstructural protein nsp2, one of which being predicted to be soluble. This arterivirus frameshift site with dual shift capacity, which is controlled by *trans*-activation by the upstream nsp1β subunit, is the only one of its kind known to date (33;34).

In addition to the nsps that are directly translated from the genome, group-specific structural proteins and – in the case of coronaviruses – accessory proteins are trans-

lated from an in part extensive set of sg mRNAs (35). These mRNAs, which carry the same 5'- and 3'-terminal sequences as the genome in most nidoviruses, are transcribed with the help a unique mechanism involving subgenome-size negative-stranded (-) templates that in part arise from discontinuous RNA synthesis (see below). As for all positive-stranded (+) RNA viruses, RNA replication (amplification of the genome) and transcription (synthesis of sg mRNAs) are thought to take place in association with an extensive network of modified membranes (36;37). For nidoviruses this membranous web takes mainly the form of interconnected double-membrane vesicles (DMVs) and convoluted membranes (CMs). It was speculated that these membrane structures may provide a scaffold for replication-transcription complex (RTC) assembly inside DMVs. Hence, it is thought that these structures support viral replication in two ways; on the one hand, by increasing local concentrations of NTPs, RNAs, and proteins required for RNA synthesis and, on the other, by shielding viral replication products, especially double-stranded replication intermediates, from detection by the host's innate immune system. After encapsidation of the viral genome, particles bud into the lumen of the smooth endoplasmic reticulum (ER) or Golgi complex. From there they are transported via the cellular secretory pathway to be released from the plasma membrane (24).

MOLECULAR DETAILS OF NIDOVIRUS REPLICATION AND TRANSCRIPTION

Prison break: how nidoviruses with large genomes overcame size constraints

The genome sizes of nidoviruses range from 12-16 kilobases (kb) for arteriviruses (from here on referred to as "small nidoviruses") to 20-34 kb for mesoni-, roni-, and coronaviruses ("intermediate and large nidoviruses"). With these sizes especially the latter group deviates substantially from the average size of most (+) RNA virus genomes that typically are smaller than 10 kb (Figure 2) (38). Still, even the largest RNA virus currently known, the recently discovered ball python nidovirus – a proposed member of the *Torovirinae,* a subfamily of the *Coronaviridae* – with a genome size of 33.5 kb (39), is dwarfed by some DNA viruses, whose genomes can reach sizes in the range of megabase pairs (Mbp), for example mimiviruses (~1 Mbp) and pandoraviruses (~2.5 Mbp) (40;41). Considering these tremendous size differences, two questions arise: in what way are RNA viruses so fundamentally different from DNA viruses that a genome expansion of the scale of the latter did not occur, and how did large nidoviruses, at least to some extent, overcome the size restrictions imposed on other (+) RNA viruses?

To answer these questions, it should be informative to explore the underlying reason for the existence of the observed size barrier in (+) RNA viruses. All of these viruses encode

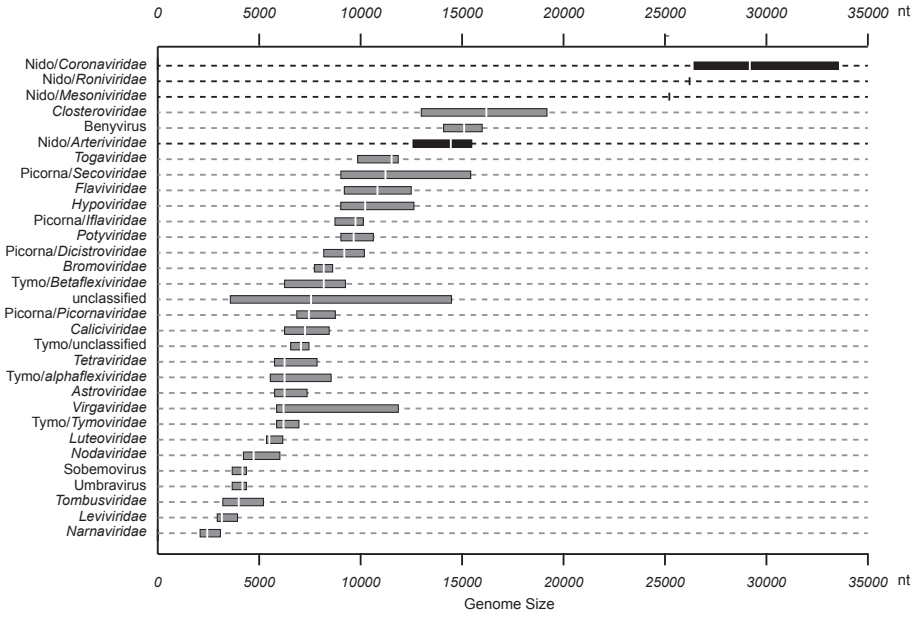

Figure 2: Genome sizes of positive-stranded RNA viruses. Size ranges of major families or genera and unclassified viruses are indicated by black (nidoviruses) or gray boxes. The median size is marked by a white vertical bar. Nido, *Nidovirales*; Picorna, *Picornavirales*; Tymo, *Tymovirales*. Adapted from (42).

an RdRp that synthesizes copies of the viral genome during the infection of a host cell. The basic mechanism by which these enzymes fulfill this central function in the viral replication cycle can be simplified to two steps: the matching of an incoming NTP to the template and the formation of a chemical bond to extend the nascent RNA chain (43;44). The first step of this mechanism basically occurs by a trial-and-error method as polymerases lack the means to determine the identity of the nucleotide that is about to be copied or of the NTP that has entered the active site. Instead the selection and, ultimately, the incorporation of an NTP is solely based on the relative difference between its dissociation rate from the active site and the rate of phosphodiester bond formation. Because a correct Watson-Crick base pair is energetically more stable than a mismatched one or any of the alternative base pairs, the correct NTP will, on average, remain at the active site for a longer period of time than an incorrect one. If this period is long enough for bond formation to occur, the RNA chain will be extended by this one nucleotide. If not, the NTP will diffuse away, and the next NTP can be tried at random. In summary, in order to minimize the number of errors but maintain RNA synthesis, the chemical reaction rate should be much lower than the dissociation rate of incorrect NTPs but higher than the dissociation rate of correct NTPs.

