

Distinct Genome Replication and Transcription Strategies within the Growing Filovirus Family

Adam J. Hume^{1,2} and Elke Mühlberger^{1,2}

1 - Department of Microbiology, Boston University School of Medicine, Boston, MA 02118, USA

2 - National Emerging Infectious Diseases Laboratories, Boston University, Boston, MA 02118, USA

Correspondence to Elke Mühlberger: National Emerging Infectious Diseases Laboratories, Boston University, 620 Albany Street, Boston, MA 02118, USA. muehlber@bu.edu
<https://doi.org/10.1016/j.jmb.2019.06.029>

Abstract

Research on filoviruses has historically focused on the highly pathogenic ebola- and marburgviruses. Indeed, until recently, these were the only two genera in the filovirus family. Recent advances in sequencing technologies have facilitated the discovery of not only a new ebolavirus, but also three new filovirus genera and a sixth proposed genus. While two of these new genera are similar to the ebola- and marburgviruses, the other two, discovered in saltwater fishes, are considerably more diverse. Nonetheless, these viruses retain a number of key features of the other filoviruses. Here, we review the key characteristics of filovirus replication and transcription, highlighting similarities and differences between the viruses. In particular, we focus on key regulatory elements in the genomes, replication and transcription strategies, and the conservation of protein domains and functions among the viruses. In addition, using computational analyses, we were able to identify potential homology and functions for some of the genes of the novel filoviruses with previously unknown functions. Although none of the newly discovered filoviruses have yet been isolated, initial studies of some of these viruses using minigenome systems have yielded insights into their mechanisms of replication and transcription. In general, the *Cuevavirus* and proposed *Dianlovirus* genera appear to follow the transcription and replication strategies employed by the ebola- and marburgviruses, respectively. While our knowledge of the fish filoviruses is currently limited to sequence analysis, the lack of certain conserved motifs and even entire genes necessitates that they have evolved distinct mechanisms of replication and transcription.

© 2019 Elsevier Ltd. All rights reserved.

Introduction

Filoviruses belong to the group of nonsegmented negative sense (NNS) RNA viruses. In the past, research mainly focused on the most prominent members of the filovirus family, Ebola virus (EBOV) and Marburg virus (MARV), both of which are highly pathogenic for humans [1,2]. EBOV and MARV are genetically distinct and belong to different filovirus genera, *Ebolavirus* and *Marburgvirus* (Fig. 1a). With the advent of high-throughput sequencing technology, the filovirus family has expanded. Newly discovered filoviruses that have not been associated with outbreaks in humans yet include Bombali virus (BOMV), a proposed new member of the *Ebolavirus* genus [4]; Lloviu virus (LLOV) [3,5,6]; Mènglà virus (MLAV) [7,8]; and some fascinating fish filoviruses

whose genome organization is very different from the “conventional” filovirus genomes [9] (Fig. 1b). Among these novel filoviruses, LLOV, MLAV, and the fish filoviruses Xīlǎng and Huángjiào viruses (XILV and HUJV, respectively) are genetically distinct from ebola- and marburgviruses and have been, or will be, assigned to distinct filovirus genera [3] (Fig. 1a).

In line with the proposed chiropteran origin of the previously identified filoviruses, some of the new filoviruses were discovered when their RNA was found in different bat species: BOMV RNA was isolated from little free-tailed bats (*Chaerephon pumilus*) and Angolan free-tailed bats (*Mops condylurus*) in Sierra Leone and Kenya [4,10], LLOV RNA was isolated from Schreiber's long-fingered bats (*Miniopterus schreibersii*) in Spain and Hungary

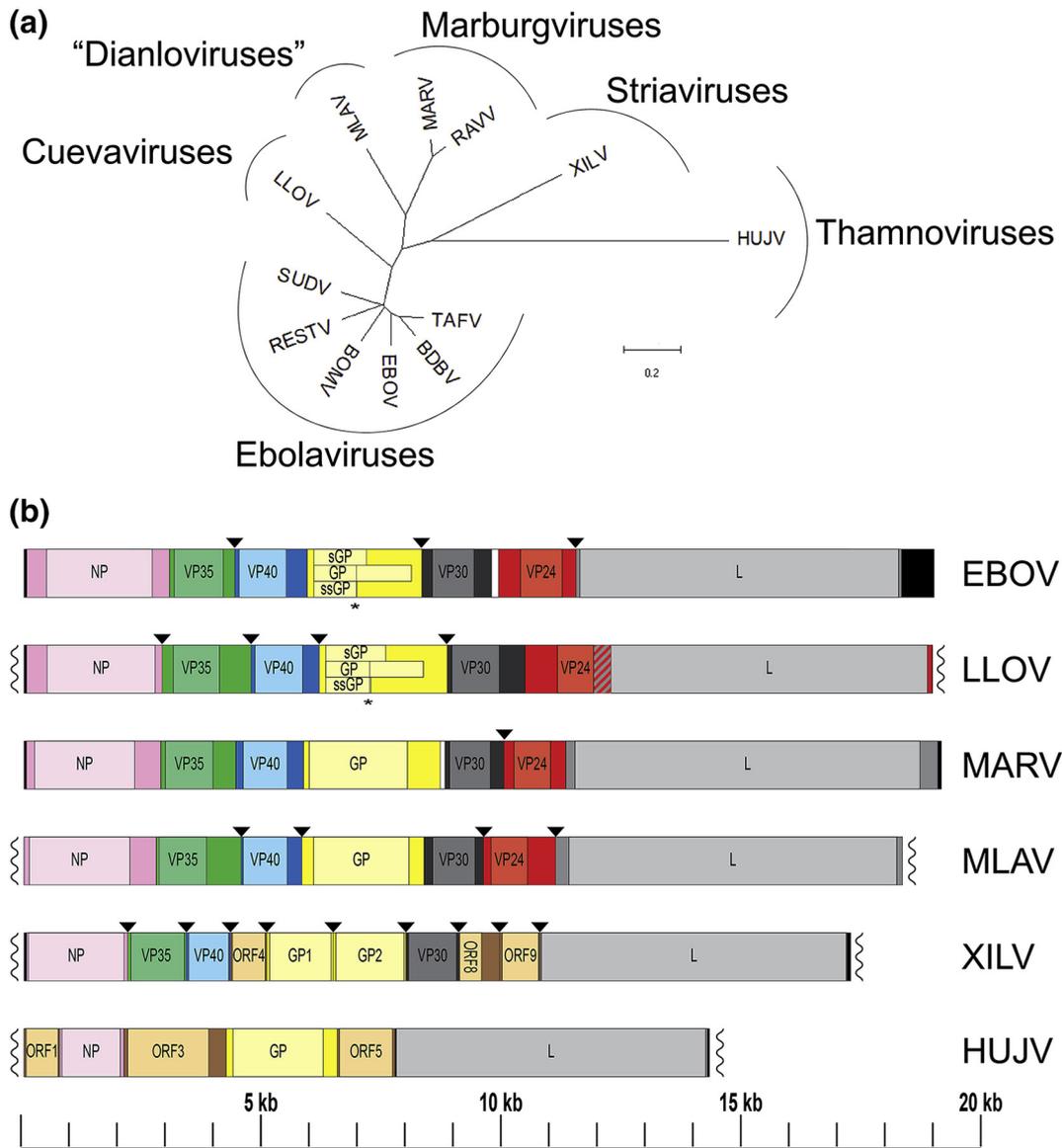


Fig. 1. Comparison of filovirus genomes. (a) A phylogenetic tree of the indicated filoviruses was performed by aligning the L protein sequences using the Geneious Prime software. Labeled arcs indicate the officially assigned filovirus genera [3], with the exception of the “Dianloviruses” genus, which has been proposed but not yet accepted. BOMV was added to the *Ebolavirus* genus, although it has not been assigned officially to the genus. (b) Schematics of filovirus genomes. Each filovirus genus is represented by the prototypical virus. Schematics are depicted to scale. Leader and trailer sequences are indicated by black bars, with missing sequence information indicated by squiggly lines. Genes are depicted as colored bars with lighter portions indicating open reading frames (ORFs). Homologous genes are colored similarly. Since the LLOV VP24 and L genes appear to form a bicistronic transcript, the untranslated region separating the VP24 and L ORFs is colored by stripes. For XILV and HUJV, genes with no homology to genes of other filoviruses and no known function are indicated in brown with lighter portions indicating ORFs. Sites of gene overlaps are indicated by black arrows. Intergenic regions are in white. Sites of cotranscriptional mRNA editing are indicated by asterisks. BDBV, Bundibugyo virus; BOMV, Bombali virus; EBOV, Ebola virus; GP, glycoprotein; HUJV, Huángjiāo virus; L, large protein; LLOV, Lloviu virus; MARV, Marburg virus; MLAV, Měnglà virus; NP, nucleoprotein; ORF, open reading frame; RAVV, Ravn virus; RESTV, Reston virus; sGP, soluble glycoprotein; ssGP, small soluble glycoprotein; SUDV, Sudan virus; TAFV, Tai Forest virus; XILV, Xīlǎng virus.

[5,6], and MLAV RNA was isolated from a bat of the genus *Rousettus* in China [7]. In contrast, RNA from the newly discovered XILV and HUJV was isolated

from saltwater fishes, the striated frogfish (*Antennarius striatus*) and the greenfin horse-faced filefish (*Thamnaconus septentrionalis*), respectively, in the

South and East China Seas [9]. While most of the known marburg- and ebolaviruses cause disease in humans, the pathogenic potential of the newly discovered filoviruses is not known. Of note, none of these viruses have been isolated from either an animal host or human patients—the available sequence information is based on viral RNA extracted from bats or fish and is, with the exception of BOMV, incomplete with various amounts of terminal genomic sequences missing (Fig. 1b).

The unexpected geographic and genetic diversity of filoviruses raises concerns about unpredictable outbreaks caused by the new members of the family. Having a thorough understanding of the filoviral genome replication mechanisms, including identifying similarities and differences in the replication strategies of the various filoviruses, will be instrumental for designing and developing countermeasures in order to be prepared for potential outbreak scenarios. Here, we provide an overview of filovirus genome replication and mRNA transcription strategies with a focus on the differences between filovirus genera. In addition to highlighting recent data regarding filoviruses, we also suggest novel functions and potential homology for some of the XILV and HUJV genes based on computational analyses.

Marburg- and Ebolavirus Replication Cycle—A Brief Overview

Structure of marburg- and ebolavirus particles

Typically, marburg- and ebolavirus particles are filamentous in shape but their shape can also be pleomorphic [11–14] (Fig. 2). The viral particles are enwrapped by a host-derived membrane covered with protrusions which are formed by the sole viral surface protein, the glycoprotein (GP). During passage through the infected cell, the GP precursor is proteolytically cleaved into two covalently linked subunits, GP_{1,2}, that are inserted in the viral membrane as trimers (reviewed in Refs. [16,17]). The inner leaflet of the viral membrane is lined with the matrix protein, viral protein 40 (VP40). VP40 drives viral budding and particle formation and is able to assemble virus-like particles in the absence of other viral proteins (reviewed in Refs. [16,18,19]). VP40 gives marburg- and ebolavirus particles their typical filamentous shape. The inner core of the marburg- and ebolavirus particles is the so-called nucleocapsid, a ribonucleoprotein complex that is composed of the viral RNA genome and five associated viral proteins, nucleoprotein (NP; encapsidates the RNA genome), VP35 (cofactor of the viral RNA-dependent RNA polymerase), VP30 (transcription activator), VP24 (nucleocapsid maturation factor), and large protein (L; catalytic subunit of the

RNA-dependent RNA polymerase), of which three—NP, VP35, and VP24—are essential for the formation of mature nucleocapsids [14,20–27] (Fig. 7). The nucleocapsid contains all the viral components required for genome replication and transcription. A more detailed description of the nucleocapsid and its components is provided in the [Nucleocapsid](#) section.

The filoviral replication cycle

Filovirus entry is a complex, multi-step process that is very well reviewed in Ref. [17]. Briefly, attachment of marburg- and ebolavirus particles to the cell occurs through interaction of the GP₁ subunit with various host cell factors, including C-type lectins and glycosaminoglycans, or through interaction of phosphatidylserine on the outer leaflet of the viral membrane with phosphatidylserine receptors, such as members of the Tyro3, Axl, and Mer (TAM), and T-cell immunoglobulin and mucin domain (TIM) families of receptors (reviewed in Ref. [17]). Following attachment, the dominant uptake mechanism of the viral particles seems to be macropinocytosis [29–31]. In the endosomes, GP_{1,2} is processed by cathepsins [32]. This cleavage step exposes the receptor binding site on GP₁ and facilitates binding of GP_{1,2} to the filovirus receptor, endosomal membrane protein Niemann Pick C1 (NPC1) [33,34]. Similar to marburg- and ebolavirus glycoproteins, LLOV and MLAV glycoproteins are able to use NPC1, including human NPC1, as an entry receptor [7,35,36]. This not only confirms the close relationship of these viruses to marburg- and ebolaviruses but also shows their ability to infect human cells. It is not yet known whether XILV and HUJV use NPC1 as an entry receptor. The low sequence similarity of the putative XILV and HUJV glycoproteins with mammalian filovirus glycoproteins opens the possibility that the mechanism of entry is divergent. Dramatic structural rearrangements of GP_{1,2} precede GP₂-driven fusion of the viral and the endosomal membranes (reviewed in Ref. [16]). Following fusion, the nucleocapsid is released into the cytoplasm of the infected cell where the viral RNA-dependent RNA polymerase consisting of L and VP35 starts to sequentially transcribe the viral genes using the encapsidated viral genome as a template (Fig. 3). For the ebolaviruses and LLOV, transcription is assisted by the transcription factor VP30. The role of VP30 in marburgvirus transcription is less well understood. This early or “primary” transcription is exclusively facilitated by the viral proteins that were brought into the cell [18,26,37]. Each filoviral gene is flanked by transcription start and stop signals that determine the beginning and the end of the transcribed mRNA. During transcription, the viral polymerase enters the genome at a single binding site at the genomic 3' end and scans for the

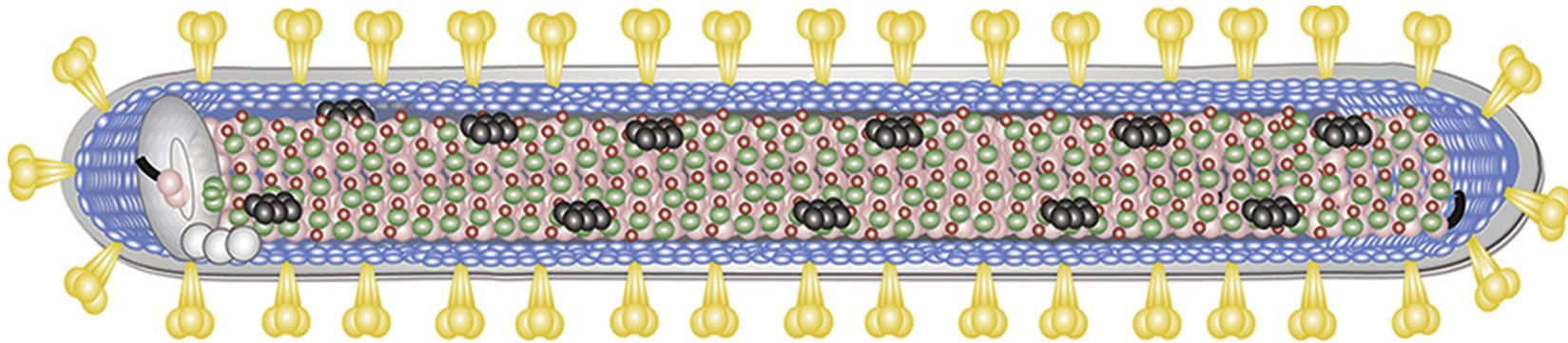


Fig. 2. Scheme of filovirus particle. The cell membrane-derived envelope (outer gray oval) is studded with viral glycoprotein trimers (yellow lollipops). The assembly of the matrix protein VP40 (blue) into long, filamentous particles imparts the filovirus particles with their namesake virion shape. The viral nucleocapsid consists of the negative sense genome (black line) enwrapped in the nucleoprotein (NP, pink spheres) along with the polymerase cofactor VP35 (green spheres), the transcription factor VP30 (dark gray spheres), the nucleocapsid protein VP24 (red spheres), and the catalytic polymerase subunit L (light gray ring with appendices). Based on the structure of the vesicular stomatitis virus L protein [15], the filoviral L is hypothetically shown as a ring with an appendage formed by three globular domains.

