Current Topics in Microbiology and Immunology

Ralph A. Tripp
S. Mark Tompkins *Editors*

Roles of Host Gene and Non-coding RNA Expression in Virus Infection

Springer

Current Topics in Microbiology and Immunology

Volume 419

More information about this series at http://www.springer.com/series/82

Ralph A. Tripp · S. Mark Tompkins
Editors

Roles of Host Gene and Non-coding RNA Expression in Virus Infection

Responsible series editor: Richard W. Compans

Springer

Editors
Ralph A. Tripp
Department Infectious Diseases,
College of Veterinary Medicine
University of Georgia
Athens, GA, USA

S. Mark Tompkins
Center for Vaccines and Immunology
University of Georgia
Athens, GA, USA

ISSN 0070-217X ISSN 2196-9965 (electronic)
Current Topics in Microbiology and Immunology
ISBN 978-3-030-05368-0 ISBN 978-3-030-05369-7 (eBook)
https://doi.org/10.1007/978-3-030-05369-7

Library of Congress Control Number: 2018963295

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Viruses are obligate parasites relying on the exploitation of host cell processes and resources for replication. The interplay between host and viruses remains largely unknown. Life cycles for individual viruses have been defined with functions ascribed for many viral proteins. The discovery of RNA interference (RNAi) and the subsequent development of tools to specifically silence individual cellular genes enabled genome-wide studies, interrogating gene function in a spectrum of processes, advancing developmental biology, and infectious disease. The advent of functional genomics allowed for the interrogation of the virus–host cell interactions and probing the genome for a role in the virus replication.

Libraries of arrayed siRNAs against human or mouse genomes have been available for more than a decade. More recently, microRNA (miRNA) mimic and inhibitor libraries have also become available for genome-wide screening, and most recently gene editing, e.g. CRISPR/Cas, has also become available. These approaches combined with transcriptomic and proteomic analyses have enabled the identification of new players in the host–virus interactome. Importantly, advances in recombinant technology, virology, and systems biology have allowed mapping of the interaction between cellular and viral gene products, including viral and cellular non-coding RNAs allowing for a better understanding of novel gene functions and pro- and anti-viral activities. These discoveries have provided an opportunity for the development of novel therapeutics and approaches to improve viral vaccines and vaccine production.

This volume presents a current understanding of the interplay between host cells and viruses during infection and replication. The first chapters present our knowledge of coronavirus, flavivirus, and human immunodeficiency virus (HIV), virus–cell interactions, i.e. three positive-sense RNA viruses (*Coronaviridae*, *Flaviviridae*, and *Retroviridae*), respectively. The volume then moves to address to negative-sense RNA viruses, with chapters on Ebola virus (*Filoviridae*), influenza virus (*Orthomyxoviridae*), and two viruses from the *Paramyxoviridae* family. The respiratory syncytial virus (RSV) chapter discusses the role of miRNAs in infection, while the henipavirus chapter explores diverse aspects of virus–host interactions. The volume finishes with a chapter on non-coding RNAs involved in

herpesvirus infection, a double-stranded DNA virus (*Herpesviridae*). These chapters capture many aspects of viral genomes and life cycles, including segmented, integrating, and latent genomes, acute, chronic, and latent infections, as well as vector-borne viruses. This volume provides a representation of virus–host interactions and a valuable resource for advancing our understanding. We are grateful to the authors for their expertise and contributions to this remarkable volume.

Athens, USA

Ralph A. Tripp
S. Mark Tompkins

Contents

Host Factors in Coronavirus Replication

Adriaan H. de Wilde, Eric J. Snijder, Marjolein Kikkert and Martijn J. van Hemert