In general, error rates of RdRps were estimated to range between 10^{-3} and 10^{-5} errors per nucleotide incorporated (45). That means most individual genomes of an average (+) RNA virus would differ by at least one nucleotide from each other. To emphasize this variation, the concept of a quasispecies was introduced, essentially representing a cloud of different variants of a consensus sequence that are heterogeneous in respect to their fitness (46;47). Depending on the environmental conditions, the composition of these quasispecies may differ. Interestingly, decreasing the variation within a quasispecies by increasing the replication fidelity of an RdRp was shown to strongly diminish the overall fitness of a virus population (48-52). It is therefore believed that on an evolutionary scale it is the quasispecies, rather than individual variants, that is targeted by selection. On the other hand, decreasing replication fidelity will lead to the accumulation of too many detrimental mutations, which will eventually prevent virus replication, a consequence that was termed "error catastrophe". Because of these two opposing principles, RNA viruses are thought to be optimized to exist close to the threshold of this error catastrophe. In summary, this implies that the size of the genome is limited by the error rate of the RdRp it encodes. Interestingly, there seems to be a correlation between the size of RNA genomes and their RdRp genes (45). Whether these larger RdRps indeed operate with a lower error rate, however, remains to be seen.

To express the interdependence between replication fidelity, genome size, and genome complexity, the term "Eigen trap" has been coined (53). This term essentially conveys the fact that none of these three parameters can be increased without simultaneously increasing the other two. When comparing (+) RNA virus genomes, two instances where both genome complexity and size expanded by the introduction of a new enzyme have been recognized. First, an RNA helicase is encoded by all viruses with genomes larger than 7 kb (54). It was proposed that this enzyme may support the RdRp by removing double-stranded regions from the template. However, how this would directly affect fidelity, which depends, as explained above, on an interplay between NTP affinities and the chemical reaction rate, is unclear. Therefore, it cannot be excluded that there may be other reasons underlying the presence of a helicase in (+) RNA virus genomes.

A more straightforward explanation was proposed for a second instance of genome expansion. In this case a 3'-5' exoribonuclease was acquired, which gave rise to (+) RNA genomes of more than 20 kb (55). A similar enzymatic activity, a 3'-5' exonuclease, is a vital part of typical DNA polymerases encoded by DNA viruses and cellular organisms. Contrary to general believe, the intrinsic error rates of DNA polymerases are, as a matter of fact, not significantly lower (10^{-4}-10^{-5} bp^{-1}) than that of RNA polymerases. It is the presence of this associated exonuclease activity that enables the reduction of the error rate to 10^{-5}-10^{-7} bp^{-1} by removing incorrectly incorporated NTPs during DNA synthesis (45).

By analogy, it was assumed and recently experimentally confirmed that the nidovirus exoribonuclease confers proofreading activity to the viral RTC (32;56). Furthermore, in reverse genetics experiments it was demonstrated that a knock-out of this proofreading activity led to a more than 10-fold increase of the overall error rate during MHV and SARS-CoV replication in cell culture (57;58). In contrast to large nidoviruses, arteriviruses do not encode an exoribonuclease subunit (55). Still, their genomes are substantially larger than those of most other (+) RNA viruses. It thus remains to be seen if another domain acquisition event may be linked to this expansion.

Nidovirus discontinuous RNA synthesis

As already hinted at, the transcription mechanism of nidoviruses is unique in the virus world (Figure 3). Although several non-nidovirus families utilize sg mRNAs, none generates those by a mechanism equivalent to that of most nidoviruses, involving discontinuous (-) subgenome-length RNA synthesis (roni- and toroviruses do not or only in part employ this mechanism) (59). In contrast to, for example, alphavirus sg mRNA synthesis that is driven from an internal promoter in the full-length negative strand, nidovirus sg mRNAs are transcribed from several co-terminal (-) subgenome-length RNAs of different lengths. Essential protagonists in the still not well understood mechanism to produce those templates are so-called transcription-regulating sequences (TRSs), which are conserved AU-rich elements of a length – depending on the virus – of 5-18 nucleotides located near the genome's 5′ end (leader TRS) and upstream of most of the 3′ ORFs (body TRSs). During negative-strand synthesis, which always initiates at the genome's 3′ end, the viral RdRp may pause at one of these sequences. Subsequently, the part of the template between the body and leader TRS is skipped before RNA synthesis resumes at the genome's 5′ end at the so-called leader sequence. How exactly this skipping occurs is still not understood, but a dissociative step during body-leader joining may be