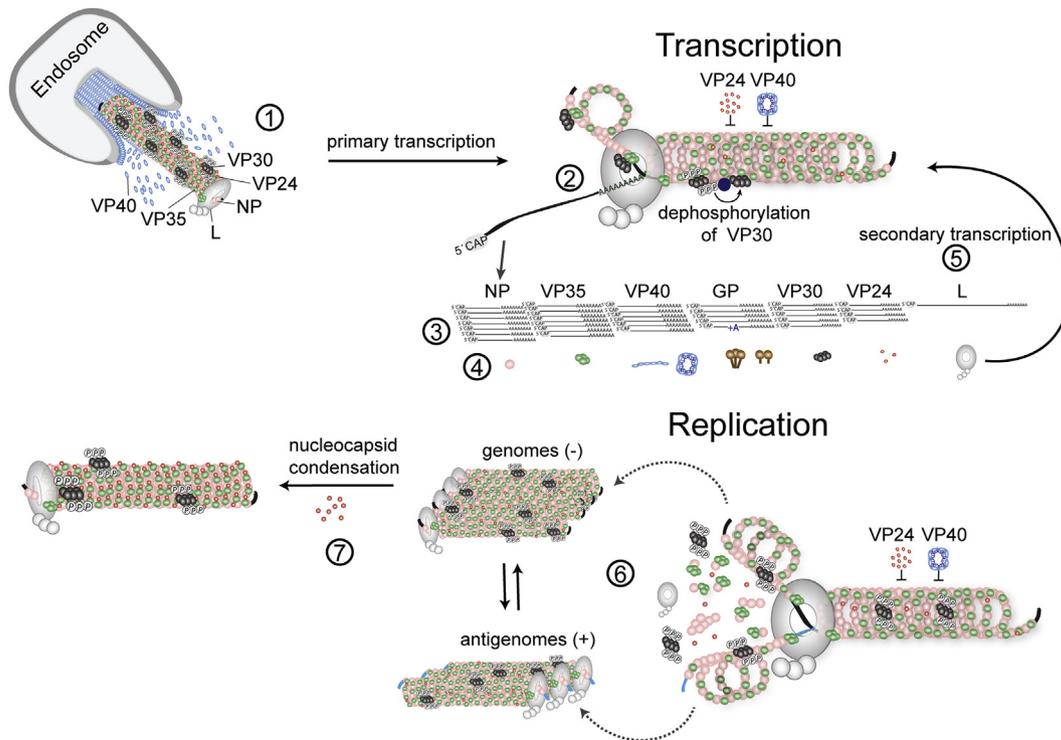


Fig. 3. Hypothetical model of Ebola virus genome replication and transcription. (1) EBOV nucleocapsids are released into the cytoplasm of the infected cells after fusion of the viral and the endosomal membranes. The incoming nucleocapsids consist of the NP–RNA complex (NP, pink spheres), VP35 (green spheres), L (light gray ring with appendices), VP24 (red spheres), and contain mostly phosphorylated EBOV VP30 (dark gray spheres). Based on the structure of the vesicular stomatitis virus L protein [15], EBOV L is hypothetically shown as a ring with an appendage formed by three globular domains. (2) The nucleocapsids are released into the cytoplasm of the infected cells where primary transcription occurs. The viral RNA-dependent RNA polymerase formed by the catalytic polymerase subunit L and the cofactor VP35 initiates mRNA synthesis. VP35 binds to NP and redirects L to the NP–RNA complex. VP30 acts as a transcription activator and must be dephosphorylated to be functional. Dephosphorylation of VP30 is carried out by cellular phosphatases (dark blue sphere) interacting with the nucleocapsid. Since VP24 condensates the nucleocapsids, we hypothesize that transcription-competent nucleocapsids contain no or only small amounts of VP24. (3) L caps the 5' ends of the mRNAs and methylates the caps. Polyadenylation of the nascent mRNAs occurs through polymerase stuttering. The GP mRNA is cotranscriptionally edited by polymerase stuttering as indicated by the “+ A” in the GP transcript. Dissociation of the viral polymerase at the gene borders results in a transcript gradient. (4) The viral mRNAs are translated by the cellular translation machinery. (5) This enables secondary transcription that requires newly synthesized viral proteins. (6) Phosphorylation of VP30 diminishes its function in viral transcription and helps to initiate viral genome replication. VP40 (blue) and VP24 inhibit viral RNA synthesis. VP35 serves as a chaperone to prevent NP from aggregation. Genome replication involves the synthesis of a positive sense antigenome (blue line) that serves as a replication intermediate. (7) VP24 is involved in nucleocapsid condensation and shifts the viral replication cycle towards assembly and budding. Binding of phosphorylated VP30 to NP ensures packaging of VP30 into virions. Fully assembled nucleocapsids are transported by an actin-dependent transport mechanism to the budding sites. Note that there are significant differences to Marburg virus transcription as explained in the text.

transcription start signal of the first gene. It then proceeds along the RNA template by stopping and reinitiating at each gene junction, resulting in the synthesis of mainly monocistronic mRNAs [38–40] (Fig. 3). Viral transcription includes the addition of a 5' cap and a poly(A) tail to the nascent mRNAs by the viral polymerase. The viral mRNAs are then translated by the cellular translation machinery. Viral protein synthesis is the prerequisite for both replication to amplify the viral genomes and secondary transcription to produce more viral mRNA [18] (Fig. 3). During genome replication, the viral poly-

merase synthesizes a full-length complementary copy of the negative sense RNA genome, the positive sense antigenome, which in turn serves as the template for the production of more genomes (Fig. 3) [40,41]. Similar to the genome, the antigenome is encapsidated by the nucleocapsid proteins. The antigenome is a mere replication intermediate—only genomes serve as templates for transcription and are packaged into viral particles. Viral RNA synthesis and nucleocapsid assembly take place within viral inclusions that are formed in the cytoplasm of the infected cell [18,37,42–44].

Mature marburg- and ebolavirus nucleocapsids are transported from the inclusions to the sites of budding at the plasma membrane via an actin-dependent mechanism [45,46]. For EBOV, it has been shown that three proteins—NP, VP35, and VP24—are sufficient to mediate this process [24]. The final step of the viral replication cycle, budding of the assembled virions, is mainly driven by matrix protein VP40 via hijacking components of the cellular endosomal sorting complex required for transport (ESCRT). During this process, GP_{1,2} is targeted to VP40-enriched membranes that enwrap the arriving nucleocapsids in a directional manner (reviewed in Ref. [18]).

This review focuses on filovirus genome replication and transcription and does not cover entry, assembly, and budding. Excellent overviews of these steps of the viral replication cycle are provided in Refs. [16–18]. Despite following the general strategies employed by NNS RNA viruses to replicate and transcribe their genomes, filoviruses have evolved many unique features that will be described in this review. We will also summarize what is known about genome replication and transcription of the newly discovered filoviruses.

Tools to Study Filovirus Replication and Transcription

Minigenomes

Transfection-based minigenome systems for NNS RNA viruses are useful tools to identify and characterize both cis-acting signals in the viral RNA genome and protein factors required for genome replication and transcription (reviewed in Refs. [47,48]). The first filovirus minigenome systems were established for MARV and EBOV in 1998 and 1999, respectively, and for Reston virus (RESTV), they were established in 2005 [49–52]. More recently, chimeric minigenome systems have been created for the novel filoviruses LLOV [53,54] and MLAV [7] (see the [Replication and transcription of novel filoviruses](#) section for more information on these minigenomes). Currently, the development of minigenome systems for the fish filoviruses is hindered by the significant lack of terminal genomic sequences. Since work with filoviruses requires the highest biosafety level (BSL-4), minigenome systems, which enable molecular biological studies under low biosafety level conditions, have been crucial for the analysis of the replication cycle of these viruses.

Filoviral minigenomes are truncated versions of the viral genomes and typically contain the 3' and 5' genome ends flanking a reporter gene. The genome ends contain all regulatory elements required for encapsidation, replication, and transcription initiation.

The reporter gene is additionally flanked by the viral transcription start and stop signals that govern the initiation and termination of mRNA synthesis by the viral polymerase [40,41,48,55]. Genetic manipulation of any aspect of the minigenomes, including any regulatory element or the reporter gene, is easy since the minigenome cDNA is contained within a plasmid. Expression plasmids encoding the viral proteins required for transcription and genomic replication are co-transfected along with the minigenome plasmid (Fig. 4a). Inside the cells, the minigenome and the viral genes are transcribed from their plasmids by a DNA-dependent RNA polymerase. In most systems, the minigenomes are under the control of the T7 RNA polymerase promoter (reviewed in Refs. [48,55]). RNA polymerase I- and II-driven minigenome systems have also been described [52,56]. It is of utmost importance that the minigenomes have the correct genome ends after being transcribed by the DNA-dependent RNA polymerase. While some DNA-dependent RNA polymerases, including the T7 RNA polymerase and RNA polymerase I, start transcription at a defined position and generate transcripts with precise 5' ends, other polymerases, including RNA polymerase II, do not. Generation of precise 3' ends is generally difficult. To overcome these issues, the minigenomes are flanked by ribozyme sequences in the plasmid backbone. Commonly, the hepatitis delta ribozyme is used to generate 3' ends and a hammerhead ribozyme is used to generate defined 5' ends, if required [57–59]. Transcription and replication of the minigenome by the viral polymerase complex assembled in the transfected cells lead to the production of replicated and transcribed minigenome RNA, and reporter gene expression (Fig. 4a). The generated RNA products can be analyzed and quantified by Northern blot analysis, primer extension analysis (a method to analyze initiation of RNA synthesis [60]), and RT-PCR (reviewed in Refs. [40,41,48,55]). Minigenome systems have been instrumental in identifying the viral proteins required for genome replication and transcription, studying the function of these proteins, and analyzing regulatory elements on the genome as described below in more detail. An in-depth description of technical tips for the use of filovirus minigenome systems can be found at <https://blog.addgene.org/minigenomes-a-safe-way-to-study-dangerous-viruses>.

Filoviral minigenomes typically contain a single gene (monocistronic minigenomes), but minigenomes with several genes encoding distinct reporter proteins have also been used. In these minigenomes, the genes are separated by filovirus-specific gene borders [61–65]. Co-transfection of expression plasmids encoding VP40, GP, and VP24 in addition to plasmids encoding the minigenome and L, NP, VP35, and VP30 results in the assembly and budding of infectious virus-like particles (iVLPs),

allowing the study of additional steps in the viral replication cycle, including virus assembly, budding, entry, and primary transcription [26,66,67] (Fig. 4b). A variant of this approach, dubbed the transcription- and replication-competent virus-like particle (trVLP) system, encodes the VP40, GP, and VP24 genes in addition to a reporter gene within a tetracistronic minigenome, obviating the need to co-transfect additional plasmids encoding these genes [68].

Various minigenome-based systems have been developed to investigate different aspects of the filoviral replication cycle (reviewed in Refs. [48,55]). The main drawback of the minigenomes is their considerably shorter length compared to the full viral genomes, which might influence different steps of the viral replication cycle, including virus entry, RNA encapsidation, RNA elongation, assembly, and budding. In addition, the viral protein concentrations and ratios in transfected cells are different from those in infected cells, which might affect the results obtained with minigenome, iVLP, and trVLP systems. This highlights the need to confirm results obtained from these reductionist minigenome-based systems using authentic viral infection.

Recombinant viruses

Recombinant filoviruses can be generated by transfecting cells with a plasmid containing the full-length cDNA of the viral genome in the antigenomic (+) sense orientation along with plasmids encoding L, NP, VP35, and VP30 [55,69] (Fig. 4c). Similar to the minigenome system, the viral antigenomic RNA is transcribed from the plasmid in the transfected cells, and the nucleocapsid and polymerase components are expressed. If the antigenomic RNA is recognized as a template for replication by the viral proteins, viral genomic RNA is produced than can then be replicated and transcribed, leading to the production of all viral proteins, and eventually to viral particle formation [47,70]. Using these virus rescue systems, it is possible to introduce mutations into viral genes or cis-acting elements and to insert additional transcription units into the viral genome, including genes encoding fluorescent proteins. To

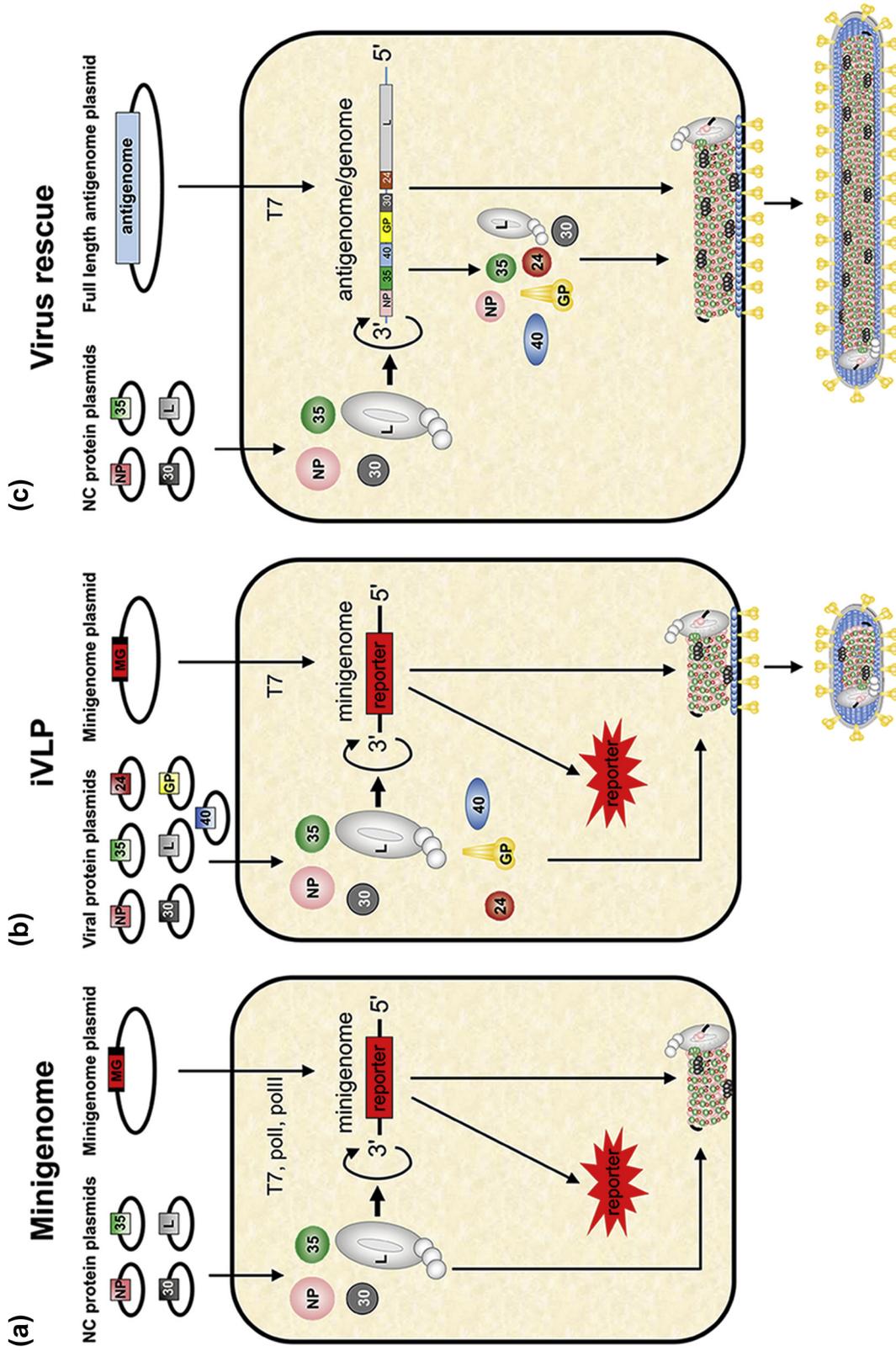
date, rescue systems have been developed for EBOV, MARV, and RESTV (first descriptions: [71–73]). A thorough overview of the generation and application of recombinant filoviruses is beyond the scope of this article and has been reviewed elsewhere [55,74].