Abstract Coronaviruses are pathogens with a serious impact on human and animal health. They mostly cause enteric or respiratory disease, which can be severe and life threatening, e.g., in the case of the zoonotic coronaviruses causing severe acute respiratory syndrome (SARS) and Middle East Respiratory Syndrome (MERS) in humans. Despite the economic and societal impact of such coronavirus infections, and the likelihood of future outbreaks of additional pathogenic coronaviruses, our options to prevent or treat coronavirus infections remain very limited. This highlights the importance of advancing our knowledge on the replication of these viruses and their interactions with the host. Compared to other +RNA viruses, coronaviruses have an exceptionally large genome and employ a complex genome expression strategy. Next to a role in basic virus replication or virus assembly, many of the coronavirus proteins expressed in the infected cell contribute to the coronavirus-host interplay. For example, by interacting with the host cell to create an optimal environment for coronavirus replication, by altering host gene expression or by counteracting the host's antiviral defenses. These coronavirus–host interactions are key to viral pathogenesis and will ultimately determine the outcome of infection. Due to the complexity of the coronavirus proteome and replication cycle, our knowledge of host factors involved in coronavirus replication is still in an early stage compared to what is known for some other +RNA viruses. This review summarizes our current understanding of coronavirus–host interactions at the level of the infected cell, with special attention for the assembly and function of the viral RNA-synthesising machinery and the evasion of cellular innate immune responses.

A.H. de Wilde · E.J. Snijder · M. Kikkert · M.J. van Hemert (✉)
Molecular Virology Laboratory, Department of Medical Microbiology,
Leiden University Medical Center, Leiden, The Netherlands
e-mail: m.j.van_hemert@lumc.nl

Current Topics in Microbiology and Immunology (2018) 419:1–42
DOI 10.1007/82_2017_25

Published Online: 23 June 2017

Contents

1 Introduction

Around the end of 2002, an outbreak of a previously unknown severe acute respiratory syndrome (SARS) started in South East China and Hong Kong. Accelerated by air travel, the disease rapidly spread to several parts of the world and displayed pandemic potential. SARS-coronavirus (SARS-CoV) was identified as the causative agent of this zoonotic infection (Drosten et al. 2003; Ksiazek et al. 2003; Kuiken et al. 2003; Peiris et al. 2003), which resulted in >8000 laboratory-confirmed cases and 774 associated deaths worldwide (WHO 2004). Although in terms of death toll not comparable to influenza, HIV or HCV, the 2003 SARS-CoV outbreak caused worldwide public concern and seriously affected the global economy [estimated losses $30–100 billion; (Keogh-Brown and Smith 2008)]. SARS-CoV initially causes lower respiratory tract disease, which can lead to a progressive and potentially lethal atypical pneumonia with clinical symptoms that include fever, malaise, lymphopenia, and in some cases also diarrhea. Two years after the outbreak, horseshoe bats were identified as the likely reservoir of the SARS virus, whereas civet cats probably have served as intermediate host during the zoonotic transfer to humans (Lau et al. 2005; Li et al. 2005b). Adaptation to the human host required a small number of mutations in the receptor-binding domain of the SARS-CoV spike (S) protein, which mediates cell binding and entry (Li et al. 2005c) (see Chap. 2). There is increasing evidence that SARS-like coronaviruses continue to circulate in bats and that these may have the potential to readily cross the species barrier and emerge as human pathogens (Ge et al. 2013; Menachery et al. 2015). Such zoonotic scenarios therefore remain a serious public health concern.

Almost a decade after the SARS-CoV outbreak, the next zoonotic coronavirus emerged: Middle East Respiratory Syndrome coronavirus (MERS-CoV) (de Groot et al. 2013). The virus was first isolated in June 2012 from a 60-year-old Saudi Arabian male who died from acute respiratory distress syndrome (ARDS) and