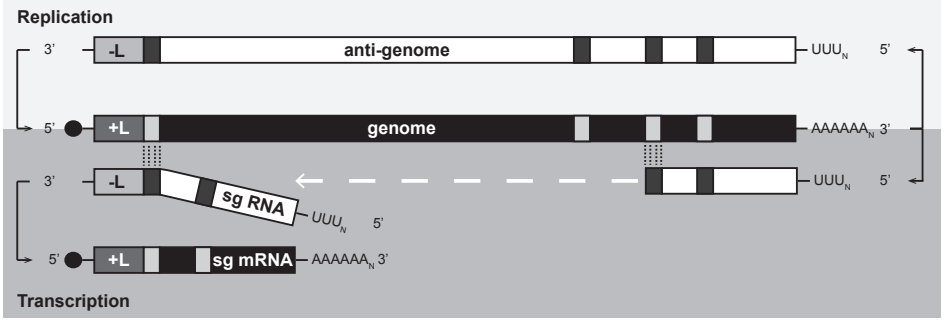

Figure 3: Discontinuous negative-strand transcription model. Transcription-regulating sequences are indicated by gray (positive strands) or black (negative strands) boxes. Leader and anti-leader regions are labeled +L and –L, respectively. Presumed 5′ cap structures of mRNAs are depicted as black dots. Adapted from (24).

involved since co-infection experiments with two different MHV strains showed that the leader sequence may derive from a different template than the rest of the transcribed product (60). Also base-pairing between the leader and anti-body TRS is important since the amount of each sg mRNA correlates with the calculated stability of the respective TRS duplex (61-64). However, it was speculated that base-pairing may not be the only factor involved. For instance, it was also shown that TRSs serve a function independent of base-pairing, potentially in secondary structure-dependent recruitment of specific proteins (59). Three viral proteins that were implicated in arterivirus transcription regulation are nsp1 (65;66), the nidovirus-wide conserved helicase nsp10 (67;68), and the endoribonuclease nsp11 (69). For these proteins, mutations either altered the balance between genome replication and transcription or selectively abolished sg mRNA synthesis altogether. Since genome replication requires the synthesis of full-length negative strands, and hence a read-through through all TRSs, these results indicated that the three proteins are directly or indirectly involved in the discontinuous step. Interestingly, nsp1 as well as nsp10 contain a zinc-binding domain that may be instrumental in establishing interactions with the RNA or proteins of the RTC to serve this regulatory function.

As a consequence of discontinuous RNA synthesis, all nidovirus mRNAs, except the smallest, are structurally polycistronic. However, with a few exceptions, only the most 5'-located ORF is actually translated, meaning that the mRNAs are functionally monocistronic (70). Thus, the question arises what the advantage of this complicated transcription mechanism is compared to structurally monocistronic mRNAs expressed from multiple promoters or polycistronic mRNAs enabling internal ribosome entry or other non-canonical translation initiation mechanisms frequently employed by other viruses. Obviously, the nidovirus mechanism ensures that all RNAs carry the same 5'- and 3'-terminal sequences as the genome or anti-genome. This could be advantageous if regulatory elements are located at the ends. For example, sequences of negative strands may be required to initiate positive-strand synthesis or capping. It would also be possible that the genome ends contain translational enhancers (59). Finally, those elements could also serve to discriminate viral from host RNAs. In view of the notion that all vertebrate nidoviruses encode an endoribonuclease (55), whose substrate is still elusive but may well be a host RNA, this possibility is especially intriguing.

Means to an end: nidovirus mRNA modification

Given the complexity of ribosomes, comprising 80 proteins and 4 rRNAs in higher eukaryotes (71), RNA viruses cannot encode information for their components. To ensure that viral mRNAs are translated in the host cell, a variety of strategies is employed by different virus families. The most obvious is to adopt the essential modifications of cel-

lular mRNAs, a 5′ cap structure and a 3′ polyA tail. Alternatively, viral mRNAs may contain special secondary structures, called internal ribosome entry sites (IRESs) or 3′ cap-independent translation enhancers (3′ CITEs), that allow the non-canonical recruitment of the translational apparatus. Finally, some viruses encode proteins that may replace certain cellular translation initiation factors. Their mRNAs may thus lack some of the modifications of host mRNAs, for instance the cap-1 or polyA tail (72-74).

Where characterized, the 3′ ends of positive-stranded RNAs of a number of nidoviruses contained polyA tails (53;75-79). Furthermore, a cap-1 ($^{m}GpppN_{m}$) structure was found to be present at the 5′ end of the genome and/or sg mRNAs of equine torovirus (80), the coronavirus MHV (81;82), and the arterivirus simian hemorrhagic fever virus (83). Based on common ancestry, it is thus assumed that all nidoviruses equip their mRNAs with these modifications, which would allow them to enter the cellular translation pathway. In line with this hypothesis, it was shown that MHV infection leads to phosphorylation of the cap-binding translation initiation factor eIF4E, which is required for the cellular pathway (84). This phosphorylation, which is a known regulatory mechanism in eukaryotic cells to strengthen the interaction between the cap and the protein, in turn increased the translation efficiency of viral mRNAs. Furthermore, overexpression of an inhibitor of eIF4E, 4E-BP, abolished replication of human coronavirus 229E in HeLa cells (85). Nevertheless, contradicting evidence with regard to the nature of its 5′ end has been brought forward for the arterivirus lactate dehydrogenase-elevating virus, whose genome appeared to be devoid of a cap and instead was monophosphorylated (86). Given this result and the fact that none of the members of the *Roniviridae* and *Mesoniviridae* was characterized so far, care should be taken in assuming that all nidovirus mRNAs carry the same 5′ end modification. Such a deviation in respect to translation strategy was also observed in the *Flaviviridae*, whose members may utilize cap-dependent or -independent mechanisms (87;88). In addition to cap-dependent translation initiation, IRES elements may, at least in coronaviruses, drive expression of a second gene product from a single sg mRNA (89-91).