Filovirus genomes

Filovirus genomes consist of a single negative sense RNA molecule that contains a variable number of monocistronic genes in a linear order. The genes are flanked by conserved gene start (GS) and gene end (GE) signals that either overlap or are separated by short intergenic regions (IRs). Located at the filovirus genome ends are the 3' leader (Le(-)) and the 5' trailer (Tr(-)). The Le(-) is defined as the sequence upstream of the GS signal of the first gene and the Tr(-) as the sequence downstream of the GE signal of the last gene (Fig. 1b). Both the genomic Le(-) and the antigenomic Tr(+) contain important regulatory elements for replication and transcription as described in the [Filoviral replication promoters](#) and [Cis-Acting Elements Involved in Filovirus Transcription](#) sections.

Until recently, filovirus genomes were believed to be significantly longer than the genomes of other NNS RNA viruses (around 19 kb). While this is still true for the mammalian filoviruses, the genomes of the fish filoviruses are considerably shorter (about 17 and 15 kb; Fig. 1b). This genetic diversity is also reflected in the gene number. Ebola- and marburgvirus genomes contain seven genes in a linear order (3' NP-VP35-VP40-GP-VP30-VP24-L) (Fig. 1b). This gene arrangement is also found with LLOV and MLAV [5,7], whereas the XILV and HUVJ genomes are different, with the XILV genome containing 10 open reading frames (ORFs) and the HUVJ containing 6 [3,9] (Fig. 1b). The XILV genome was initially reported to encode NP, VP40, VP30, and L homologs, two putative glycoproteins, and four proteins with unknown functions [3,9]. There were no reported homologs for VP35 or VP24 encoded in the XILV genome. HUVJ, on the other hand, was

Fig. 4. Tools to study filovirus replication and transcription. (a) Filovirus minigenome systems involve transfecting cells with expression plasmids encoding the minigenome and the filovirus proteins required for transcription and replication. In most systems, the minigenome is transcribed by ectopically expressed T7 RNA polymerase (T7). There are also minigenome systems in which the minigenome is transcribed by RNA polymerase I (polI) or RNA polymerase II (polII). If the expressed nucleocapsid proteins (NP, VP35, L, and VP30) accept the minigenome RNA as a template for replication and transcription, this can be monitored by reporter gene expression. (b) Supplementing the minigenome system with expression plasmids encoding GP, VP40, and VP24 results in the release of single-round infectious virus-like particles (iVLPs) which consist of minigenome nucleocapsids which are packaged and released from the cells similar to viral particles. (c) Filovirus rescue systems involve transfecting cells with a plasmid encoding the full-length viral antigenome along with expression plasmids encoding the nucleocapsid proteins. The nucleocapsid proteins will drive replication of the antigenome, leading to the production of viral genomes. The viral genomes will then be replicated and transcribed, which results in the expression of all viral proteins and production of more genomes. Virus rescue will proceed similarly to infection, with infectious particle assembly, egress, and release.



reported to encode NP, GP, and L homologs, with no clear homologs to VP35, VP40, VP30, or VP24. Thus, three of the six HUJV genes were reported to encode proteins of unknown function [3,9] (Fig. 1b). Additional analysis of the XILV and HUJV ORFs with previously unassigned homologs and potential functions is presented in the [Putative functions of unassigned XILV and HUJV proteins](#) section. These analyses strongly suggest that ORF2 of the XILV genome encodes a VP35 homolog (Fig. 1b).

Filoviral replication promoters

The filoviral genomic and antigenomic replication promoters are located at the 3' end of the genome and antigenome, respectively [60,75–78]. Serial passaging of EBOV-infected cells led to the generation of defective interfering particles that contained 155 nucleotides of the 3' end of the genome fused to 176 nucleotides of the 5' end, indicating that the genomic and antigenomic replication promoters are contained in these regions [79]. NNS RNA viral polymerases enter the genomes and antigenomes exclusively at the 3' promoter regions. It is not understood how the filoviral polymerase recognizes the promoter and gets access to the RNA template. However, as described in the [Filovirus replication and transcription](#) section, only viral RNA that is complexed with NP and other nucleocapsid proteins is accepted as a template for replication and transcription. It has been shown for other NNS RNA viruses that purified viral polymerase is capable of initiating replication in the absence of NP using short RNA templates [80–83]. Based on these observations, it is conceivable that encapsidation of the extreme 3' ends of filoviral genomes and antigenomes is not required for initiation of RNA synthesis, whereas efficient RNA elongation requires a properly packaged NP–RNA template.

Genomic replication promoter

The filoviral genomic replication promoters are not well studied—there is only experimental data available for EBOV and MARV. Both the MARV and EBOV genomic replication promoters are bipartite in nature [75,76]. The first promoter element (PE) resides within the Le(–), which is 55 nucleotides in length for EBOV and 48 nucleotides for MARV, and is separated from the second PE by a short spacer region. For EBOV, the spacer consists of the GS signal of the first gene and the following 12 nucleotides (Fig. 5a). In the case of MARV, the two promoter elements are separated by the GS signal only (Fig. 5a). Whereas the sequence of the spacer is irrelevant, its length has to be divisible by six to allow efficient replication [75,76]. The first PE has not been thoroughly mapped for any filovirus. Limited

mutational analysis of the EBOV Le(–) showed that clusters of mutations affecting nucleotides 44–55 decreased replication activity in a minigenome system to non-detectable levels, indicating that this part of the Le(–) is important for replication [75]. Computer predictions suggest that marburgvirus, ebolavirus, and LLOV Le(–) regions adopt an internal hairpin structure (Fig. 5b). Since the first 13–23 nucleotides of the ebola- and marburgvirus genome termini show a high degree of complementarity, the Le(–) has also the potential to form a panhandle structure with complementary sequences in the Tr(–) region [84,85]. Formation of the internal leader hairpin was experimentally confirmed using *in vitro* transcribed MARV and EBOV Le(+) RNAs. For technical reasons, the analyzed RNAs were in positive sense orientation and not encapsidated with NP [75,76]. Panhandle structure formation between the Le(–) and Tr(–) was observed using *in vitro* transcribed EBOV negative sense minigenome RNA. These studies were also performed with non-encapsidated RNA [85]. Mutations in the Le(–) hairpin region that interfered with stable secondary structure formation did not dramatically affect replication efficiency of EBOV minigenomes [84]. Also, clustered compensatory mutations of nucleotides 10–13 that changed the primary sequence but restored the secondary structure eliminated minigenome replication activity, indicating that secondary structure formation *per se* is not sufficient to support replication [75]. Together, these data suggest that formation of the internal Le(–) hairpin structure is not a prerequisite for efficient replication. Mutations in the Tr(–) region that affected panhandle structure formation with the Le(–) led to a decrease in minigenome activity and were prohibitive for viral rescue, indicating that the respective nucleotides in the Tr(–) are required for efficient replication [85]. Since the mutated nucleotides are also part of the antigenomic replication promoter located in the Tr(+) region of the antigenomes (see below, *antigenomic replication promoter*) and are involved in the interaction with cellular host factors that might promote viral replication [85], interpretation of these data is difficult.

The putative internal Le(–) hairpin could also serve a different purpose. Recent studies of the replication initiation mechanisms of EBOV, Sudan virus (SUDV), and RESTV revealed that regardless of the identity of the 3' terminal nucleotide(s), the viral polymerase initiates replication accurately opposite the first C residue of a highly conserved motif (3' CCUGUG), which is typically located at position 2 of the genomes and antigenomes, respectively [60] (Fig. 5a). Even if additional nucleotides are added to the 3' end of the CCUGUG motif, the polymerase still initiates replication internally opposite the first C residue of this motif. If this C

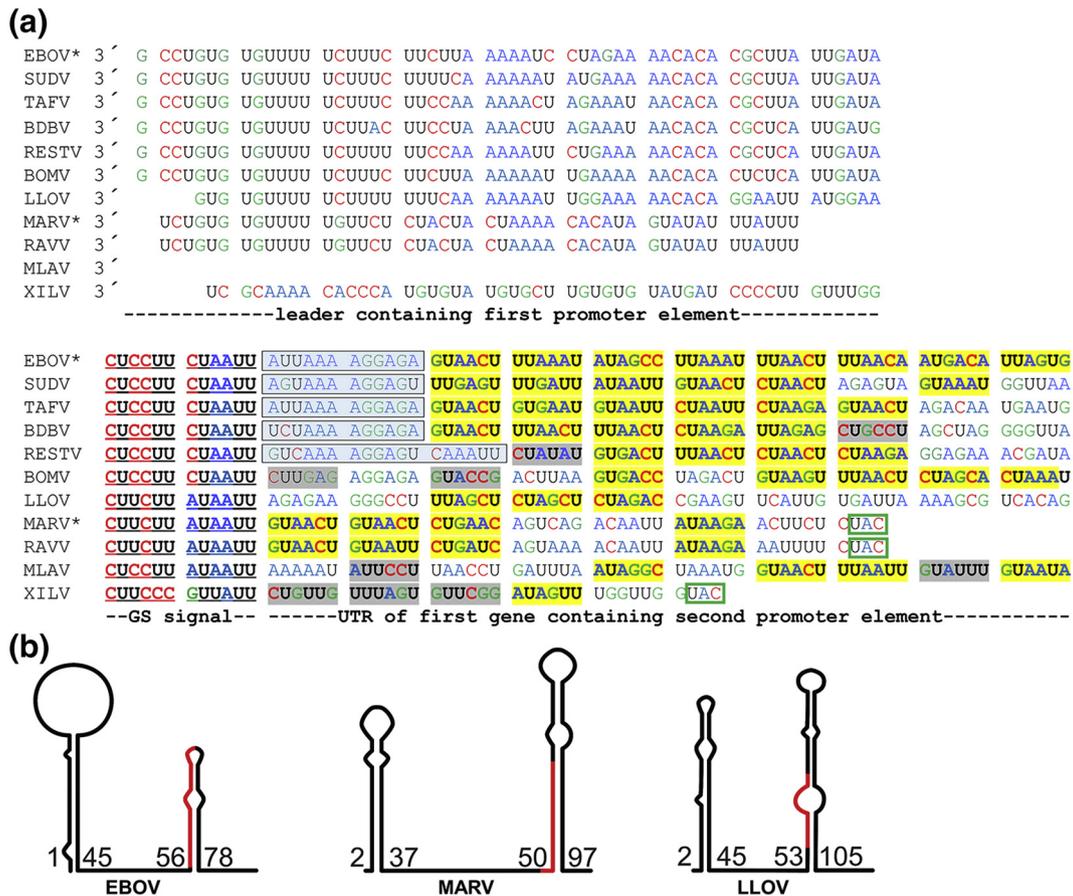


Fig. 5. Filovirus genomic replication promoters. (a) An alignment of the leader regions, initial gene start (GS) signals, and the untranslated regions (UTRs) of the first gene of filovirus genomes was performed manually. GS signals are indicated in bold and underlined. Spacer regions in the ebolavirus promoters, located between the GS signal and the second promoter element, are indicated by outlined gray boxes. Hexamers in the second promoter element containing a uridine residue in the second position followed by two purines are indicated in yellow, while other hexamers containing a uridine in the second position are indicated in gray. Translation start anticodons are indicated by green boxes. The asterisks indicate experimentally determined replication promoter structures. (b) RNA secondary structures for the 3' ends of the EBOV, MARV, and LLOV genomes, predicted using the RNAfold web server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>), are illustrated in negative sense RNA orientation. Numbers indicate the first and last nucleotide of the stem-loop structures. The LLOV sequence used in this analysis was appended with the four 3' terminal nucleotides of the EBOV sequence (3' GCCU). The second stem-loop structure is formed by the GS signal of the first gene, the NP gene, by base-pairing with downstream sequences. The location of the NP GS sequence in the stem-loop is marked in red. Abbreviations of virus names are explained in the legend to Fig. 1.

residue is replaced with a U residue (which reflects the 3' end of the MARV genome), replication is not initiated [60]. This unusual replication initiation mechanism leads to heterogeneous ebolavirus genome ends with a 3' overhang that is not templated by the complementary strand. The 3' overhang is usually a single G or A residue, but a small fraction of genomes and antigenomes has extended 3' ends [60,86]. A possible mechanism for the addition of the non-templated 3' terminal nucleotide(s) is back-priming in which the 3' end of the viral RNA folds into a hairpin structure and the internal RNA sequence is elongated by the viral polymerase [60]. So it is conceivable that the predicted internal Le(-) hairpin is utilized by the

viral polymerase to add the non-templated 3' nucleotide(s). This unusual heterogeneity of ebolavirus genome ends might help to avoid detection of viral RNA by cellular RNA sensors, such as RIG-I-like receptors, whose binding to double-stranded RNA is impaired by 3' overhangs [87,88].

The second PE of the EBOV and MARV replication promoters resides within the 3' untranslated region (UTR; negative sense orientation) of the NP gene and consists of a stretch of nucleotide hexamers, three for MARV and eight for EBOV, in which every sixth nucleotide is a uridine residue, usually followed by two purine residues (NURRN₂) (Fig. 5a). The minimal requirement for EBOV and MARV replication is a stretch of three of these

hexamers with the EBOV promoter gaining strength when more hexamers are present [75,76]. This unusual replication promoter structure is conserved among all marburg- and ebolaviruses and is also found at the 3' end of the LLOV genome (Fig. 5a). Of note, a NURRN₂ hexamer is also part of all filovirus GS signals with the exception of the fish filoviruses.

Thorough analysis of the promoter structures is not yet possible for some of the more recently discovered filoviruses, including MLAV, XILV, and HUJV, because of the lack of significant sequence information at both ends of their genomes [7,9]. The LLOV Tr(-) region is also lacking significant sequence information [5]. The Le(-) and Tr(-) sequences of MLAV have not yet been published, but the first GS signal has a filovirus-typical structure, and the UTR of the MLAV NP gene contains a putative second PE (Fig. 5a). Although it appears that the GS signal for the first gene is present in the XILV genome, poor conservation of the the Le(-) and second PE leaves many questions (Fig. 5a). Does XILV have an extended Le(-) sequence? Or is the XILV Le(-) merely poorly conserved? Why is the proposed second PE so different? For this last question, there remains an intriguing possible explanation. When looking at the second PE of XILV, it contains only a single NURRN₂ motif (Fig. 5a). It is noteworthy, however, that XILV has a somewhat divergent GS signal, which does not include the NURRN₂ sequence. Thus, it is conceivable that the XILV second PE contains other sequences that serve as elements of the viral replication promoter.

The only other NNS RNA virus family with a bipartite promoter structure are the paramyxoviruses [89–91]. In contrast to filovirus genomes, the length of paramyxovirus genomes is a multiple of six. This hexameric genome length results from each NP protomer in the NP–RNA complex interacting with exactly six nucleotides. Consequently, each nucleotide has a defined position within the binding groove of the NP protomers [91,92]. This ability to precisely position the nucleotides in the NP–RNA complex determines the paramyxovirus promoter structure. The first PE has a minimum length of 12 nucleotides and is positioned at NP protomers 1 and 2 (starting from the 3' end). The second promoter element is defined by the presence of one or two conserved nucleotides (usually C or GC) at the same position in three consecutive NP protomers. Proper spacing of the two PEs places them at the same vertical face of the NP–RNA helix [92].

Although marburg- and ebolavirus genome lengths are not a multiple of six, six seems to be a magic number for these viruses. The bipartite replication promoters follow the “rule of six”, and similar to the paramyxovirus nucleocapsids, each MARV and EBOV NP protomer interacts with six nucleotides in the NP–RNA complex (more detail in the [Filovirus](#)

[replication and transcription](#) section). The length of the MARV Le(-) is divisible by six (48 nts) as is the length of the filoviral GS signals (12 nts). Although the ebolavirus Le(-) is not divisible by 6 (55 nts), the unusual replication initiation strategy excludes the first nucleotide from the promoter, making the part of the Le(-) region that is accessed by the polymerase divisible by six (54 nts) (Fig. 5a) [60].