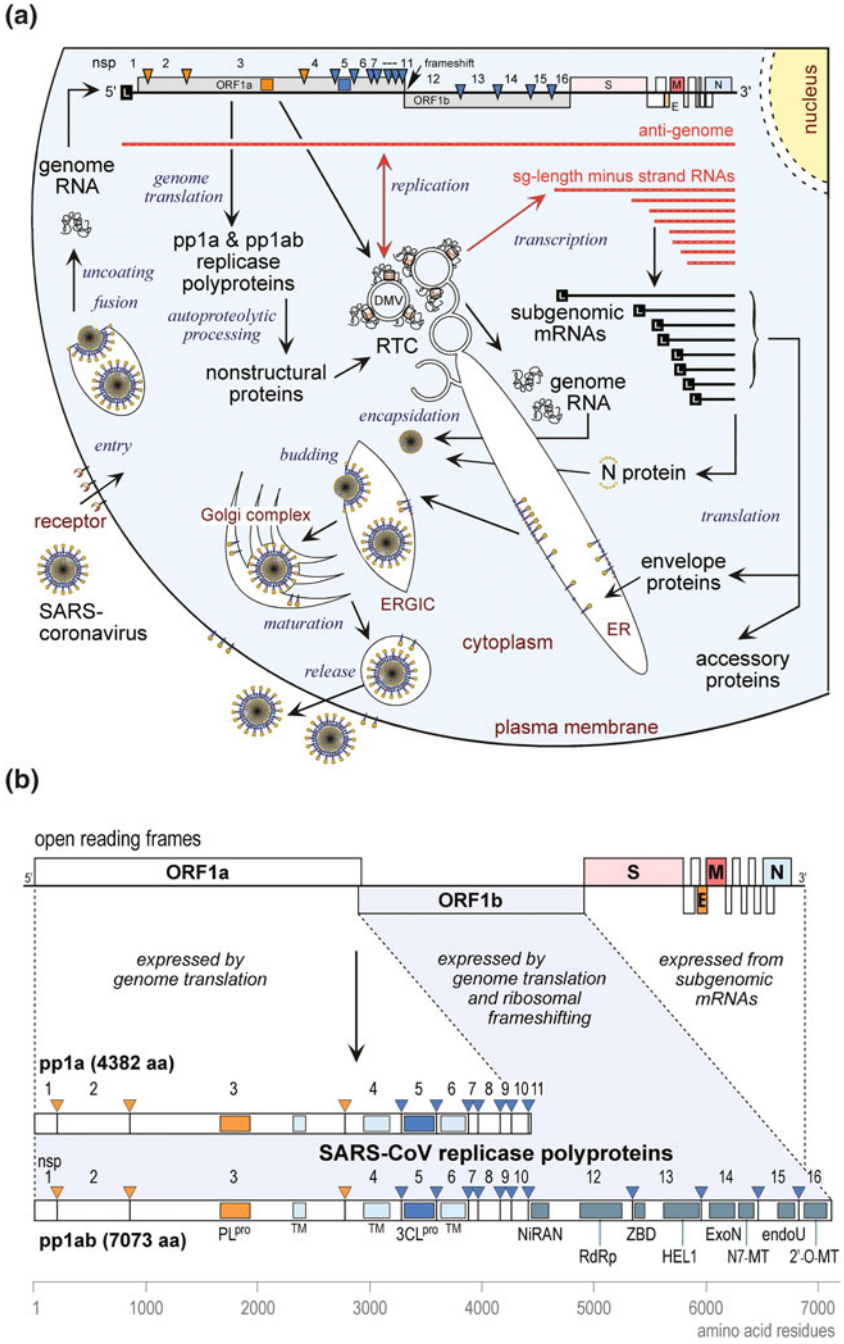
multiple organ failure, including renal failure (Zaki et al. 2012; van Boheemen et al. 2012). Also MERS-CoV can cause a lower respiratory tract infection with symptoms that include coughing and high fever. By the end of 2016, more than 1850 laboratory-confirmed MERS-CoV cases had been recorded, with a mortality rate of about 35% (WHO 2016). MERS-CoV is assumed to be transmitted to humans from camels and serological studies in the latter animals revealed that they have harbored MERS-CoV or MERS-CoV-like viruses for decades (Muller et al. 2014).

Besides the zoonotic SARS- and MERS-CoVs, the coronavirus family includes four 'established' human coronaviruses (HCoVs), of which HCoV-OC43 and -229E have already been known since the 1960s. These two viruses cause mild respiratory disease and, after rhinoviruses, are a leading cause of common colds (10–30% of the cases) (van der Hoek 2007; McIntosh et al. 1967; Hamre and Procknow 1966). More recently, following intensified screening for coronaviruses, two additional HCoVs were discovered, HCoV-NL63 (van der Hoek et al. 2004) and HCoV-HKU1 (Woo et al. 2005). Interestingly, recent findings suggest that also HCoV-NL63, -229E, and -OC43 originate from zoonotic transfer from bats (Huynh et al. 2012; Corman et al. 2016; Vijgen et al. 2006; Corman et al. 2015). Coronaviruses also cause a range of infectious diseases in animal species, some with serious (economical) consequences for the livestock industry. This is illustrated by the recent emergence of a novel variant of porcine epidemic diarrhea virus, which is closely related to a strain that caused a large outbreak in China in 2010, killing almost one million piglets [for a recent review, see (Lin et al. 2016)].