In the host cell cap-1 and polyA tails are strictly generated in the nucleus during and shortly after RNA polymerase II-dependent transcription (92). Since nidoviruses and many other (+) RNA viruses replicate in the cytoplasm, they cannot benefit from this cellular machinery. Instead the polyA tail may be synthesized by (one of) the viral RdRp(s) – coronaviruses are believed to encode a main RdRp (nsp12) and an accessory RdRp (nsp8) (93-95). How exactly this is achieved was not investigated so far. However, since negative-stranded RNAs were shown to contain a short polyU stretch at their 5′ end, it was speculated that iterative copying of this stretch may be involved (96). Alterna-

tively, SARS-CoV nsp8 in complex with nsp7 was shown to possess terminal transferase, that is, non-templated extension activity, on single-stranded RNAs (95).

In contrast to polyA-tail addition, the assembly of the cap-1 structure appears to be better understood in regard to the proteins involved, at least in large nidoviruses. In general the synthesis of the cap involves four steps and three different enzymatic activities that may be present in a single subunit with multiple domains or in multiple individual proteins (97;98). In case of the conventional capping pathway, which is employed by all eukaryotes and a number of viruses, the triphosphate end of a newly synthesized RNA is trimmed back to a diphosphate by an RNA-triphosphatase (RTPase). As this activity is mechanistically identical to the cleavage of NTPs, the NTPase domains of a viral helicase, if encoded, may execute it. In the second step, a guanylyltransferase (GTase) transfers a GMP-moiety to the RNA diphosphate end. In contrast to nucleotide bonds established by polymerases, this bond is formed via a 5′-5′ linkage to generate a GpppN-RNA structure. While this unusual bond cannot be cleaved by regular exo- and endoribonucleases, specially regulated cytoplasmic host decapping enzymes are employed for the removal of cap structures (99). As a consequence, capping confers protection against 5′-3′ exoribonucleases, and hence capped RNAs exhibit much longer half-lives than uncapped ones. In order to make the second step irreversible, a methyl group is attached to the N7-position of the guanine by an N7-methyltransferase (N-MT). Although this so-called cap-0 structure is due to the specific recognition of the methyl group by eIF4E the basic requirement for translation initiation (100), a second methylation usually occurs at the 2′ oxygen of the ribose of the first (cap-1) or second (cap-2) nucleotide following the cap. This second methylation step, which is catalyzed by a 2′-O-methyltransferase (O-MT) that may or may not be different from the domain utilized for N7 methylation, is connected to host mRNA surveillance mechanisms for self versus non-self discrimination (101;102). Next to this conventional pathway, alternative viral mechanisms have evolved that include a different order of steps leading to the same mature cap structure (98).

It has been proposed that nidoviruses employ the canonical pathway of cap synthesis described above (98). However, this hypothesis is far from proven especially with respect to the universal conservation of this pathway in all nidoviruses. For instance, the GTase has not been identified in any of the nidoviruses, while RTPase activity was demonstrated for only two coronavirus helicases (nsp13) (103;104). Whether or not this enzyme, which belongs to the most conserved proteins of the order, actually exerts this activity in the context of capping, however, remains to be verified. Finally, two methyltransferases (MTases) residing in nsp14 (N-MT) and nsp16 (O-MT) have been experimentally identified in coronaviruses (105-108). Interestingly, while other large nidoviruses – with the exception of toroviruses, which seem to lack the N-MT activity – encode homologs of

both MTases (55), neither of them was identified in arteriviruses. Since arteriviruses encode a unique protein (nsp12) at a genome position equivalent to that of coronavirus nsp16, the capping mechanism could be another example of biochemical variability within the diverse *Nidovirales* order.

OUTLINE OF THIS THESIS

The work described in this thesis addresses several poorly or uncharacterized (domains of) nsps that are likely involved in one or multiple steps during RNA replication and/or transcription of the prototypic arterivirus EAV. After the above short introduction on the nidovirus replication cycle and known molecular details of the unusual transcription and mRNA processing mechanisms, **chapter 2** presents the crystal structure of the enzymatically active EAV helicase nsp10, which was obtained and analyzed in close collaboration with Chinese colleagues. Interestingly, a strong resemblance between this viral protein and the conserved cellular helicase Upf1, in particular with respect to their N-terminal zinc-binding domains, became obvious. Since this cellular helicase is implicated in a number of eukaryotic post-transcriptional quality control mechanisms, a role for nsp10 and its nidovirus homologs in genome expansion is proposed. This and other potential functions of the nidovirus helicase in RNA replication, transcription, and translation, as well as virion biogenesis are further discussed in **chapter 3**, which presents a review of our current knowledge about nidovirus helicases. Special emphasis is placed on gaps that still remain, facts that cannot be easily reconciled with our current understanding of the nidovirus replication mechanisms, and questions that need to be addressed in future.