Antigenomic replication promoter

The filoviral antigenomic replication promoters are poorly defined. Per definition, they are located at the 3' end of the antigenomes and are complementary to the 5' ends of the genomes. Sequence comparison suggests that the overall structure of the antigenomic and genomic replication promoters is similar [75,76]. Indeed, EBOV defective interfering copy-back type particles and copy-back MARV minigenomes in which the leader region was replaced with a complementary copy of the trailer were readily replicated [49,79]. Removal of the first 25 nucleotides of the EBOV Tr(+) region (i.e., the last 25 nucleotides of the genome) reduced replication activity of negative sense minigenomes to a single round, confirming that these nucleotides are an essential part of the antigenomic promoter [60,78]. The single-round replicating minigenomes were still able to initiate replication and transcription from the genomic promoter, indicating that the trailer region is not required for these processes. Rescue of full-length EBOV clones lacking the last 56 nucleotides of the genome was not successful, confirming the importance of this region for viral replication [85]. Similar to the Le(-) regions, the Tr(+) regions are also predicted to form RNA secondary structures, resulting either in internal Tr(+) hairpins or in panhandle structures formed with complementary regions in the Le(+) [84,85,93,94]. It is likely that the filoviral antigenomic replication promoters are also bipartite in nature as the 3' ends of the marburg- and ebolavirus antigenomes contain clusters of hexameric NURRN₂ motifs downstream of Tr(+) [75].

An important difference between the genomic and antigenomic promoters is that the antigenomic promoters do not drive transcription. Studies on the replication promoters of other NNS RNA viruses revealed that the antigenomic promoters are stronger than the genomic promoters, leading to more efficient amplification of the genomic RNA during replication. This difference in promoter strength can be attributed to two mechanisms: first, nucleotide differences between the genomic and antigenomic promoter regions, and second, competing replication and transcription activity at the genomic promoter [95–98]. There are currently no comparative studies on the relative strength of the filoviral genomic and antigenomic promoters. It has been shown, though,

that transcription-deficient MARV minigenomes replicate more efficiently than transcription-competent minigenomes, supporting the idea of competing replication and transcription activity at the genomic promoter [76].

Cis-Acting Elements Involved in Filovirus Transcription

Transcription promoter

The structure of the filoviral transcription promoters has not been elucidated yet, but analogous to other NNS RNA viruses, it must be located in the Le(-) region, upstream of the first GS signal. Indeed, MARV copy-back minigenomes, in which the leader region including the first GS signal has been replaced by a complementary copy of the trailer, are replicated but not transcribed [49]. Similarly, deletions in the trailer region abolish antigenome replication but do not affect transcription [77,78]. Since there is a single entry site for the polymerase at the 3' end of the genome, the polymerase cannot initiate transcription internally if it falls off the template. For this reason, the promoter-proximal genes are more frequently transcribed than the promoter-distal genes, leading to a transcription gradient (Fig. 3) [99,100].

GS signals

With the possible exception of a single LLOV bicistronic gene, each filovirus gene is flanked by conserved GS and GE signals (Figs. 1b and 6a). After the polymerase has accessed the NP-RNA template at the 3' end, it scans the genome sequence for the first GS signal to initiate transcription of the first gene. The first nucleotide of all filoviral GS signals is a cytidine residue, and consequently, all filoviral mRNAs start with a guanosine (Fig. 6a). This seems to be important for mRNA capping (see the [Transcription-specific functions of filoviral proteins](#) and [Transcription-specific functions of L sections](#)). The 12-nucleotide stretch that comprises the GS signals of most of the filoviruses is fairly well conserved (Fig. 6a). A unique feature of filovirus GS signals is that they are predicted to adopt RNA hairpin structures [99,101]. With the exception of the hairpin formed by the GS signal of the EBOV NP gene (Fig. 5a; see the [Transcription-specific functions of filoviral proteins](#) and [The role of VP30 in filovirus transcription sections](#)), it remains to be determined if these RNA secondary structures are of functional relevance.

The consensus sequence of the GS signal for ebolaviruses, marburgviruses, LLOV, and MLAV is 3' C₁UNCUUN₇UAAUU₁₂. The main difference between these sequences is that ebolaviruses have a preference for a cytidine residue at position 7 in this motif, whereas the marburgviruses, LLOV, and MLAV prefer a purine residue (Fig. 6a). Despite significant

differences in sequence and genome organization, XILV has a fairly similar GS sequence of 3' C₁UCCCCG₇UAAUU₁₂. In contrast, nothing resembling the conserved filovirus GS signals, including the XILV GS signals, is present upstream any of the HUVV ORFs. Computer-assisted search of the HUVV noncoding regions for conserved motifs identified a highly conserved 14-nucleotide sequence upstream of the NP ORF, ORF3, and the GP ORF (3' C₁GGGUUNUANUAU₁₄; Fig. 6a). Somewhat similar sequences are observed upstream of ORF5 and the L ORF, although differences are noteworthy. A lack of this motif upstream of ORF1 could be due to missing sequence at the 3' end of the genome.

GE signals

After transcription initiation and capping, the polymerase moves along the template and elongates the mRNA until it reaches the first GE signal. The GE signals are highly conserved across the different filoviral genera and end with a stretch of 5–6 uridine residues (Fig. 6a). These induce stuttering of the polymerase, leading to the addition of a poly(A) tail at the 3' end of the mRNAs [102]. There are few variations in the filoviral GE signals, including in the highly divergent fish filoviruses (Fig. 6a).

Gene borders and IRs

Filovirus transcription follows the typical stop-start mechanism of NNS RNA viruses. After terminating transcription at the GE signal of the upstream gene, the polymerase scans for the GS signal of the next gene. The EBOV polymerase is able to scan for GS signals in both directions of the GE signal [64]. Most of the filoviral GS and GE signals contain a conserved 3' UAAUU pentamer, allowing for gene overlaps in which the 3' UAAUU sequence is shared by the GS signal of the downstream gene and the GE signal of the upstream gene, requiring the polymerase to scan slightly upstream of the GE signal for the next GS signal (Fig. 6a and b). For XILV, this conserved sequence is one nucleotide different; a 3' UUAUU pentamer is conserved in all GS and GE signals. Although the HUVV GE signals contain a conserved UAAUU pentamer, this sequence is not present in the putative HUVV GS signals, suggesting that typical filoviral gene overlaps cannot occur (Fig. 6a and b). The gene overlaps are as efficient in transcription initiation and termination as the GS and GE signals that are separated by IRs [64]. While the marburg- and ebolavirus genomes contain one to three gene overlaps, the MLAV and LLOV genomes contain four, and XILV takes it to the extreme with all genes overlapping (Fig. 6b). It seems that the VP24 and L ORFs of LLOV are not separated by GE and GS signals, indicating that

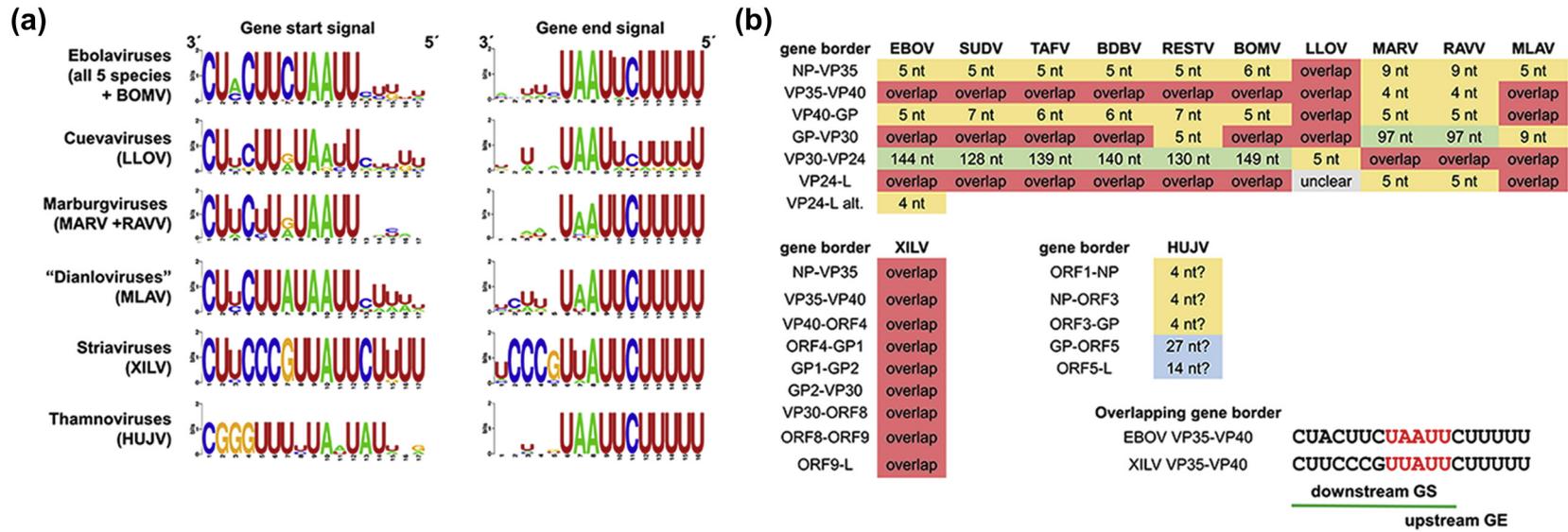


Fig. 6. Filovirus gene start and end signals, and gene borders. (a) Alignments of all gene start (GS) and gene end (GE) sequences of the indicated viruses were performed using Clustal Omega. Logo plots were created using WebLogo version 2.8.2 (<https://weblogo.berkeley.edu/logo.cgi>). Sequences are shown in negative sense orientation, with five additional nucleotides downstream and upstream of the GS and GE signals, respectively. For the fish filoviruses, GS and GE borders we predicted based on homology to corresponding sequences in the other filoviruses, with the exception of the HUJV GS signals which were predicted by identifying a similar sequence found upstream multiple genes. (b) Gene borders in filovirus genomes. Gene borders with overlapping GS-GE sequences are indicated with “overlap” (in red), gene borders with intergenic regions (IRs) are indicated with the length of the IRs (short IRs in orange, intermediate IRs in blue, and long IRs in green), and gene borders with unclear or possibly lacking GS or GE sequences are indicated with “unclear” (gray). The EBOV VP24 gene has two GE signals, resulting in two different gene borders (VP24-L and VP24-L alt.). As the GS sequences for HUJV are speculative and divergent from the other filoviruses, the proposed gene borders are indicated with the length of the IR colored as the other IRs with a question mark. Examples of overlapping gene borders are also shown. Filovirus overlapping gene borders are all structured similarly: these overlaps begin with the GS signal of the downstream gene (green line), followed by an overlap region consisting of the end of the GS signal and the beginning of the GE signal of the upstream gene (overlap sequence colored in red), and are followed by the end of the GE signal (red line). The EBOV VP35-VP40 gene border is shown as a representative of mammalian filovirus gene border overlaps. The XILV VP35-VP40 gene border is shown as a representative of XILV gene border overlaps. Abbreviations of virus names are explained in the legend to Fig. 1.

VP24 and L are produced from a single mRNA (Fig. 6b) [5]. The mechanism by which such a bicistronic transcript could serve as a template for the expression of both genes remains unclear.

While some of the filovirus genes overlap, others are separated by diverse IRs, which vary both in length and in sequence. Except a single long IR in each of the ebola- and marburgvirus genomes (97–149 nts), the IRs are quite short (4–9 nts) (Fig. 6b). The non-overlapping gene borders for LLOV and MLAV are also short, with no long IRs (Fig. 6b). Based on the proposed GS signals in the HUJV genome, there would be no overlapping genes. All genes would be separated by IRs of short (4 nts) and intermediate (14 and 27 nts) length IRs (Fig. 6b).

Mutational analysis of the EBOV IRs indicated that they are involved in transcription regulation. However, the main regulators of transcript levels are the UTRs of the viral genes [63–65,103]. If the viral polymerase fails to recognize the GS and GE signals, this leads to the production of readthrough mRNAs, which are present in EBOV- and MARV-infected cells [65,100]. Readthrough transcription occurs independently of the structure of the gene borders. The sequential transcription of the viral genes is highly regulated. If the polymerase is unable to terminate transcription of a gene because the GE signal is destroyed, it cannot reinitiate transcription at the following gene, even if the GS signal of this gene is intact [64].

UTRs

The ORFs of filoviral genes, particularly the mammalian filovirus genes, are flanked by unusually long UTRs (Fig. 1b). The upstream UTRs of the EBOV, LLOV, MARV, and MLAV genes are on average 214, 275, 131, and 157 nts in length, and the average lengths of the downstream UTRs are 298, 372, 476, and 399 nts, respectively. In analogy to cellular mRNAs, it is conceivable that the filoviral UTRs play a regulatory role at mRNA level, for example, in mRNA stabilization and translation. Indeed, the 5' UTR of the EBOV L mRNA encodes a small, upstream ORF (uORF) that overlaps with the L ORF. This uORF controls L translation by diminishing ribosome initiation at the L AUG start codon. However, in the event of cellular stress, when translation of capped mRNA is impaired, the uORF is required to maintain translation of the L mRNA [103]. Meanwhile, the UTRs of the fish filoviruses are conspicuously shorter, with the downstream UTRs of XILV ORF8 and HUJV ORF3 as well as the upstream and downstream UTRs of the HUJV GP gene being the exceptions (Fig. 1b).

Filovirus Replication and Transcription

Nucleocapsid

A common feature of all NNS RNA viruses is that naked viral RNA is not accepted as a template for replication and transcription [47]. Only viral genomes and antigenomes that are complexed with the viral nucleocapsid proteins, forming the helical nucleocapsid, are recognized as templates for viral RNA synthesis and, in the case of the genomes, will be packaged into virions. Five of the seven marburg- and ebolavirus structural proteins are associated with the nucleocapsid. These are NP, VP35, VP30, VP24, and L [14,20–22,42,104–107]. An overview of the multiple interactions of these proteins is provided in Ref. [108]. Though not identical, the MARV and EBOV nucleocapsids show striking architectural homology. The core of the nucleocapsids is a left-handed NP–RNA helix (Fig. 7a). VP35 and VP24 are required for the formation of complete nucleocapsids in addition to NP (Fig. 7b, right and c). The mature EBOV nucleocapsids have an average pitch of 74 Å and 11.9 or 12.9 subunits per turn and are slender and longer than the MARV nucleocapsids which have an average pitch of 77 Å and 14.8 or 15.8 subunits per turn [23].

Because viral particles for any of the new members of the filovirus family have not yet been isolated, the virion structures are not known. However, based on gene homology, it is very likely that the composition of the BOMV, LLOV, and MLAV particles is similar to that of the known filoviruses, whereas it is difficult to predict the structures of XILV and HUJV particles because of considerable differences in their gene composition (Fig. 1b and the [Putative functions of unassigned XILV and HUJV proteins](#) section).

Filovirus proteins involved in RNA synthesis

Using minigenome systems that allow the separation of genome replication and transcription, it was shown that three nucleocapsid proteins, NP, VP35, and L, are sufficient to mediate MARV and EBOV replication (Fig. 3). NP encapsidates the viral RNA to form the template for the polymerase complex composed of L, the catalytic subunit of the RNA-dependent RNA polymerase, and its cofactor VP35 [40,49,50,109].