The economic impact of coronavirus infections, the past and likely future emergence of pathogenic zoonotic coronaviruses and the lack of effective antiviral strategies have made it painfully clear that our preparedness to treat or prevent coronavirus infections are very limited. This highlights the importance of advancing our knowledge on the replication of these viruses and their interactions with the host.

Coronaviruses are positive-stranded RNA (+RNA) viruses with, for this kind of viruses, exceptionally large genomes of ~30 kb. They have a polycistronic genome organization and employ a unique transcription mechanism to generate a nested set of subgenomic (sg) mRNAs. These are used to express the open reading frames (ORFs) located downstream of the replicase ORFs 1a and 1b (see Fig. 1a), which encode structural and accessory proteins. The sg mRNAs are 3′ co-terminal but they also contain a common 5′ leader sequence. The leader and 'body' segments of the sg RNAs are joined during discontinuous negative strand RNA synthesis, which produces a subgenome-length template for each of the sg mRNAs [(Sawicki and Sawicki 1995), for a recent review, see (Sola et al. 2015)].

Following entry and uncoating, the coronavirus replicative cycle (see Fig. 1a) starts with the translation of the 5′-proximal ORFs of the viral genome (ORF1a and ORF1b), which results in the synthesis of two large replicase polyproteins (pp1a and pp1ab). Synthesis of pp1ab, a C-terminally extended form of pp1a, involves a -1 ribosomal frameshift (RFS) into ORF1b occurring near the 3′ end of ORF1a. This regulatory mechanism is thought to have evolved to downregulate expression levels of ORF1b-encoded proteins compared to ORF1a-encoded nonstructural proteins (nsps) (Brierley and Dos Ramos 2006; Brierley et al. 1989). Ultimately, 15 or 16


(a)
nsp
frameshift
ORF1a
ORF1b
S
M
N
E
nucleus
anti-genome
genome RNA
genome translation
replication
sg-length minus strand RNAs
transcription
pp1a & pp1ab replicase polyproteins
uncoating
fusion
autoproteolytic processing
DMV
RTC
subgenomic mRNAs
nonstructural proteins
genome RNA
encapsidation
entry
budding
N protein
translation
receptor
Golgi complex
envelope proteins
SARS-coronavirus
ERGIC
maturation
cytoplasm
ER
accessory proteins
release
plasma membrane
(b)
open reading frames
ORF1a
ORF1b
S
M
N
E
expressed by genome translation
expressed by genome translation and ribosomal frameshifting
expressed from subgenomic mRNAs
pp1a (4382 aa)
SARS-CoV replicase polyproteins
pp1ab (7073 aa)
PLpro
TM
3CLpro
NiRAN
RdRp
ZBD
HEL1
ExoN
N7-MT
endoU
2'-O-MT
amino acid residues

◂**Fig. 1** Outline of the coronavirus replicative cycle and replicase polyprotein organization, based on SARS-CoV. **a** Schematic overview of the coronavirus replicative cycle. Following entry by receptor-mediated endocytosis and release of the genome into the cytosol, genome translation yields the pp1a and pp1ab replicase polyproteins. Following polyprotein cleavage by multiple internal proteases, the viral nsps assemble into an RTC that engages in minus-strand RNA synthesis. Both full-length and subgenome (sg)-length minus strands are produced, with the latter templating the synthesis of the sg mRNAs required to express the structural and accessory protein genes residing in the 3′-proximal quarter of the genome. Ultimately, novel genomes are packaged into nucleocapsids that become enveloped by budding from smooth intracellular membranes, after which the new virions leave the cell by following the exocytic pathway. See text for more details. **b** The 14 open reading frames in the genome are indicated, i.e., the replicase ORFs 1a and 1b, the four common CoV structural protein genes (S, E, M, and N) and the ORFs encoding so-called 'accessory proteins.' The bottom panel explains the organization and proteolytic processing of the pp1a and pp1ab replicase polyproteins, the latter being produced by −1 ribosomal frameshifting. The nsp3 (PL^{pro}) and nsp5 ($3CL^{pro}$) proteases and their cleavage sites are indicated in matching colors. The resulting 16 cleavage products [nonstructural proteins (nsps)] are indicated, as are the conserved replicase domains. Domain abbreviations and corresponding nsp numbers: PL^{pro}, papain-like proteinase (nsp3); $3CL^{pro}$, 3C-like protease (nsp5); TM, transmembrane domain (nsp3, nsp4, and nsp6); NiRAN, nidovirus RdRp-associated nucleotidyl transferase (nsp12); RdRp, RNA-dependent RNA polymerase (nsp12); ZBD, zinc-binding domain (nsp13); HEL1, superfamily 1 helicase (nsp13); ExoN, exoribonuclease (nsp14); N7-MT, N7-methyl transferase (nsp14); endoU, uridylate-specific endoribonuclease (nsp15); 2′-O-MT, 2′-O-methyl transferase (nsp16). Adopted with permission from (Snijder et al. 2016)