Chapters 4 and **5** focus on one of the central arterivirus replication proteins, nsp9, which harbors the RdRp domain. **Chapter 4** describes a carefully controlled study to investigate different polymerase activities that nsp9 may have, including a previously claimed primer-independent RdRp activity. Despite considerable efforts, involving experiments with different preparations of nsp9 and assays performed in the presence of putative polymerase co-factors, no *in vitro* activity was observed that could be clearly attributed to this protein. Moreover, circumstantial evidence suggested that the previously reported activity may have been caused by a contamination of the recombinant nsp9 preparation with the T7 RNA polymerase used to drive its expression in *E. coli*. In arteriviruses, the RdRp domain is located in the C-terminal two-thirds of nsp9. In **chapter 5**, it is now described for the first time that the RdRp domain is flanked at its N-terminus by another domain that is conserved in all nidoviruses. However, unlike the situation for the RdRp domain, no homologs of this domain have been found in other

RNA viruses. This domain is thus proposed to be a second marker for the *Nidovirales* order, besides the N-terminal zinc-binding domain of the helicase subunit. Residues that are part of three conserved sequence motifs were without exception associated with a newly discovered nucleotidylation activity of recombinant nsp9. It is thus proposed that this activity could play a role in the modification of the 5′ end of viral RNAs through either RNA ligation, protein priming of RNA synthesis, or guanylyl transfer during RNA capping. Further research is required to definitely tie nsp9 to one of these pathways. Nevertheless, alanine substitution of any of the conserved residues was either lethal to EAV and SARS-CoV or severely crippled these viruses, eventually resulting in reversion of the mutation. These results thus demonstrate the essential nature of this domain for virus replication, whatever its exact function will turn out to be.

Two MTase activities, commonly required for capping of mRNAs, were previously identified in two ORF1b-encoded coronavirus proteins, nsp14 and nsp16. While the former has no counterpart among the arterivirus nsps, the latter and the arterivirus C-terminal subunit nsp12 occupy equivalent positions in the ORF1b-encoded part of the replicase although the two proteins share no detectable sequence similarity. It is thus a long standing question, how arteriviruses may catalyze the 5′ end modification of mRNAs, and we therefore performed a first characterization of the entirely uncharacterized EAV nsp12 subunit (**chapter 6**). Based on the genomic position of its coding sequence, sequence alignment, and secondary structure prediction it is hypothesized that nsp12 might represent a unique arterivirus MTase, which has diverged from its homologs beyond sharing appreciated similarity. To test this hypothesis, recombinant nsp12 was expressed in and purified from *E. coli* and tested alone and in combination with potential co-factors for N-MT and O-MT activity. Although positive controls represented by the SARS-CoV MTases (nsp14 and the nsp10:nsp16 complex) demonstrated the functionality of the assay, no activity was detected for EAV nsp12. Guided by the sequence alignment, an extensive set of EAV mutants was generated and characterized with respect to their plaque phenotype and progeny titer, as well as their protein expression. These reverse genetics experiments revealed a number of phenotypes ranging from wild-type-like via non-spreading to replication-incompetent, which indicated that nsp12 is essential for viral replication.

The above chapters describing biochemical properties of selected proteins may ultimately contribute to the identification of drug targets to combat nidovirus infections. In **chapter 7** the prerequisites under which the marketing of such an antiviral drug would be economically viable are analyzed. This project was realized under guidance of several specialists of one of the industrial partners, Janssen Infectious Diseases, of the EUVIRNA consortium, the Marie Curie Initial Training Network to which my research project be-

longed. This study concludes that, at the moment, none of the circulating nidoviruses constitutes a sufficiently sized market to warrant the considerable investments required for drug development. The situation may be different if a new highly-pathogenic virus would emerge, as exemplified in 2002 by SARS-CoV or 2012 by MERS-CoV. In view of such threats, pre-pandemic drug stockpiling could be considered. However, also under those circumstances, it seems likely that the inherent financial risk would preclude an independent private initiative, even though market parameters and approval procedures appear to be favorable.

Finally, **chapter 8** connects some of the main findings described in this thesis with previously described data. In particular, potential differences between small and large nidoviruses on the level of the molecular mechanisms of RNA synthesis initiation and mRNA capping are highlighted. To this end, alternative mechanisms are considered that would be consistent with the data on arteriviruses presented in this thesis and elsewhere. Furthermore, potential roles of cellular helicases in nidovirus replication and the host's immune response against nidoviruses are discussed.

REFERENCE LIST

1. Lefkowitz EJ. ICTV/MSL: International Committee on Taxonomy of Viruses / Master Species List updated. 2014.

2. Curtis TP, Sloan WT, Scannell JW. Estimating prokaryotic diversity and its limits. Proc.Natl.Acad. Sci.U.S.A 2002;99(16):10494-10499.

3. Mora C, Tittensor DP, Adl S, *et al.* How many species are there on Earth and in the ocean? PLoS.Biol. 2011;9(8):e1001127.

4. Woolhouse M, Gaunt E. Ecological origins of novel human pathogens. Crit Rev.Microbiol. 2007; 33(4):231-242.

5. Lauber C, Ziebuhr J, Junglen S, *et al. Mesoniviridae*: a proposed new family in the order *Nidovirales* formed by a single species of mosquito-borne viruses. Arch.Virol. 2012;157(8):1623-1628.

6. Cowley JA, Walker PJ, Flegel TW, *et al.* Family *Roniviridae*. In King AMQ, Adams MJ, Carstens EB *et al.* editors, Virus taxonomy. Ninth report of the international committee on taxonomy of viruses, Amsterdam, Elsevier Academic Press, 2012;829-834.

7. de Groot RJ, Baker SC, Baric R, *et al.* Family *Coronaviridae*. In King AMQ, Adams MJ, Carstens EB *et al.* editors, Virus taxonomy. Ninth report of the international committee on taxonomy of viruses, Amsterdam, Elsevier Academic Press, 2012;806-828.

8. Faaberg KS, Balasuriya UB, Brinton MA, *et al.* Family *Arteriviridae*. In King AMQ, Adams MJ, Carstens EB *et al.* editors, Virus Taxonomy. Ninth report of the international committee on taxonomy of viruses, Amsterdam, Elsevier Academic Press, 2012;796-805.

9. Sang Y, Rowland RR, Blecha F. Antiviral regulation in porcine monocytic cells at different activation States. J.Virol. 2014;88(19):11395-11410.