NP

All NNS RNA viruses have a helical nucleocapsid with an NP–RNA core. NP–RNA complex formation is driven by NP self-aggregation [110,111]. This is also true for filoviruses. Both NP aggregation and NP–RNA complex formation are mediated by the

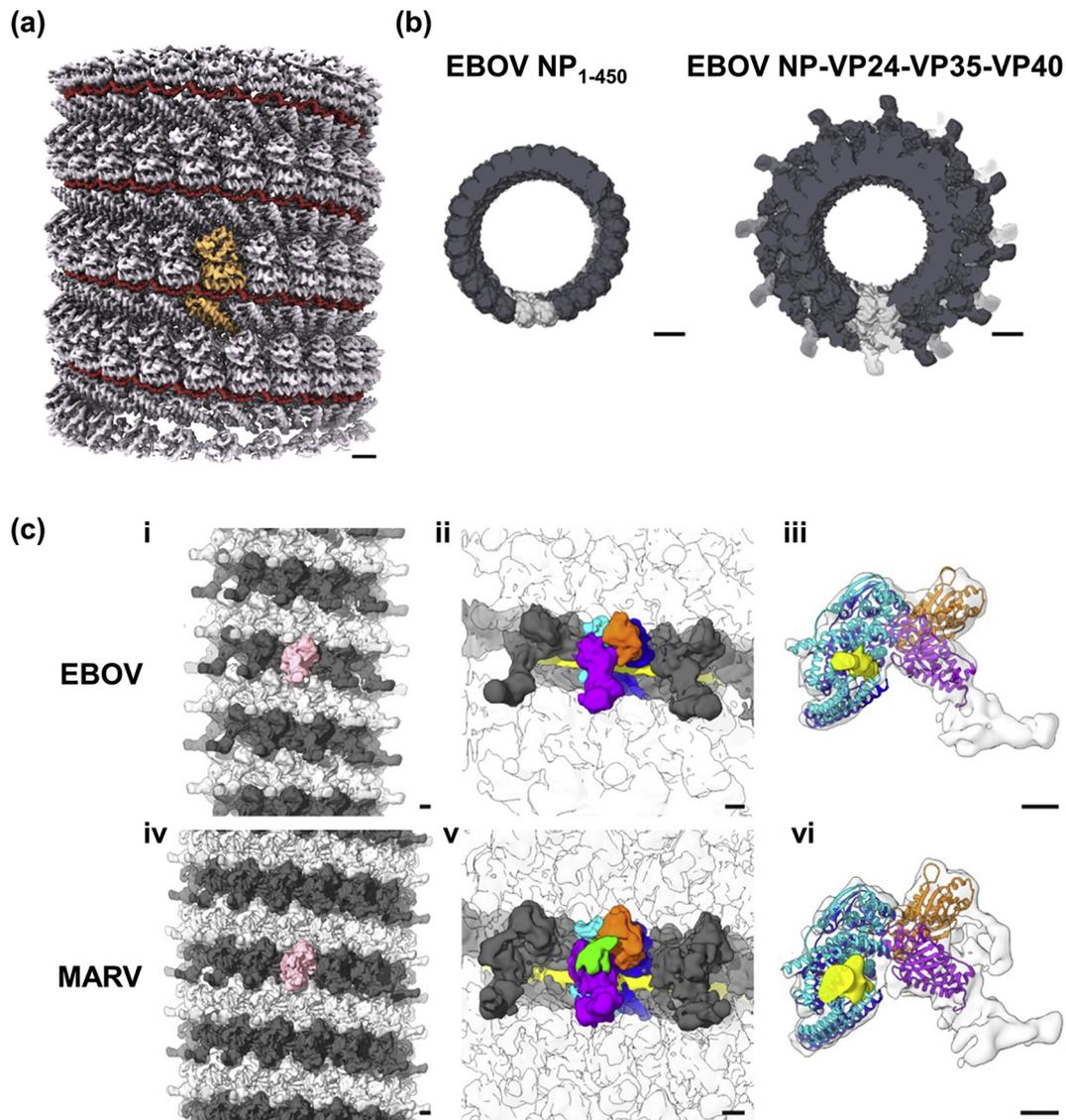


Fig. 7. Cryo-EM reconstructions of the EBOV and MARV nucleocapsids. (a) Cryo-EM reconstruction of the EBOV NP–RNA helix using a truncated version of NP (NP₁₋₄₅₀ + RNA). Single NP molecule in orange, RNA in red. The scale bar represents 20 Å. Reprinted by permission from Springer Nature [28] (<https://www.nature.com/articles/s41586-018-0630-0>). (b) Comparison of cross-sections of EBOV nucleocapsid-like structures. Left, structure of purified EBOV NP₁₋₄₅₀ helices; right, structure of EBOV VLPs produced in cells expressing NP, VP24, VP35, and VP40. In light gray are two nucleocapsid subunits. Scale bars, 50 Å. Adapted from Extended Data Fig. 7 in Ref. [23]. Courtesy of John Briggs, MRC Laboratory of Molecular Biology, Cambridge, UK. (c) Structure of EBOV and MARV nucleocapsids from intact viruses. (i, iv) Visualizations of the EBOV (i) and MARV nucleocapsid helix (iv) as determined by subtomogram averaging. Adjacent rungs are in dark and light gray; a single subunit is highlighted in pink. (ii, v) Subunit densities of EBOV (ii) and MARV nucleocapsids (v). NP densities are in cyan and blue, the small outer protrusions are in orange, and the large outer protrusions are in purple. Green density is extra disordered density in the MARV nucleocapsid. Gray densities show other subunits on the same rung; densities from adjacent rungs are transparent. RNA density is in yellow. (iii, vi), Molecular models of NP and VP24 fitted into the EM densities of EBOV (iii) and MARV nucleocapsid subunits (vi). The two NP models are in blue and cyan while the two VP24 models are shown in orange and purple. Scale bars, 20 Å. Reprinted by permission from Springer Nature [23] (<https://www.nature.com/articles/nature24490>).

amino-terminal parts of EBOV NP (aa 1–450; total length: 739 aa) and MARV NP (aa 1–390; total length: 695 aa) [14,20,22,112–117] (Fig. 7a and b,

left). Crystal structure and high-resolution reconstruction analyses provided a deeper insight into the structure of the EBOV and MARV NP–RNA

complexes [23,28,116–121]. The amino-terminal NP core folds into two lobes. Self-assembly of the NP core leads to the formation of a left-handed helical tube. The RNA cleft is located on the outside of the NP helix between the N- and C-terminal lobes (Fig. 7a). Each NP protomer in the helix interacts with 6 nucleotides [14,22,28]. This interaction occurs through hydrogen bonds and electrostatic interactions between the NP core and the RNA backbone, and is RNA sequence independent [28].

VP35

The NP–RNA complex alone is not a functional template for genome replication. Both EBOV and MARV NPs interact with the polymerase cofactor VP35, which is the homolog to the phosphoproteins (P) of other NNS RNA viruses [116,117,119,120,122]. VP35–NP interaction is not only required for nucleocapsid formation but is also crucial for viral RNA synthesis. Similar to the NNS RNA virus P proteins, VP35 forms homo-oligomers (trimers in the case of MARV, and trimers and tetramers in the case of EBOV) through a coiled-coil motif located in the N-terminal part of the protein [123–126]. Homo-oligomerization is essential for VP35's function in replication and transcription [124,127]. Both EBOV and MARV VP35 proteins interact with the catalytic polymerase subunit L to form the functional polymerase complex [105,109,122,127,128]. It is believed that VP35 serves as a bridge to connect L with the NP–RNA complex [105,128]. VP35 also interacts with NP to keep it in a monomeric, RNA-free state [116,117,119,120]. This chaperon function of VP35 might help to prevent premature and unspecific RNA binding (cellular RNA instead of viral RNA), or it might be used as a mechanism to temporarily unpack the viral RNA template to make it available to the polymerase complex. The NP binding domain within VP35 that facilitates NP chaperoning is conserved among all filoviruses [117] with the exception of XILV and HUJV (Fig. 8b). This is perhaps not surprising given the weak homology of XILV VP35 to other filovirus VP35 proteins and the lack of a clear VP35 homology in HUJV (Figs. 1b and 9). For more information regarding possible VP35 functional homologs in XILV and HUJV, see the [Putative functions of unassigned XILV and HUJV proteins](#) section.

L

The catalytic subunit of the viral polymerase, L, interacts with VP35 to form the functional polymerase complex [49,50,105,122,127,128]. Though associated with the nucleocapsid, L is not required for nucleocapsid formation, but it is indispensable for replication and transcription. The VP35 binding site

resides within the amino-terminal part of L [105,122,127,128]. EBOV L also interacts with VP30 [129], but VP30 is not required for RNA synthesis activity [109]. EBOV L has also the potential to homo-oligomerize [128]. Similar to the L proteins of other NNS RNA viruses, the filovirus L proteins have various enzymatic activities required for RNA synthesis and mRNA processing that are contained within six conserved regions (I–VI) [93,94,130–132]. Cryo-EM structure analysis of the vesicular stomatitis virus (VSV) L showed that the protein is organized in distinct domains: the RNA-dependent RNA polymerase domain (contains regions I–III), the capping domain (contains regions IV and V), the connector domain, the methyltransferase domain (region VI), and a carboxy-terminal domain that is variable among the NNS RNA viruses [15]. The catalytic center of the RNA polymerase domain contains a highly conserved GDN motif. Amino acid exchanges within this motif render the polymerase inactive, as shown for EBOV, RESTV, and MARV L [53,56,109,131,133]. A recently established *in vitro* polymerase assay for EBOV showed that the polymerase uses Mg^{2+} ions as catalytic cofactors for RNA synthesis. The aspartic acid of the GDN motif within the catalytic domain likely coordinates the interaction with the metal ions [109]. The amino acid sequences surrounding the active center are identical in most filovirus L proteins except for the fish viruses, which appear to be more similar to other NNS RNA viruses (Fig. 8d, Catalytic domain (GDN motif)). Another domain of the L protein found in all filoviruses is the KEKE motif (Fig. 8d, template binding), which is involved in positioning and binding of the RNA template [94,130,134], and is also conserved across the filoviruses. All NNS RNA viruses encode L, and it is the best-conserved protein among these viruses [130,135].

Negative regulators of viral RNA synthesis: VP24 and VP40

VP24 is the only filovirus protein for which there is no homolog among the other NNS RNA viruses, and intriguingly, there does not appear to be a VP24 homolog in the fish filoviruses either (Fig. 1b). For a long time, the function of VP24 was obscure, and it was considered to be a minor matrix protein. It is now known that VP24 plays a crucial role in the formation of mature, transport-competent nucleocapsids [24–26,136]. In addition to NP and VP35, VP24 is an essential structural component of the nucleocapsid complex and indispensable for the assembly of rigid and condensed nucleocapsids [20–22,25,137] (Figs. 3 and 7b and c). VP24 is not required for genome replication and transcription, which implies that there is a structural difference between replication-competent nucleocapsids that lack or contain only

small amounts of VP24 and condensed nucleocapsids that contain VP24 [68,138]. A model based on cryo-electron tomography and subtomogram averaging suggests that the NP–RNA helix in a semi-condensed state recruits VP24 and VP24–VP35 to complete condensation [23]. According to this model, the NP:VP24:VP35 stoichiometry in the condensed nucleocapsids is 2:2:1, whereby VP24 and VP24–VP35 form alternating small and large protrusions [14,22,23] (Fig. 7b, right and c).

Overexpression of VP24 in MARV, EBOV, and LLOV minigenome systems had a negative effect on minigenome activity [54,66,77,139]. This inhibitory effect is likely due to premature VP24-driven nucleocapsid condensation, which interferes with the movement of the polymerase along the RNA template [68]. The MARV and EBOV VP40 matrix proteins have also been shown to block minigenome activity [66,77,140] (Fig. 3). The role of VP24 and VP40 in viral RNA synthesis has not been studied for other filoviruses.

VP30

The fifth viral protein associated with marburg- and ebolavirus nucleocapsids is VP30 [104,105]. Its function during viral transcription is discussed in the [Transcription-specific functions of filoviral proteins](#) and [The role of VP30 in filovirus transcription](#) sections.

Inclusions

The marburg- and ebolavirus nucleocapsids form large inclusions in the infected cells. These viral inclusions are the sites of viral RNA synthesis, nucleocapsid formation, and nucleocapsid maturation [13,37,42,43,45,46,141,142]. Inclusion formation in infected cells has also been observed for other NNS RNA viruses. Recent work on viral inclusions in rabies virus- and VSV-infected cells shows that the viral inclusions have the characteristics of liquid droplets formed by phase separation, similar to stress granules, P bodies, and other cellular

RNA–protein complexes [143,144]. Given that EBOV inclusions show some of the features characteristic of liquid organelles as observed by live-cell imaging (e.g., they fuse and form larger spheres) [37], it is tempting to assume that the filovirus inclusions, too, are distinct liquid organelles. However, this has not yet been experimentally shown. In EBOV-infected cells, cellular proteins involved in RNA regulation and translation form distinct granules that are sequestered within the viral inclusions, but are separate from them, supporting the idea of liquid–liquid phase separation [44]. The functional relevance of these inclusion-bound granules is not clear. It is not known whether the sequestration of distinct cellular protein granules is unique for ebolaviruses or also occurs during infection with other filoviruses. Similar to MARV and EBOV, the LLOV nucleocapsid proteins form inclusions when expressed in cells [53]. Given the high degree of similarity regarding the genome organization and encoded viral proteins of filoviruses (with exception of the fish filoviruses), it is likely that both filoviral nucleocapsid structure and inclusion formation are conserved. The fish filoviruses are clearly distinct from the other filoviruses; neither encodes a homolog to VP24, and HUVJ additionally does not encode clear homologs to VP35 or VP30 (Figs. 1b and 9). It would be highly interesting to explore if the differences in the genome organization of these viruses are also reflected in significant structural and functional differences compared to the more “conventional” mammalian filoviruses, including whether these viruses form inclusions upon infection.

Transcription-specific functions of filoviral proteins

While the protein requirements for viral genome replication and transcription overlap, there are also some key differences, including differences between the filovirus genera. Similar to filovirus genome replication, transcription requires encapsidation of the genome by NP. In addition to VP35 and L, the transcription factor VP30 is involved in filovirus

Fig. 8. Conserved regions in filovirus proteins. Alignments of key conserved motifs in filovirus proteins involved in replication and transcription, (a) NP, (b) VP35, (c) VP30, and (d) L, were performed manually. Alignment of the L protein motifs includes similar motifs in the L proteins of other NNS RNA viruses. Motif names or functions are indicated above. Critical residues of conserved filovirus motifs are shown above alignments and are indicated in bold and underlined. Similar amino acids to the conserved motifs are in black but not bold or underlined and dissimilar amino acids are in gray. Surrounding amino acids not part of the consensus sequence are in blue with conserved sequences in bold. Amino acids highlighted in magenta and blue indicate a divergence between the mammalian filoviruses and the other NNS RNA viruses, including XILV and HUVJ. Sequences in which no similar motifs were found are indicated with “not found.” Instances where similar motifs were found in different proteins (e.g., in panel c), are indicated by including the divergent viral gene name. Instances where multiple potential motifs were identified in the same protein are indicated with an underscore and number. Abbreviations of filovirus names are explained in the legend to Fig. 1. BDV, Borna disease virus; HeV, Hendra virus; MV, measles virus; NDV, Newcastle disease virus; NiV, Nipah virus; PRNTase, GDP polyribonucleotidyl-transferase; RABV, rabies virus; RSV, respiratory syncytial virus; SAM, S-adenosyl-methionine; VSV, vesicular stomatitis virus.