mature replicase proteins are released from pp1a and pp1ab due to proteolytic cleavages performed by two or three ORF1a-encoded proteases. Nsp3 contains one or two papain-like protease domains ($PL1^{pro}$ and $PL2^{pro}$, or PL^{pro} for SARS-CoV and infectious bronchitis virus) that process the nsp1-4 part of the replicase polyproteins. The remaining cleavage sites are processed by the viral main protease that resides in nsp5, a chymotrypsin-like enzyme also known as 3C-like protease (Snijder et al. 2016). A schematic overview of the proteolytic processing and domain structure of the SARS-CoV replicase is presented in Fig. 1b. The replicase proteins contain a variety of (enzymatic) activities and functions that are required for viral RNA synthesis and capping (Perlman and Netland 2009; Snijder et al. 2016), such as the RNA-dependent RNA polymerase (RdRp; nsp12), a helicase (nsp13), RNA cap-modifying methyltransferases (nsp14 and nsp16), and an exoribonuclease (nsp14). Together with recruited host cell proteins, the coronavirus nsps form membrane-associated replication and transcription complexes [RTCs; (van Hemert et al. 2008)], which localize to a network of virus-induced membrane structures in the perinuclear region of the infected cell (Knoops et al. 2008; Gosert et al. 2002; van der Meer et al. 1999; Brockway et al. 2003; Stertz et al. 2007; Ulasli et al. 2010). Many of the nsps appear to have multiple functions in the synthesis or processing of viral RNA, or in virus–host interactions aiming to create an optimal environment for coronavirus replication, for example by facilitating viral entry, gene expression, RNA synthesis or virus release. Moreover, to further enhance viral replication, host gene expression and antiviral defenses are targeted in several ways. Coronavirus–host interactions also play a decisive role in viral pathogenesis and the ultimate outcome of infection.

Due to the exceptional size of their +RNA genome and proteome, and the resulting complexity of the interactions with the host, our knowledge of host factors involved in coronavirus replication is still in an early stage compared to what is known for some other +RNA virus groups. In this review, we will summarize our current understanding of coronavirus–host interactions at the level of the infected cell, with special attention for the assembly and function of the viral RNA-synthesizing machinery and the evasion of cellular innate immune responses.

2 Host Receptors Involved in Coronavirus Entry

Entry into the target cell constitutes the first critical step in the coronavirus replication cycle. The major determinant for this step is the efficient binding of the coronavirus S glycoprotein to a protein-receptor on the cell surface. The coronavirus S protein is a type 1 glycoprotein that consists of S1 and S2 subunits and is present on the virion surface as a trimer. (Li 2016; Hulswit et al. 2016). The S1 region is involved in receptor binding and contains N- and C-terminal domains (S1-NTD and S1-CTD, respectively) (Walls et al. 2016) that may both act as receptor-binding domain (RBD), with the major determinants of cell tropism residing in S1-CTD. The elongated S2 regions form the stalk of the spike trimer and are mainly involved in triggering the fusion of the viral envelope and target cell membranes [for recent reviews on coronavirus entry and spike protein organization, see (Li 2016; Hulswit et al. 2016)].

The S1-NTD is mainly involved in facilitating virus binding and entry, by interacting with glycans on the host cell surface. Based on the crystal structure of the betacoronavirus S1-NTD and the sequence conservation among the S1-NTDs of other coronaviruses, all coronavirus S1-NTDs are thought to share a galectin fold that mediates binding to sialic acids, like N-glycolylneuraminic acid (Neu5Gc), N-acetylneuraminic acid (Neu5Ac), and/or 5-N-acetyl-9-O-acetylneuraminic acid (Neu5,9Ac2) (see (Li 2016), and references herein). An exception is the murine hepatitis virus (MHV) S1-NTD, which binds the N-terminal D1 domain of carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1), a type-I membrane protein belonging to the immunoglobulin superfamily (Walls et al. 2016; Williams et al. 1991).