10. Cowley JA, Dimmock CM, Spann KM, *et al.* Gill-associated virus of Penaeus monodon prawns: an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. J.Gen.Virol. 2000;81(Pt 6):1473-1484.

11. Song D, Park B. Porcine epidemic diarrhoea virus: a comprehensive review of molecular epidemiology, diagnosis, and vaccines. Virus Genes 2012;44(2):167-175.

12. Zhang Q, Hu R, Tang X, *et al.* Occurrence and investigation of enteric viral infections in pigs with diarrhea in China. Arch.Virol. 2013;158(8):1631-1636.

13. Cavanagh D. Coronaviruses in poultry and other birds. Avian Pathol. 2005;34(6):439-448.

14. Saif LJ. Bovine respiratory coronavirus. Vet.Clin.North Am.Food Anim Pract. 2010;26(2):349-364.

15. Lai MM, Perlman S, Anderson LJ. *Coronaviridae*. In Mahan Knipe D, Howley PM, editors, Fields' Virology Volume One, Philadelphia, Lippincott, Williams & Wilkins, 2007;1305-1397.

16. Coleman CM, Frieman MB. Coronaviruses: important emerging human pathogens. J.Virol. 2014; 88(10):5209-5212.

17. Hilgenfeld R, Peiris M. From SARS to MERS: 10 years of research on highly pathogenic human coronaviruses. Antiviral Res. 2013;100(1):286-295.

18. Shi Z, Hu Z. A review of studies on animal reservoirs of the SARS coronavirus. Virus Res. 2008; 133(1):74-87.

19. Ge XY, Li JL, Yang XL, *et al.* Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature 2013;503(7477):535-538.

20. Cherry JD, Krogstad P. SARS: the first pandemic of the 21st century. Pediatr.Res. 2004;56(1):1-5.

21. World Health Organization. SARS outbreak contained worldwide. [press release], 5 July 2003. Available at: http://www.who.int/mediacentre/news/releases/2003/pr56/en [accessed 10 November 2014]

22. Drosten C, Kellam P, Memish ZA. Evidence for camel-to-human transmission of MERS coronavirus. N.Engl.J.Med. 2014;371(14):1359-1360.

23. World Health Organization. Middle East respiratory syndrome coronavirus (MERS-CoV): Summary of current situation, literature update, and risk assessment - as of 5 February 2015. [online] Available at: http://www.who.int/csr/disease/coronavirus_infections/mers-5-february-2015.pdf [accessed 17 April 2014]

24. Snijder EJ, Siddell SG, Gorbalenya AE. The order *Nidovirales*. In Mahy BW, ter Meulen V, editors, Topley and Wilson's Microbiology and Microbial Infections: Virology Volume, London, Hodder Arnold, 2005;390-404.

25. Den Boon JA, Snijder EJ, Chirnside ED, *et al.* Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. J.Virol. 1991;65(6):2910-2920.

26. Firth AE, Brierley I. Non-canonical translation in RNA viruses. J.Gen.Virol. 2012;93(Pt 7):1385-1409.

27. Brierley I, Boursnell ME, Binns MM, *et al.* An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J. 1987;6(12):3779-3785.

28. Bredenbeek PJ, Pachuk CJ, Noten AF, *et al.* 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. 1990;18(7):1825-1832.

29. Sutton G, Fry E, Carter L, *et al.* The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. Structure. 2004;12(2):341-353.

30. Bouvet M, Lugari A, Posthuma CC, *et al.* Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. J.Biol.Chem. 2014;289(37):25783-25796.

31. Egloff MP, Ferron F, Campanacci V, *et al.* The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc.Natl.Acad.Sci.U.S.A 2004;101(11):3792-3796.

32. Subissi L, Posthuma CC, Collet A, *et al.* One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc.Natl.Acad. Sci.U.S.A 2014;111(37):e3900

33. Li Y, Treffers EE, Napthine S, *et al.* Transactivation of programmed ribosomal frameshifting by a viral protein. Proc.Natl.Acad.Sci.U.S.A 2014;111(21):e2172-2181.

34. Fang Y, Treffers EE, Li Y, *et al.* Efficient -2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein. Proc.Natl.Acad.Sci.U.S.A 2012;109(43):e2920-2928.

35. Sawicki SG, Sawicki DL, Siddell SG. A contemporary view of coronavirus transcription. J.Virol. 2007;81(1):20-29.

36. Hagemeijer MC, Rottier PJ, de Haan CA. Biogenesis and dynamics of the coronavirus replicative structures. Viruses. 2012;4(11):3245-3269.

37. Paul D, Bartenschlager R. Architecture and biogenesis of plus-strand RNA virus replication factories. World J.Virol. 2013;2(2):32-48.

38. Gorbalenya AE, Enjuanes L, Ziebuhr J, *et al. Nidovirales*: evolving the largest RNA virus genome. Virus Res. 2006;117(1):17-37.

39. Stenglein MD, Jacobson ER, Wozniak EJ, *et al.* Ball python nidovirus: a candidate etiologic agent for severe respiratory disease in Python regius. MBio. 2014;5(5):e01484-14.

40. Yamada T. Giant viruses in the environment: their origins and evolution. Curr.Opin.Virol. 2011;1(1): 58-62.

41. Philippe N, Legendre M, Doutre G, *et al.* Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. Science 2013;341(6143):281-286.

42. Lauber C. On the evolution of genetic diversity in RNA virus species - Uncovering barriers to genetic divergence and gene length in picorna- and nidoviruses. 2012. [dissertation]

43. Arnold JJ, Cameron CE. Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg^{2+}. Biochemistry 2004;43(18): 5126-5137.