(a)		(b)		(c)	
VP30 binding domain		NP binding domain		C3H1 Zinc Finger	
VP30 binding domain	NP	VP35	VP30	VP35	VP30
PP2A-B56 binding domain	PP2A-B56 binding domain	NP binding domain	C3H1 Zinc Finger	NP binding domain	C3H1 Zinc Finger
<u>PPAPXY</u>	<u>L*PxxE</u>	<u>ISEQLMTGK-IPVSDIF</u>	<u>PKDICTLKKGFLCDSDF-CKKDHQLES</u>	<u>ISEQLMTGK-IPVSDIF</u>	<u>PKDICTLKKGFLCDSDF-CKKDHQLES</u>
EBOV	EBOV	EBOV	EBOV	EBOV	EBOV
SUDV	SUDV	SUDV	SUDV	SUDV	SUDV
BDBV	TAFV	TAFV	TAFV	TAFV	TAFV
RESTV	BDBV	BDBV	BDBV	BDBV	BDBV
TAFV	RESTV	RESTV	RESTV	RESTV	RESTV
BOMV	BOMV	BOMV	BOMV	BOMV	BOMV
LLOV	LLOV	LLOV	LLOV	LLOV	LLOV
MARV	MARV	MARV	MARV	MARV	MARV
RAVV	RAVV	RAVV	RAVV	RAVV	RAVV
MLAV	MLAV	MLAV	MLAV	MLAV	MLAV
XILV_1	XILV_1	XILV	XILV	XILV	XILV
XILV_2	XILV_2	HUJV_ORF3	HUJV_ORF5_1	HUJV_ORF5_1	HUJV_ORF5_1
XILV_3	XILV_3	HUJV not found	HUJV_ORF5_2	HUJV_ORF5_2	HUJV_ORF5_2
HUJV_1	HUJV_1				
HUJV_2	HUJV_2				
HUJV_3	HUJV_3				
HUJV not found	HUJV not found				

(d)		(b)		(c)	
Template binding	Catalytic domain (GDN motif)	SAM binding domain	Methyltransferase tetrad (K-D-K-E)	PRNTase	
<u>KEKE</u>	<u>GDN</u>	<u>GxGxG</u>	<u>K</u>	<u>D</u>	<u>K</u>
EBOV	AV <u>GDNQ</u>	EBOV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VVL <u>KVFL</u> ..RSSE <u>WYL</u>	<u>E</u>	<u>HR</u>
SUDV	AV <u>GDNQ</u>	SUDV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..KSSE <u>WYL</u>	EBOV	YSGNIV <u>HR</u> NDQY
TAFV	AV <u>GDNQ</u>	TAFV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VVL <u>KVFL</u> ..KSSE <u>WYL</u>	SUDV	YSGNIV <u>HR</u> NDQY
BDBV	AV <u>GDNQ</u>	BDBV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..KSSE <u>WYL</u>	TAFV	YSGNIV <u>HR</u> NDQY
RESTV	AV <u>GDNQ</u>	RESTV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..KSSE <u>WYL</u>	BDBV	YSGNIV <u>HR</u> NDQY
BOMV	AV <u>GDNQ</u>	BOMV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	RESTV	YSGNIV <u>HR</u> NDQY
LLOV	AV <u>GDNQ</u>	LLOV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	BOMV	YSGNIV <u>HR</u> NDQY
MARV	SV <u>GDNQ</u>	MARV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	LLOV	YSGNIV <u>HR</u> NDQY
RAVV	SV <u>GDNQ</u>	RAVV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	MARV	YSGNIV <u>HR</u> NDQY
MLAV	CV <u>GDNQ</u>	MLAV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	RAVV	YSGNIV <u>HR</u> NDQY
XILV	AV <u>GDNQ</u>	XILV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	MLAV	YSGNIV <u>HR</u> NDQY
HUJV	MI <u>GDNQ</u>	HUJV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	XILV	YSGNIV <u>HR</u> NDQY
RSV	LIN <u>GDNQ</u>	RSV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	HUJV	YSGNIV <u>HR</u> NDQY
VSV	LA <u>GDNQ</u>	VSV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	RSV	YSGNIV <u>HR</u> NDQY
RABV	LA <u>GDNQ</u>	RABV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	VSV	YSGNIV <u>HR</u> NDQY
MV	LV <u>GDNQ</u>	MV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	RABV	YSGNIV <u>HR</u> NDQY
NDV	MV <u>GDNQ</u>	NDV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	MV	YSGNIV <u>HR</u> NDQY
BDV	LG <u>GDNQ</u>	BDV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	NDV	YSGNIV <u>HR</u> NDQY
NiV	IV <u>GDNQ</u>	NiV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	BDV	YSGNIV <u>HR</u> NDQY
HeV	IV <u>GDNQ</u>	HeV	<u>MHYK</u> LDDE..ITND <u>DAET</u> ..VIL <u>KVFL</u> ..RSSE <u>WYL</u>	NiV	YSGNIV <u>HR</u> NDQY
				HeV	YSGNIV <u>HR</u> NDQY

(a) XILV			(b) HUJV		
gene	classic filovirus homolog	conserved domains	gene	classic filovirus homolog	conserved domains
NP	NP	Ebola NP superfamily (aa 9-471) LxxLF late domain (LIKLF, aa 227-231) P(T/S)AP late domain (PSAP, aa 496-499)	ORF1	matrix?	PPxY late domain (PPVY, aa 182-185)
VP35*	VP35	Filo_VP35 Pfam domain (aa 86-274) Coiled-coil domain (aa 120-170)	NP	NP	Ebola NP superfamily (aa 270-340)
VP40	VP40	VP40 Pfam domain (aa 30-192) P(T/S)AP late domain (PSAP, aa 206-209)	ORF3	VP35?	Coiled-coil domain (aa 389-423)
ORF4		Transmembrane helix region (aa 5-27)			
GP1	GP	Signal peptide (aa 1-23) Transmembrane region (aa 81-98) Coiled-coil domain (aa 100-168) Transmembrane region (aa 390-412)	GP	GP	Fusion_gly Pfam domain (aa 74-405) Furin (PACE) cleavage site (RRSRR, aa 156-160) Coiled-coil domain (aa 190-233) Transmembrane region (aa 550-572)
GP2	GP?	Transmembrane region (aa 7-29)			
VP30	VP30	Zinc finger C3H1-type profile (aa 75 - 102) Transcript VP30 superfamily (aa 155-279)	ORF5	VP30?	Zinc finger C3H1-type profile (aa 56 - 83) Zinc finger C3H1-type profile (aa 239-261)
ORF8		none found			
ORF9		Orn/DAP/Arg decarboxylases family 2 signature 2 (aa 4 - 20) PPxY late domain (PPEY, aa 161-164), LYPX _N L late domain (LYPGL, aa 209-213) P(T/S)AP late domain (PTAP, aa 241-244)			
L	L	Mononegavirales RNA polymerase superfamily (aa 16-1084) Mononegavirales mRNA cap superfamily (aa 1104-1357) G-7-MTase superfamily (aa 1216-2035)	L	L	Mononegavirales RNA polymerase superfamily (aa 95-1061) Mononegavirales mRNA cap superfamily (aa 1099-1341) G-7-MTase superfamily (aa 1203-1938)

*not previously identified as a homolog to other filovirus genes

Fig. 9. Conserved domains in XILV and HUJV proteins. XILV and HUJV proteins were searched for homology to other filovirus proteins and for the presences of key conserved domains that are known to be important for different aspects of filovirus replication. Domain searches were performed using protein BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>), ELM (<http://elm.eu.org/>), Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>), ScanProsite (<https://prosite.expasy.org/scanprosite/>), SMART (<http://smart.embl-heidelberg.de/>), and manually. Each (a) XILV and (b) HUJV gene is listed with the discovered domains, including the location of the domain as indicated by amino acid number, and any predicted or suggested homology to ebola- and marburgvirus proteins. Proteins are colored as in Fig. 1, with the exceptions of HUJV ORF3 and ORF5, which are here colored similarly to VP35 and VP30, respectively. Abbreviations of filovirus names are explained in the legend to Fig. 1. MTase, methyltransferase.

transcription, although there are significant differences in the requirement for VP30 between the different filovirus genera. Nothing is yet known about XILV and HUJV transcription.

Transcription-specific functions of L

Shortly after transcription initiation, the nascent mRNA chain is capped by the viral polymerase [145]. Although EBOV mRNAs are capped [61], there is not much known about the filoviral capping mechanisms. The L protein of the closely related VSV uses an unconventional mechanism to add the ^{m7}Gppp cap to the 5' ends of the viral mRNAs, which involves various enzymatic activities, including a guanosine 5'-triphosphatase, a GDP polyribonucleotidyl-transferase (PRNTase), and a dual methyltransferase (MTase) activity [135,146,147]. The PRNTase activity depends on a His₁₂₂₇–Arg₁₂₂₈ (HR) amino acid motif in domain V of VSV L that is also found in the L proteins of other NNS RNA viruses, including all filoviruses, suggesting that the unusual VSV capping mechanism is conserved among NNS RNA viruses (Fig. 8d, PRNTase) [146]. The MTase domain, which is another conserved activity of most NNS RNA virus polymerases, includes a GxGxG motif and catalytic K-D-K-E tetrad [148] and is found in all filoviral L proteins (Fig. 8d, SAM binding domain and MTase tetrad (K-D-K-E)). Functional analysis of the MTase activity of filoviral L has only been performed for SUDV L. The C-terminal domain of SUDV L containing region VI is able to transfer methyl groups to the guanine N7 position of the cap when RNA substrates are used that start with a 2'-O-methylated guanosine residue. However, substrates starting with methylated adenosine residues are not accepted [132]. This matches beautifully with the conserved filovirus GS signals—all filovirus mRNAs start with a guanosine residue (Fig. 6a). SUDV L is also able to methylate the ribose 2'-O position of the first guanosine residue of the RNA substrate, but only if the substrate is longer than 30 nts [132]. This is reminiscent of VSV for which it was shown that capping only occurs when the nascent mRNA reaches a length of 31 nts [149]. Martin and colleagues [132] propose the following model: SUDV transcription starts with the synthesis of an approximately 30-nt-long uncapped transcript which remains bound to the PRNTase domain until the non-methylated cap guanosine is added. Following addition of the guanosine cap to the nascent RNA, two methyl groups are added to the cap by the MTase domain of L. The GxGxG motif within the MTase domain binds S-adenosyl-L-methionine which is the source of the methyl groups added to the guanosine RNA cap. The first methylation step, 2'-O methylation of the first nucleotide of the nascent mRNA (GpppG_m), is catalyzed via the

surrounding K-D-K-E tetrad. Following 2'-O methylation, N7 methylation of the cap guanosine (^{m7}GpppG_m) is mediated by the same MTase domain. In addition to cap methylation, SUDV L has an unconventional internal adenosine-specific 2'-O-MTase activity that is cap-independent and might play a role for epigenetic control mechanisms or RNA packaging [132].

RNA editing

Ebolaviruses increase the coding capacity of the GP gene by cotranscriptional editing, resulting in the synthesis of multiple GP mRNA species [150–153]. The editing site within the GP gene contains a stretch of seven uridine residues that are transcribed into adenosine residues during transcription. The full-length membrane-bound GP_{1,2} is only expressed if an additional, non-templated adenosine residue is incorporated into the nascent mRNA strand (8A version), resulting in a frame shift [150,151]. If the mRNA is not edited (7A version), a soluble form of GP, sGP, is expressed that lacks the membrane anchor [154]. It has been suggested that GP mRNA editing is required for downregulating GP_{1,2} expression in order to reduce its cytotoxic effects [71,155]. The 7U stretch is also used as a cryptic transcription termination and polyadenylation site. About 40% of the synthesized EBOV GP mRNAs are prematurely terminated at the 7U stretch [156]. Approximately a quarter of the remaining 60% of the mRNAs contain a single non-templated adenosine residue, resulting in GP_{1,2} expression [150,152]. A mechanistic study on EBOV editing revealed that VP30 is involved in this process. According to the proposed model, an RNA hairpin is formed upstream of the editing site in the nascent GP mRNA, leading to polymerase stuttering. VP30 is required to overcome this hairpin, similar to its proposed function during transcription initiation at the NP gene GS signal (see below, [The role of VP30 in filovirus transcription](#)) [157].

Editing is not restricted to mRNA synthesis but also occurs during EBOV replication. A common cell culture adaptation is the insertion of a uridine residue at the GP gene editing site (8U mutant) that leads to predominant expression of membrane-bound GP_{1,2} and increased cytotoxicity [155,158,159]. This 8U cell culture adaptation was also observed for SUDV [153]. Intriguingly, the 8U mutant reverts back to the 7U wildtype in animal models of EBOV infection [153,158,159]. Follow-up studies showed that recombinant EBOV clones overexpressing GP_{1,2} were attenuated in a guinea pig infection model, possibly due to GP_{1,2}-induced premature cell death. This suggests that controlled GP_{1,2} expression is an important virulence factor in EBOV infection [155]. The reversion back to the 7U wildtype was less pronounced in a nonhuman primate infection model

for SUDV, indicating that the different ebolavirus species might use distinct strategies to modulate GP expression [153].

Besides the ebolaviruses, only LLOV uses mRNA editing to express membrane-bound GP_{1,2} [35,36]. The GP ORFs of marburgviruses, MLAV, and the fish filoviruses are open, and each only encodes a single version of GP_{1,2} that does not require transcriptional editing (Fig. 1b). While RNA editing has mostly been studied in the context of EBOV, a recent deep sequencing analysis of RNA from cells and non-human primates infected with MARV Angola revealed novel editing sites in the NP and L ORFs, potentially resulting in novel forms of these proteins [100]. It is currently unknown whether cotranscriptional editing is a pan-filovirus feature or if it is only conserved within a subset of the filovirus family.

The role of VP30 in filovirus transcription

Marburg- and ebolavirus VP30 proteins share some structural and functional features with the M2-1 proteins of the pneumoviruses but are not found in other NNS RNA viruses [160–163]. As mentioned above, they are associated with the nucleocapsid [104,105]. While MARV and EBOV have the same protein requirements for genome replication, there are differences in the transcription strategies of these viruses regarding VP30 as outlined below.

The role of Ebola virus VP30 in transcription

In addition to NP, VP35, and L, ebolavirus transcription requires the transcription activator VP30 (Fig. 3) [50–52]. EBOV VP30 is a single-stranded RNA binding protein that forms hexamers and interacts with NP, VP35, and L [107,129,164–171]. It contains an unusual C3H1 Zinc binding motif that is required for transcriptional activity and RNA binding [166,168,172]. Although VP30 is tightly associated with the nucleocapsid complex, it is not essential for nucleocapsid formation [14,22]. VP30 is involved in various steps of EBOV transcription. During transcription initiation of the first gene, the NP gene, VP30 is required to overcome an RNA hairpin structure adopted by the NP GS signal with downstream located sequences (Fig. 5b). Disruption of this secondary structure renders EBOV minigenome transcription VP30-independent [61,171]. If transcription does not proceed beyond the first GS signal, none of the other viral mRNAs can be synthesized. It has been proposed that VP30 forms a ternary complex with the Le(–) RNA and the EBOV polymerase through interaction with VP35 to stabilize the transcription initiation complex and prevent the polymerase from falling off the template due to destabilization caused

by the RNA hairpin [168]. VP30 is also involved in later steps of the viral transcription cycle [173,174]. It has been proposed that it acts as a reinitiation factor at the internal gene boundaries [174]. However, transcription reinitiation in a bicistronic minigenome system occurred independently of VP30 [61]. It has also been suggested that VP30 is involved in the switch from transcription to replication [167,173]. Finally, VP30 is required for GP mRNA editing, as described above in the mRNA editing section [157]. The transcriptional activity of EBOV VP30 depends on its phosphorylation state (Fig. 3). Phosphorylation of VP30 occurs at serine and threonine residues mainly located in the N-terminal domain of the protein [164,175]. Non-phosphorylated or hypo-phosphorylated VP30 enhances transcription, whereas the phosphorylated form is transcriptional inactive. Phosphorylation of VP30 interferes with its ability to interact with NP, VP35, and RNA, and excludes VP30 from the transcription complex. This leads to enhanced replication activity which can be carried out without VP30 (Fig. 3) [167,168,176]. The presence of hypo-phosphorylated VP30 is crucial for the EBOV replication cycle. Intriguingly, NP recruits the cellular phosphatase PP2A-B56 through a B56 binding motif (Fig. 8a, PP2A-B56 binding domain). This leads to VP30 dephosphorylation and transcriptional activation [177]. The PP2A-B56 binding domain within NP is conserved in all the filoviruses, with the possible exception of HUVV, which does not encode a clearly identifiable VP30 homolog. Blocking the activity of cellular phosphatases involved in VP30 dephosphorylation has therapeutic potential because hyper-phosphorylated VP30 is transcriptional inactive. Indeed, phosphatase inhibitors interfere with EBOV replication [164,175,177].

Crystal structure analysis showed that the C-terminal domain of VP30 interacts with a peptide within the C-terminal domain of NP [107,170,171]. This interaction is crucial for VP30 activity and seems to be highly regulated. Mutational analysis in one study suggested that the VP30–NP binding affinity has to be in an optimal range. Mutants that resulted in too high or too low binding affinities did not perform well in the minigenome system [170]. Another study suggested that even minimal VP30–NP binding is sufficient to support transcriptional activity. This study also showed that the NP-VP30 interaction is required for the early activities of VP30, which are to overcome the RNA hairpin structure at the first GS signal [171]. The VP30-interacting peptide in NP contains a PPxPxY motif that is conserved across all mammalian filoviruses (Fig. 8a, VP30 binding) [170,171]. The cellular ubiquitin ligase RBBP6 contains the same binding motif and is able to sequester VP30, leading to reduced transcriptional activity [178]. As for the fish filoviruses, XILV NP contains a handful of slightly less conserved VP30-

binding motifs, and HUJV NP lacks any similar motif (Fig. 8a, VP30 binding). The absence of these motifs in HUJV NP is perhaps not surprising since the virus lacks a clear VP30 homolog.