To mediate entry into host cells, the S1-CTD of most known members of the alphacoronavirus genus interacts with aminopeptidase N (APN) (for an overview and references, see Table 1). However, the alphacoronavirus HCoV-NL63 uses a different type-I membrane glycoprotein, angiotensin-converting enzyme 2 (ACE2) (Wu et al. 2009), which contains a large N-terminal ectodomain composed of two alpha-helical lobes. The same molecule, ACE2, has been identified as a receptor for the zoonotic betacoronavirus SARS-CoV (Li et al. 2003). The betacoronaviruses MERS-CoV and bat coronavirus HKU4 use yet another cellular peptidase for virus entry: dipeptidyl peptidase 4 (DPP4) (Yang et al. 2014; Raj et al. 2013). The MERS-CoV S protein has a higher affinity for human DPP4, while the HKU4 S

Table 1 Overview of known coronavirus entry receptors

Genus	Species:	S1-NTD	S1-CTD	References
Alphacoronavirus	Alphacoronavirus 1	Neu5Gc and Neu5Ac*	APN	(Tresnan et al. 1996; Delmas et al. 1992)
	PEDV	Neu5Ac	APN	(Liu et al. 2015; Li et al. 2007)
	PRCV		APN	(Schultze et al. 1996)
	HCoV-229E		APN	(Yeager et al. 1992)
	HCoV-NL63		ACE2	(Wu et al. 2009)
Betacoronavirus	Betacoronavirus 1	Neu5,9Ac2		(Schultze and Herrler 1992; Krempl et al. 1995)
	MERS-CoV		DPP4	(Raj et al. 2013)
	MHV	CEACAM1		(Williams et al. 1991)
	HKU1	Neu5,9Ac2		(Huang et al. 2015b)
	HKU4		DPP4	(Yang et al. 2014)
	SARS-CoV		ACE2	(Li et al. 2003)
Gammacoronavirus	IBV	Neu5Gc		(Schultze et al. 1993)
Deltacoronavirus	PDCV	Unknown	unknown	

(*Abbreviations PEDV* Porcine epidemic diarrhea virus; *TGEV* Transmissible gastroenteritis coronavirus; *PRCV* Porcine Respiratory coronavirus; *FCoV* Feline coronavirus; *CCoV* Canine coronavirus; *HCOV* Human coronavirus; *BCoV* Bovine coronavirus; *MHV* Murine hepatitis virus; *IBV* Infectious bronchitis virus; *PDCV* Porcine delta coronavirus). *Within the alphacoronavirus 1 species, only for TGEV the sialic acids Neu5Gc and Neu5Ac has been identified as attachment factors

protein binds more strongly to bat DPP4 (Yang et al. 2014). Chemical peptidase inhibitors do not affect virus entry, indicating that SARS-CoV and MERS-CoV receptor usage and entry are independent of the receptor's peptidase activity and merely depend on binding to these particular host receptors (Li et al. 2005c; Raj et al. 2013).

Besides the receptors discussed above, also extracellular, cell surface-associated and/or lysosomal proteases play a role in coronavirus entry by activating the fusion activity of the S protein [for a recent review, see (Li 2016)]. For SARS-CoV, fusion of the viral and cellular membrane is triggered upon cleavage of the S protein by the cell surface-associated transmembrane protease, serine 2 (TMPRSS2) (Glowacka et al. 2011). The same protease is important for cleavage and activation of the HCoV-229E and MERS-CoV S protein (Shirato et al. 2013; Bertram et al. 2013). After endocytosis, the SARS-CoV S protein is cleaved by the lysosomal proteases cathepsin L and cathepsin P in early endosomes, leading to fusion of the virus envelop with the endosome membranes and release of the viral RNA into the cytosol of the infected cell (Huang et al. 2006a, b; Simmons et al. 2005). MERS-CoV entry occurs by a similar mechanism (Shirato et al. 2013; Burkard et al. 2014), although inhibition of the cellular protease furin abolished the entry of MERS-CoV but not SARS-CoV, indicating that furin-mediated cleavage is pivotal