44. Ng KK, Arnold JJ, Cameron CE. Structure-function relationships among RNA-dependent RNA polymerases. Curr.Top.Microbiol.Immunol. 2008;320:137-156.

45. Belshaw R, Gardner A, Rambaut A, *et al.* Pacing a small cage: mutation and RNA viruses. Trends Ecol.Evol. 2008;23(4):188-193.

46. Sallie R. Replicative homeostasis II: influence of polymerase fidelity on RNA virus quasispecies biology: implications for immune recognition, viral autoimmunity and other "virus receptor" diseases. Virol.J. 2005;2:70.

47. Vignuzzi M, Stone JK, Arnold JJ, *et al.* Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 2006;439(7074):344-348.

48. Coffey LL, Beeharry Y, Borderia AV, *et al.* Arbovirus high fidelity variant loses fitness in mosquitoes and mice. Proc.Natl.Acad.Sci.U.S.A 2011;108(38):16038-16043.

49. Graham RL, Becker MM, Eckerle LD, *et al.* A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. Nat.Med. 2012;18(12):1820-1826.

50. Pfeiffer JK, Kirkegaard K. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. PLoS.Pathog. 2005;1(2):e11.

51. Vignuzzi M, Wendt E, Andino R. Engineering attenuated virus vaccines by controlling replication fidelity. Nat.Med. 2008;14(2):154-161.

52. Weeks SA, Lee CA, Zhao Y, *et al.* A Polymerase mechanism-based strategy for viral attenuation and vaccine development. J.Biol.Chem. 2012;287(38):31618-31622.

53. Nga PT, Parquet MC, Lauber C, *et al.* Discovery of the first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. PLoS.Pathog. 2011;7(9):e1002215.

54. Gorbalenya AE, Koonin EV. Viral proteins containing the purine NTP-binding sequence pattern. Nucleic Acids Res. 1989;17(21):8413-8440.

55. Lauber C, Goeman JJ, Parquet MC, *et al.* The footprint of genome architecture in the largest genome expansion in RNA viruses. PLoS.Pathog. 2013;9(7):e1003500.

56. Bouvet M, Imbert I, Subissi L, *et al.* RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc.Natl. Acad.Sci.U.S.A 2012;109(24):9372-9377.

57. Eckerle LD, Lu X, Sperry SM, *et al.* High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. J.Virol. 2007;81(22):12135-12144.

58. Eckerle LD, Becker MM, Halpin RA, *et al.* Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. PLoS.Pathog. 2010;6(5):e1000896.

59. Pasternak AO, Spaan WJ, Snijder EJ. Nidovirus transcription: how to make sense…? J.Gen.Virol. 2006;87(Pt 6):1403-1421.

60. Makino S, Stohlman SA, Lai MM. Leader sequences of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. Proc.Natl.Acad.Sci.U.S.A 1986; 83(12):4204-4208.

61. Pasternak AO, van den Born E, Spaan WJ, *et al.* Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J. 2001;20(24):7220-7228.

62. Pasternak AO, van den Born E, Spaan WJ, *et al.* The stability of the duplex between sense and antisense transcription-regulating sequences is a crucial factor in arterivirus subgenomic mRNA synthesis. J.Virol. 2003;77(2):1175-1183.

63. van Marle G, Dobbe JC, Gultyaev AP, *et al.* Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. Proc.Natl. Acad.Sci.U.S.A 1999;96(21):12056-12061.

64. Zuniga S, Sola I, Alonso S, *et al.* Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J.Virol. 2004;78(2):980-994.

65. Tijms MA, van Dinten LC, Gorbalenya AE, *et al.* A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. Proc.Natl.Acad.Sci.U.S.A 2001;98(4):1889-1894.

66. Tijms MA, Nedialkova DD, Zevenhoven-Dobbe JC, *et al.* Arterivirus subgenomic mRNA synthesis and virion biogenesis depend on the multifunctional nsp1 autoprotease. J.Virol. 2007;81(19): 10496-10505.

67. van Dinten LC, van Tol H, Gorbalenya AE, *et al.* The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. J.Virol. 2000;74(11):5213-5223.

68. van Marle G, van Dinten LC, Spaan WJ, *et al.* Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. J.Virol. 1999;73(7):5274-5281.

69. Posthuma CC, Nedialkova DD, Zevenhoven-Dobbe JC, *et al.* Site-directed mutagenesis of the Nidovirus replicative endoribonuclease NendoU exerts pleiotropic effects on the arterivirus life cycle. J.Virol. 2006;80(4):1653-1661.

70. Enjuanes L, Almazan F, Sola I, *et al.* Biochemical aspects of coronavirus replication and virus-host interaction. Annu.Rev.Microbiol. 2006;60:211-230.

71. Melnikov S, Ben-Shem A, Garreau de LN, *et al.* One core, two shells: bacterial and eukaryotic ribosomes. Nat.Struct.Mol.Biol. 2012;19(6):560-567.

72. Lopez-Lastra M, Ramdohr P, Letelier A, *et al.* Translation initiation of viral mRNAs. Rev.Med.Virol. 2010;20(3):177-195.

73. Terenin IM, Andreev DE, Dmitriev SE, *et al.* A novel mechanism of eukaryotic translation initiation that is neither m7G-cap-, nor IRES-dependent. Nucleic Acids Res. 2013;41(3):1807-1816.

74. Nicholson BL, White KA. 3' Cap-independent translation enhancers of positive-strand RNA plant viruses. Curr.Opin.Virol. 2011;1(5):373-380.