The role of Marburg virus VP30 in transcription

Although there is a high degree of structural homology between MARV and EBOV VP30 proteins, the role of VP30 in MARV transcription is less clear. Similar to EBOV VP30, MARV VP30 is phosphorylated and binds to NP and VP35 [105,106,170,179]. While VP30 is an essential component of the EBOV minigenome system and enhances transcriptional activity at least 50-fold, MARV VP30 is dispensable for minigenome activity and increases reporter gene activity only moderately [49,66,133,179]. Despite this difference in the minigenome systems, MARV VP30 is essential for viral gene expression and for the rescue of full-length viral clones [72,180]. Similar to EBOV VP30, the phosphorylation state of MARV VP30 determines its function in the minigenome system, and treatment of MARV-infected cells with phosphatase 1 inhibitors reduced viral titers [179]. Crystal structure analysis revealed a striking conformational similarity of the EBOV and MARV VP30–NP complexes [170]. Interestingly, the binding affinity of the MARV VP30–NP complex is lower compared to EBOV [170]. It remains to be determined whether this difference has functional consequences.

Replication and Transcription of Novel Filoviruses

Since there are no virus isolates available for any of the recently discovered filoviruses, replication and transcription can only be studied using minigenome systems. Besides ebola- and marburgviruses, minigenome systems have been established for LLOV and MLAV [7,53,54]. A caveat of these minigenome systems is that the genome end sequences of these viruses are not complete and therefore require complementation with sequences from other filoviruses. The 3' terminus of the published LLOV genome, for example, lacks 3 or 4 nucleotides [5] (Fig. 5a). When complemented with EBOV genome ends (3' GCCU), chimeric LLOV minigenomes are accepted as templates for replication and transcription by the LLOV polymerase, whereas minigenomes complemented with the MARV genome ends (3' UCU) are not [53]. This suggests that LLOV follows an ebolavirus-like replication strategy (see the [Filoviral replication promoters](#) and [Genomic replication promoter](#) sections). In contrast to this highly specific recognition of the 3' end, the LLOV

polymerase complex recognizes both the EBOV and MARV 5' trailer regions [53]. Similar to EBOV, LLOV minigenome expression is strongly enhanced in the presence of VP30, indicating that LLOV utilizes an ebolavirus-like transcription strategy [53,54]. Due to the lack of sequence information for the leader and trailer regions, the function of the MLAV NP, VP35, VP30, and L proteins was tested with minigenomes that contained MARV or EBOV leader and trailer regions. The MLAV proteins supported reporter gene expression from both minigenomes [7]. Nothing is known about the replication and transcription strategies of XILV and HUJV.

Putative Functions of Unassigned XILV and HUJV Proteins

Some, but not all, of the XILV and HUJV genes were previously assigned gene names and presumed functions based on homology to other filovirus genes [9]. A more targeted search for conserved filovirus protein motifs and domains helped to assign additional putative functions and putative filovirus homologs for the XILV and HUJV genes.

NP, L, and GP

Given the importance of the NP–RNA complex, it is not surprising that XILV and HUJV possess NP homologs, each of which contains a conserved filovirus NP domain (Fig. 9). The same holds true for the L proteins, which contain all of the highly conserved NNS RNA virus L protein domains (Figs. 8d and 9). Fish filoviruses are divergent from the other filoviruses as highlighted by the conserved amino acid residue preceding the catalytic GDN motif in L; instead of the filovirus-specific methionine residue, the fish filoviruses have a glutamine residue, which is more characteristic of the L proteins of other NNS RNA viruses (Fig. 8d, Catalytic domain (GDN motif)). Both of the fish filoviruses also encode one (HUJV) or even two (XILV) GP candidates that would enable them to attach to the host cells (Fig. 9). Which cellular receptor(s) these glycoproteins utilize for attachment and entry has not yet been explored.

VP35

Surprisingly, initial characterization of the XILV and HUJV genomes suggested that neither of these viruses encode a clear VP35 homolog. As a P protein homolog, VP35 is an indispensable component of the polymerase complex. It is therefore highly unlikely that the fish filoviruses do not encode a homolog to this protein. As mentioned above in the [Filovirus proteins involved in RNA synthesis](#) section, a general feature of the NNS RNA virus P proteins, including filovirus VP35, is the possession of a

(a) VP35 IID		(b) VP24 Keap1 binding domain	
	<u>QKSLR</u>		<u>D/E</u> <u>GE</u>
EBOV	<u>PRAC</u> <u>QKSLR</u> <u>FV</u>	Nrf2	<u>QLDEET</u> <u>GE</u>
LLOV	<u>PKAC</u> <u>QKSLR</u> <u>PL</u>	p62	<u>EVDPST</u> <u>GE</u>
MARV	<u>PRPC</u> <u>QKSLR</u> <u>AV</u>	MARV_VP24	<u>DI</u> <u>EPCCGE</u>
MLAV	<u>PRPC</u> <u>QKSLR</u> <u>AV</u>	EBOV_VP24	not found
XILV	<u>VAEI</u> <u>Q</u> - <u>SLRAS</u>	LLOV_VP24	not found
HUJV	not found	MLAV_VP24	not found
		XILV_ORF8	not found
		XILV_ORF9	not found
		HUJV_ORF5_1	<u>HD</u> <u>LKCGE</u>
		HUJV_ORF5_2	<u>EE</u> <u>EGSEGE</u>
		HUJV_ORF5_3	<u>SSG</u> <u>DPGE</u>

Fig. 10. Conserved domains involved in interferon inhibition and antioxidant pathway induction. Alignments of conserved motifs in filovirus proteins involved in the inhibition of (a) interferon induction (VP35 IID) and (b) the induction of a cytoprotective antioxidant response (VP24 Keap1 binding domain) were performed manually. Critical residues of conserved motifs are shown above alignments and are indicated in bold and underlined. Dissimilar amino acids and gaps are in gray. Surrounding amino acids not part of the consensus sequence are in blue with conserved amino acids in bold. Sequences in which no similar motifs were found are indicated with “not found.” Instances where similar motifs were found in different proteins (e.g., in panel b), are indicated by including the divergent viral gene name. Instances where multiple motifs or potential motifs were identified in the same protein are indicated with an underscore and number. For panel b, two human cellular proteins, Nrf2 and p62, that are known to bind to Keap1 via a similar protein motif were included in the alignment. Abbreviations of virus names are explained in the legend to Fig. 1. IID, interferon inhibition domain; Keap1, Kelch-like ECH-associated protein 1; Nrf2, nuclear factor, erythroid 2 like 2.

coiled-coil motif required for homo-oligomerization and polymerase function. Additional scanning revealed that the previously unnamed ORF2 of XILV does indeed have weak homology to VP35, including the coiled-coil motif (Fig. 9a). HUJV ORF3, while not displaying any noticeable homology to other filovirus VP35 genes, also has a coiled-coil motif which could hint at a possible VP35-like function for this protein (Fig. 9b).

In addition to their function in replication and transcription, marburgvirus, ebolavirus, and LLOV VP35 proteins play important roles in antagonizing the host immune response, including the type I interferon response (reviewed in Ref. [181]). The VP35 interferon-inhibiting domain resides within the C-terminus of the protein [182–187]. Interestingly, XILV VP35 contains a somewhat similar domain but which lacks a key lysine residue (Fig. 10a). HUJV ORF3 does not contain anything remotely similar to this conserved domain. Given the lack of conservation, it is conceivable that the “VP35” proteins of the fish filoviruses do not have the interferon (IFN) inhibiting function due to differences in the antiviral responses between fish and mammals.

VP40/matrix protein

HUJV is the only known filovirus that does not encode a VP40 homolog. Although it remains to be experimentally determined which HUJV protein mediates budding and how this affects the shape of the viral particles, protein motif search provides some hints regarding the identity of the putative matrix protein. All filovirus VP40 proteins contain “late domain” motifs that are involved in binding to

components of the ESCRT machinery in order to facilitate viral budding (role of late domains in budding reviewed in Refs. [188–191]). Although HUJV has no clear matrix protein, HUJV ORF1 contains a PPxY late domain, potentially indicating a role in viral budding (Fig. 9b). XILV encodes three proteins that contain potential late domains, including NP (akin to MARV NP, which contains a late domain; [192]), VP40, and ORF9 (Fig. 9a). It remains unclear which of these potential late domains might be involved in XILV budding.

VP30

XILV was previously annotated to encode a VP30 homolog [9]. The XILV VP30 homolog contains a C3H1 zinc finger motif which is characteristic for all filoviral VP30 proteins. Interestingly, although there is no clear VP30 homolog encoded by HUJV, HUJV ORF5 contains two C3H1 zinc finger motifs (Figs. 8c and 9b; see also the [Transcription-specific functions of filoviral proteins](#) and [The role of VP30 in filovirus transcription sections](#)), and may therefore perform similar functions to that of VP30 despite low sequence similarity.

VP24

As determined by structural analysis, the ebolavirus VP24 proteins contain conserved amino acids that are involved in binding to karyopherin alpha proteins, thereby blocking interferon signaling by preventing nuclear translocation of activated STAT proteins [193–196]. Since neither of the fish filoviruses encode a clear homolog to VP24,

however, it is not surprising that neither of these viruses have proteins with similar sequence homology to this binding domain.

MARV VP24, which does not inhibit IFN signaling, was previously shown to activate the antioxidant response pathway by interacting with Kelch-like ECH-associated protein 1 (Keap1), a suppressor of this pathway [197,198]. MARV VP24 interacts with Keap1 using a domain reminiscent of the Keap 1 binding domain of two cellular proteins, nuclear factor erythroid 2-related factor 2 (Nrf2) and p62 (Fig. 10b) [197]. The Keap 1 binding domain (DX_NGE) is only conserved in marburgvirus VP24 proteins; it is not found in ebolavirus, LLOV, or MLAV VP24 proteins, or the XILV ORF8 or ORF9. Intriguingly, HUVJ ORF5 contains three domains that look similar to the MARV VP24 Keap1 interacting domain, potentially hinting at a role for this protein at inducing a cytoprotective response during infection (Fig. 10b).

Other XILV and HUVJ proteins

A scan of the XILV and HUVJ proteins with unknown functions revealed that XILV ORF9 contains an Orn/DAP/Arg decarboxylases family 2 signature 2 (ODR_DC_2_2) domain, which is found in proteins involved in the polyamine synthesis pathway. This is an intriguing finding given the recently described role of polyamines in EBOV gene expression and replication [78,199], potentially indicating that dependence on this pathway could be a pan-filovirus characteristic.

Conclusions

Among the “classical,” mammalian filoviruses, including the new members, LLOV and MLAV, there appears to be two distinct patterns with regards to the strategies employed for genome replication and transcription. LLOV appears to be more like the ebolaviruses. This includes similarities in the replication promoter structure, the requirement of VP30 for minigenome transcription, and the use of cotranscriptional mRNA editing to encode multiple GP ORFs. This pattern of similarity to ebolaviruses holds true in other aspects as well. While the marburgviruses use VP40 to inhibit interferon signaling, the ebolaviruses and LLOV use VP24 instead [200]. MLAV, on the other hand, seems to be more closely related to the marburgviruses (Fig. 1a) and has some marburgvirus-like features. For example, MLAV does not use cotranscriptional editing to expand the coding capacity of its GP gene. One noticeable difference between MLAV and the marburgviruses, however, is the lack of the Keap1 binding domain in MLAV VP24: the induction of the antioxidant pathway by Keap1 sequestration

appears to be restricted to the marburgviruses and possibly HUVJ. It remains to be determined whether MLAV follows the strategy employed by ebolaviruses or marburgviruses to block IFN signaling.

Given the genetic diversity of the newly discovered saltwater fish filoviruses, they will each be representative of new filovirus lineages [3]. Although we currently only have sequence information to analyze, many differences can be found in their genomes; divergences in replication promoters, and GS and GE signals; their lack of a VP24 homolog; the lack of clear VP35, VP40, and VP30 homologs in HUVJ; and the lack of certain key protein domains. These disparities in the XILV and HUVJ genomes necessitate that their mechanisms of viral genome replication and transcription will be different from the other filoviruses and likely different from each other.

The genetic diversity of fish filoviruses indicates that many more lineages of filoviruses may exist. The discovery of filoviruses in saltwater fish hosts opens up the likelihood that filoviruses will be found in other unexpected host species, as these fish are quite divergent from the proposed bat reservoirs of the mammalian filoviruses.

In addition to the surprising diversity of filovirus hosts, the discovery of MLAV in China and LLOV in Europe gives confidence in the larger geographical distribution of filoviruses. Previously, only RESTV was found outside of Africa. Given the diversity of filoviruses and their hosts, it seems probable that filoviruses exist in many other host species likely to be found in a larger geographic distribution.

Although we do not know the pathogenic potential of these diverse new filoviruses, given the high lethality of some of the family members, we should be prepared for possible spillover events. Using tools such as minigenome systems, we can not only study how these diverse viruses replicate and transcribe their genomes, we will also be able to test and develop antiviral therapies with the goal of building an arsenal of broad-spectrum antiviral drugs that could be useful in the face of potential outbreaks caused by newly emerging filoviruses.

Minigenome systems for the fish filoviruses do not currently exist due to the apparent lack of key promoter elements at the genome termini. As new filoviruses are discovered, the ability to ascertain as much terminal sequence as possible will be critical not only for the ability to develop useful minigenome systems but also for the creation of reverse genetic systems. This will be essential for the timely studies of these viruses, since the culturing of filoviruses from their natural hosts has been historically difficult.

Sequence information used in this study

The following GenBank accession numbers were used for sequence comparisons in this study: EBOV, NC_002549.1; SUDV, NC_006432.1; TAFV,

NC_014372.1; BDBV, NC_014373.1; RESTV, NC_004161.1; BOMV, NC_039345.1; LLOV, NC_016144.1; MARV, NC_001608.3; RAVV, NC_024781.1; MLAV, KX371887.2; XILV, MG599980.1; HUJV, MG599981.1; RSV, NC_038235.1; VSV, NC_001560.1; RABV, NC_001542.1; MV, NC_001498.1; NDV, NC_039223.1; BDV, NC_001607.1; NiV, NC_002728.1; HeV, NC_001906.3.

Acknowledgments

The authors would like to thank Kristina Brauburger, Lund University, Sweden, for valuable help with Figs. 2 and 3, and the graphical summary, and John Briggs, MRC Laboratory of Molecular Biology, Cambridge, UK, for providing the images for Fig. 7b. We are grateful to the three reviewers for their helpful comments. This work was funded by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers R01 AI133486 (to Rachel Fearn and E.M.) and R21 AI137793 (to E.M.).

Declaration of Interest Statement: None.

Received 21 February 2019;

Received in revised form 31 May 2019;

Accepted 24 June 2019

Available online 29 June 2019

Keywords:

Ebola virus;
Marburg virus;
filoviruses;
nonsegmented negative sense RNA viruses;
genome replication and transcription

Abbreviations used:

BDBV, Bundibugyo virus; BDV, BOMV, Bombali virus; EBOV, Ebola virus; ESCRT, endosomal sorting complex required for transport; GE, gene end; GP, glycoprotein; GS, gene start; HUJV, Huángjiāo virus; IFN, interferon; IR, intergenic region; iVLPs, infectious virus-like particles; Keap1, Kelch-like ECH-associated protein 1; L, large protein (RNA-dependent RNA polymerase); Le(-), genomic leader; Le(+), antigenomic leader; LLOV, Lloviu virus; MARV, Marburg virus; MLAV, Mēnglà virus; MTase, dual methyltransferase; NNS, nonsegmented negative sense; NP, nucleoprotein; NPC1, Niemann Pick C1; Nrf2, nuclear factor, erythroid 2 like 2; ORF, open reading frame; p62, 62-kDa protein; PE, promoter element; PRNTase, GDP polyribonucleotidyl-transferase; RAVV, Ravn virus; RESTV, Reston virus; sGP, soluble glycoprotein; SUDV, Sudan virus; TAFV, Tai Forest virus; TAM, Tyro3, Axl, and Mer; TIM, T-cell immunoglobulin and mucin domain; Tr(-), genomic trailer; Tr(+), antigenomic trailer; trVLP,

transcription- and replication-competent virus-like particle; VSV, vesicular stomatitis virus; XILV, Xilǎng virus.