75. Cowley JA, Dimmock CM, Walker PJ. Gill-associated nidovirus of Penaeus monodon prawns transcribes 3'-coterminal subgenomic mRNAs that do not possess 5'-leader sequences. J.Gen.Virol. 2002;83(Pt 4):927-935.

76. Dye C, Siddell SG. Genomic RNA sequence of feline coronavirus strain FCoV C1Je. J.Feline.Med. Surg. 2007;9(3):202-213.

77. Schutze H, Ulferts R, Schelle B, *et al.* Characterization of White bream virus reveals a novel genetic cluster of nidoviruses. J.Virol. 2006;80(23):11598-11609.

78. Wang C, Zhao Q, Liang C, *et al.* Complete genome sequence of a highly pathogenic porcine reproductive and respiratory syndrome virus variant. J.Virol. 2012;86(16):8906.

79. Zirkel F, Roth H, Kurth A, *et al.* Identification and characterization of genetically divergent members of the newly established family *Mesoniviridae*. J.Virol. 2013;87(11):6346-6358.

80. van Vliet AL, Smits SL, Rottier PJ, *et al.* Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. EMBO J. 2002;21(23):6571-6580.

81. Lai MM, Stohlman SA. Comparative analysis of RNA genomes of mouse hepatitis viruses. J.Virol. 1981;38(2):661-670.

82. Lai MM, Patton CD, Stohlman SA. Further characterization of mRNA's of mouse hepatitis virus: presence of common 5'-end nucleotides. J.Virol. 1982;41(2):557-565.

83. Sagripanti JL, Zandomeni RO, Weinmann R. The cap structure of simian hemorrhagic fever virion RNA. Virology 1986;151(1):146-150.

84. Banerjee S, Narayanan K, Mizutani T, *et al.* Murine coronavirus replication-induced p38 mitogen-activated protein kinase activation promotes interleukin-6 production and virus replication in cultured cells. J.Virol. 2002;76(12):5937-5948.

85. Burgui I, Yanguez E, Sonenberg N, *et al.* Influenza virus mRNA translation revisited: is the eIF4E cap-binding factor required for viral mRNA translation? J.Virol. 2007;81(22):12427-12438.

86. Chen Z, Faaberg KS, Plagemann PG. Determination of the 5' end of the lactate dehydrogenase-elevating virus genome by two independent approaches. J.Gen.Virol. 1994;75 (Pt 4):925-930.

87. Fraser CS, Doudna JA. Structural and mechanistic insights into hepatitis C viral translation initiation. Nat.Rev.Microbiol. 2007;5(1):29-38.

88. Bidet K, Garcia-Blanco MA. Flaviviral RNAs: weapons and targets in the war between virus and host. Biochem.J. 2014;462(2):215-230.

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

90. O'Connor JB, Brian DA. Downstream ribosomal entry for translation of coronavirus TGEV gene 3b. Virology 2000;269(1):172-182.

91. Thiel V, Siddell S. Translation of the MHV sM protein is mediated by the internal entry of ribosomes on mRNA 5. Adv.Exp.Med.Biol. 1995;380:311-315.

92. Neugebauer KM. On the importance of being co-transcriptional. J.Cell Sci. 2002;115(Pt 20):3865-3871.

93. Imbert I, Guillemot JC, Bourhis JM, *et al.* A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. EMBO J. 2006;25(20):4933-4942.

94. te Velthuis AJ, Arnold JJ, Cameron CE, *et al.* The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. Nucleic Acids Res. 2010;38(1):203-214.

95. te Velthuis AJ, van den Worm SH, Snijder EJ. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. Nucleic Acids Res. 2012;40(4):1737-1747.

96. Hofmann MA, Brian DA. The 5' end of coronavirus minus-strand RNAs contains a short poly(U) tract. J.Virol. 1991;65(11):6331-6333.

97. Ghosh A, Lima CD. Enzymology of RNA cap synthesis. Wiley.Interdiscip.Rev.RNA. 2010;1(1):152-172.

98. Decroly E, Ferron F, Lescar J, *et al.* Conventional and unconventional mechanisms for capping viral mRNA. Nat.Rev.Microbiol. 2012;10(1):51-65.

99. Liu H, Kiledjian M. Decapping the message: a beginning or an end. Biochem.Soc.Trans. 2006;34(Pt 1):35-38.

100. Rhoads RE. eIF4E: new family members, new binding partners, new roles. J.Biol.Chem. 2009; 284(25):16711-16715.

101. Daffis S, Szretter KJ, Schriewer J, *et al.* 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature 2010;468(7322):452-456.

102. Zust R, Cervantes-Barragan L, Habjan M, *et al.* Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. Nat. Immunol. 2011;12(2):137-143.

103. Ivanov KA, Ziebuhr J. Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. J.Virol. 2004; 78(14):7833-7838.

104. Ivanov KA, Thiel V, Dobbe JC, *et al.* Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. J.Virol. 2004;78(11):5619-5632.

105. Bouvet M, Debarnot C, Imbert I, *et al. In vitro* reconstitution of SARS-coronavirus mRNA cap methylation. PLoS.Pathog. 2010;6(4):e1000863.

106. Chen Y, Cai H, Pan J, *et al.* Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. Proc.Natl.Acad.Sci.U.S.A 2009;106(9):3484-3489.

107. Decroly E, Imbert I, Coutard B, *et al.* Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. J.Virol. 2008;82(16):8071-8084.

108. Decroly E, Debarnot C, Ferron F, *et al.* Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. PLoS.Pathog. 2011;7(5): e1002059.