References

- [1] V. Rougeron, H. Feldmann, G. Grard, S. Becker, E.M. Leroy, Ebola and Marburg haemorrhagic fever, *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology*. 64 (2015) 111–119.
- [2] K. Brauburger, A.J. Hume, E. Mühlberger, J. Olejnik, Forty-five years of Marburg virus research, *Viruses*. 4 (2012) 1878–1927.
- [3] J.H. Kuhn, G.K. Amarasinghe, C.F. Basler, S. Bavari, A. Bukreyev, K. Chandran, et al., ICTV virus taxonomy profile: Filoviridae, *J. Gen. Virol.* 100 (6) (2019) 911–912.
- [4] T. Goldstein, S.J. Anthony, A. Gbakima, B.H. Bird, J. Bangura, A. Tremeau-Bravard, et al., The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses, *Nat. Microbiol.* 3 (2018) 1084–1089.
- [5] A. Negrodo, G. Palacios, S. Vazquez-Moron, F. Gonzalez, H. Dopazo, F. Molero, et al., Discovery of an ebolavirus-like filovirus in europe, *PLoS Pathog.* 7 (2011), e1002304.
- [6] G. Kemenesi, K. Kurucz, B. Dallos, B. Zana, F. Foldes, S. Boldogh, et al., Re-emergence of Lloviu virus in *Miniopterus schreibersii* bats, Hungary, 2016, *Emerg Microbes Infect.* 7 (2018) 66.
- [7] X.L. Yang, C.W. Tan, D.E. Anderson, R.D. Jiang, B. Li, W. Zhang, et al., Characterization of a filovirus (Mengla virus) from *Rousettus* bats in China, *Nat. Microbiol.* 4 (3) (2019) 390–395.
- [8] X.L. Yang, Y.Z. Zhang, R.D. Jiang, H. Guo, W. Zhang, B. Li, et al., Genetically diverse filoviruses in *Rousettus* and *Eonycteris* spp. Bats, China, 2009 and 2015, *Emerg. Infect. Dis.* 23 (2017) 482–486.
- [9] M. Shi, X.D. Lin, X. Chen, J.H. Tian, L.J. Chen, K. Li, et al., The evolutionary history of vertebrate RNA viruses, *Nature*. 556 (2018) 197–202.
- [10] K.M. Forbes, P.W. Webala, A.J. Jaaskelainen, S. Abdurahman, J. Ogola, M.M. Masika, et al., Bombali virus in *Mops condylurus* bat, Kenya. *Emerg Infect Dis.* 25 (2019).
- [11] D. Peters, G. Müller, W. Slenczka, Morphology, development, and classification of the Marburg virus, in: G.A. Martini, R. Siebert (Eds.), *Marburg virus disease*, Springer, Berlin, Heidelberg, New York, 1971, pp. 68–83.
- [12] T.W. Geisbert, P.B. Jahrling, Differentiation of filoviruses by electron microscopy, *Virus Res.* 39 (1995) 129–150.
- [13] E. Ryabchikova, B.B.S. Price, *Ebola and Marburg Viruses: A View of Infection Using Electron Microscopy*, Battelle Press, Columbus, OH, 2004.
- [14] T.A. Bharat, J.D. Riches, L. Kolesnikova, S. Welsch, V. Krahling, N. Davey, et al., Cryo-electron tomography of marburg virus particles and their morphogenesis within infected cells, *PLoS Biol.* 9 (2011), e1001196.
- [15] B. Liang, Z. Li, S. Jenni, A.A. Rahmeh, B.M. Morin, T. Grant, et al., Structure of the L protein of vesicular stomatitis virus from electron cryomicroscopy, *Cell*. 162 (2015) 314–327.
- [16] R.N. Kirchdoerfer, H. Wasserman, G.K. Amarasinghe, E.O. Saphire, *Filovirus structural biology: the molecules in the machine*, *Curr. Top. Microbiol. Immunol.* 411 (2017) 381–417.

- [17] R.A. Davey, O. Shtanko, M. Anantpadma, Y. Sakurai, K. Chandran, W. Maury, Mechanisms of filovirus entry, *Curr Top Microbiol Immunol* 411 (2017) 323–352.
- [18] L. Kolesnikova, A. Nanbo, S. Becker, Y. Kawaoka, Inside the cell: assembly of filoviruses, *Curr. Top. Microbiol. Immunol.* 411 (2017) 353–380.
- [19] R.V. Stahelin, Membrane binding and bending in Ebola VP40 assembly and egress, *Front. Microbiol.* 5 (2014) 300.
- [20] Y. Huang, L. Xu, Y. Sun, G.J. Nabel, The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein, *Mol. Cell* 10 (2002) 307–316.
- [21] T. Noda, K. Aoyama, H. Sagara, H. Kida, Y. Kawaoka, Nucleocapsid-like structures of Ebola virus reconstructed using electron tomography, *J. Vet. Med. Sci.* 67 (2005) 325–328.
- [22] T.A. Bharat, T. Noda, J.D. Riches, V. Kraehling, L. Kolesnikova, S. Becker, et al., Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography, *Proc. Natl. Acad. Sci. U. S. A.* 109 (2012) 4275–4280.
- [23] W. Wan, L. Kolesnikova, M. Clarke, A. Koehler, T. Noda, S. Becker, et al., Structure and assembly of the Ebola virus nucleocapsid, *Nature.* 551 (2017) 394–397.
- [24] Y. Takamatsu, L. Kolesnikova, S. Becker, Ebola virus proteins NP, VP35, and VP24 are essential and sufficient to mediate nucleocapsid transport, *Proc. Natl. Acad. Sci. U. S. A.* 115 (5) (2018) 1075–1080.
- [25] M. Mateo, C. Carbonnelle, M.J. Martinez, O. Reynard, A. Page, V.A. Volchkova, et al., Knockdown of Ebola virus VP24 impairs viral nucleocapsid assembly and prevents virus replication, *J. Infect. Dis.* 204 (Suppl. 3) (2011) S892–S896.
- [26] T. Hoenen, A. Groseth, L. Kolesnikova, S. Theriault, H. Ebihara, B. Hartlieb, et al., Infection of naive target cells with virus-like particles: implications for the function of ebola virus VP24, *J. Virol.* 80 (2006) 7260–7264.
- [27] D.R. Beniac, P.L. Melito, S.L. Devarenes, S.L. Hiebert, M.J. Rabb, L.L. Lamboo, et al., The organisation of Ebola virus reveals a capacity for extensive, modular polyploidy, *PLoS One.* 7 (2012), e29608.
- [28] Y. Sugita, H. Matsunami, Y. Kawaoka, T. Noda, M. Wolf, Cryo-EM structure of the Ebola virus nucleoprotein-RNA complex at 3.6 Å resolution, *Nature.* 563 (2018) 137–140.
- [29] M.F. Saeed, A.A. Kolokoltsov, T. Albrecht, R.A. Davey, Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes, *PLoS Pathog.* 6 (2010).
- [30] A. Nanbo, M. Imai, S. Watanabe, T. Noda, K. Takahashi, G. Neumann, et al., Ebolavirus is internalized into host cells via macropinocytosis in a viral glycoprotein-dependent manner, *PLoS Pathog.* 6 (2010).
- [31] P. Aleksandrowicz, A. Marzi, N. Biedenkopf, N. Beimforde, S. Becker, T. Hoenen, et al., Ebola virus enters host cells by macropinocytosis and clathrin-mediated endocytosis, *J. Infect. Dis.* 204 (Suppl. 3) (2011) S957–S967.
- [32] K. Chandran, N.J. Sullivan, U. Felbor, S.P. Whelan, J.M. Cunningham, Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection, *Science.* 308 (2005) 1643–1645.
- [33] M. Cote, J. Misasi, T. Ren, A. Bruchez, K. Lee, C.M. Filone, et al., Small molecule inhibitors reveal Niemann–Pick C1 is essential for Ebola virus infection, *Nature.* 477 (2011) 344–348.
- [34] J.E. Carette, M. Raaben, A.C. Wong, A.S. Herbert, G. Obermsterer, N. Mulherkar, et al., Ebola virus entry requires the cholesterol transporter Niemann–Pick C1, *Nature.* 477 (2011) 340–343.
- [35] J. Maruyama, H. Miyamoto, M. Kajihara, H. Ogawa, K. Maeda, Y. Sakoda, et al., Characterization of the envelope glycoprotein of a novel filovirus, Iloviu virus, *J. Virol.* 88 (2014) 99–109.
- [36] Ng M, Ndungo E, Jangra RK, Cai Y, Postnikova E, Radoshitzky SR, et al. Cell entry by a novel European filovirus requires host endosomal cysteine proteases and Niemann–Pick C1. *Virology.* 2014;468-470:637–46.
- [37] T. Hoenen, R.S. Shabman, A. Groseth, A. Herwig, M. Weber, G. Schudt, et al., Inclusion bodies are a site of ebolavirus replication, *J. Virol.* 86 (2012) 11779–11788.
- [38] H. Feldmann, E. Mühlberger, A. Randolph, C. Will, M.P. Kiley, A. Sanchez, et al., Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle, *Virus Res.* 24 (1992) 1–19.
- [39] A. Sanchez, M.P. Kiley, Identification and analysis of Ebola virus messenger RNA, *Virology.* 157 (1987) 414–420.
- [40] E. Mühlberger, Filovirus replication and transcription, *Future Virol.* 2 (2007) 205–215.
- [41] K. Brauburger, L.R. Deflubé, E. Mühlberger, Filovirus transcription and replication, in: A.K. Pattnaik, M.A. Whitt (Eds.), *Biology and Pathogenesis of Rhabdo- and Filoviruses*, World Scientific Publishing Co. Pte. Ltd., Singapore, 2015, pp. 515–5555.
- [42] A. Nanbo, S. Watanabe, P. Halfmann, Y. Kawaoka, The spatio-temporal distribution dynamics of Ebola virus proteins and RNA in infected cells, *Sci. Rep.* 3 (2013) 1206.
- [43] O. Dolnik, L. Stevermann, L. Kolesnikova, S. Becker, Marburg virus inclusions: a virus-induced microcompartment and interface to multivesicular bodies and the late endosomal compartment, *Eur. J. Cell Biol.* 94 (2015) 323–331.
- [44] E.V. Nelson, K.M. Schmidt, L.R. Deflube, S. Doganay, L. Banadyga, J. Olejnik, et al., Ebola virus does not induce stress granule formation during infection and sequesters stress granule proteins within viral inclusions, *J. Virol.* 90 (2016) 7268–7284.
- [45] G. Schudt, L. Kolesnikova, O. Dolnik, B. Sodeik, S. Becker, Live-cell imaging of Marburg virus-infected cells uncovers actin-dependent transport of nucleocapsids over long distances, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 14402–14407.
- [46] G. Schudt, O. Dolnik, L. Kolesnikova, N. Biedenkopf, A. Herwig, S. Becker, Transport of ebolavirus nucleocapsids is dependent on actin polymerization: live-cell imaging analysis of ebolavirus-infected cells, *J. Infect. Dis.* 212 (Suppl. 2) (2015) S160–S166.
- [47] K.K. Conzelmann, Reverse genetics of mononegavirales, *Curr. Top. Microbiol. Immunol.* 283 (2004) 1–41.
- [48] T. Hoenen, A. Groseth, F. de Kok-Mercado, J.H. Kuhn, V. Wahl-Jensen, Minigenomes, transcription and replication competent virus-like particles and beyond: reverse genetics systems for filoviruses and other negative stranded hemorrhagic fever viruses, *Antivir. Res.* 91 (2011) 195–208.
- [49] E. Mühlberger, B. Löffering, H.-D. Klenk, S. Becker, Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes, *J. Virol.* 72 (1998) 8756–8764.
- [50] E. Mühlberger, M. Weik, V.E. Volchkov, H.-D. Klenk, S. Becker, Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems, *J. Virol.* 73 (1999) 2333–2342.

- [51] Y. Boehmann, S. Enterlein, A. Randolph, E. Mühlberger, A reconstituted replication and transcription system for Ebola virus Reston and comparison with Ebola virus Zaire, *Virology*. 332 (2005) 406–417.
- [52] A. Groseth, H. Feldmann, S. Theriault, G. Mehmetoglu, R. Flick, RNA polymerase I-driven minigenome system for Ebola viruses, *J. Virol.* 79 (2005) 4425–4433.
- [53] W.A. Manhart, J.R. Pacheco, A.J. Hume, T.N. Cressey, L.R. DeFlube, E. Mühlberger, A chimeric Lloviu virus minigenome system reveals that the bat-derived filovirus replicates more similarly to ebolaviruses than marburgviruses, *Cell Rep.* 24 (2018) 2573–2580, e4.
- [54] L. Kamper, L. Zierke, M.L. Schmidt, A. Müller, L. Wendt, J. Brandt, et al., Assessment of the function and intergenus-compatibility of Ebola and Lloviu virus proteins, *J Gen Virol.* 100 (2019) 760–772.
- [55] K.M. Schmidt, E. Mühlberger, Marburg virus reverse genetics systems, *Viruses*. 8 (2016).
- [56] E.V. Nelson, J.R. Pacheco, A.J. Hume, T.N. Cressey, L.R. DeFlube, J.B. Ruedas, et al., An RNA polymerase II-driven Ebola virus minigenome system as an advanced tool for antiviral drug screening, *Antivir. Res.* 146 (2017) 21–27.
- [57] A.K. Pattnaik, L.A. Ball, A.W. LeGrone, G.W. Wertz, Infectious defective interfering particles of VSV from transcripts of a cDNA clone, *Cell.* 69 (1992) 1011–1020.
- [58] A. Ghanem, A. Kern, K.K. Conzelmann, Significantly improved rescue of rabies virus from cDNA plasmids, *Eur. J. Cell Biol.* 91 (2012) 10–16.
- [59] H. Yanai, Y. Hayashi, Y. Watanabe, N. Ohtaki, T. Kobayashi, Y. Nozaki, et al., Development of a novel Borna disease virus reverse genetics system using RNA polymerase II promoter and SV40 nuclear import signal, *Microbes Infect.* 8 (2006) 1522–1529.
- [60] L.R. DeFlube, T.N. Cressey, A.J. Hume, J. Olejnik, E. Haddock, F. Feldmann, et al., Ebolavirus polymerase uses an unconventional genome replication mechanism, *Proc Natl Acad Sci U S A* 116 (17) (2019) 8535–8543.
- [61] M. Weik, J. Modrof, H.D. Klenk, S. Becker, E. Mühlberger, Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation, *J. Virol.* 76 (2002) 8532–8539.
- [62] J.A. Alonso, J.L. Patterson, Sequence variability in viral genome non-coding regions likely contribute to observed differences in viral replication amongst MARV strains, *Virology*. 440 (2013) 51–63.
- [63] G. Neumann, S. Watanabe, Y. Kawaoka, Characterization of ebolavirus regulatory genomic regions, *Virus Res.* 144 (2009) 1–7.
- [64] K. Brauburger, Y. Boehmann, V. Krähling, E. Mühlberger, Transcriptional regulation in ebola virus: effects of gene border structure and regulatory elements on gene expression and polymerase scanning behavior, *J. Virol.* 90 (2015) 1898–1909.
- [65] K. Brauburger, Y. Boehmann, Y. Tsuda, T. Hoenen, J. Olejnik, M. Schumann, et al., Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation, *J. Virol.* 88 (2014) 12558–12571.
- [66] J. Wenigenrath, L. Kolesnikova, T. Hoenen, E. Mittler, S. Becker, Establishment and application of an infectious virus-like particle system for Marburg virus, *J Gen Virol.* 91 (2010) 1325–1334.
- [67] S. Watanabe, T. Watanabe, T. Noda, A. Takada, H. Feldmann, L.D. Jasenosky, et al., Production of novel ebola virus-like particles from cDNAs: an alternative to ebola virus generation by reverse genetics, *J. Virol.* 78 (2004) 999–1005.
- [68] A. Watt, F. Moukambi, L. Banadyga, A. Groseth, J. Callison, A. Herwig, et al., A novel life cycle modeling system for Ebola virus shows a genome length-dependent role of VP24 in virus infectivity, *J. Virol.* 88 (2014) 10511–10524.
- [69] T. Hoenen, H. Feldmann, Reverse genetics systems for filoviruses, *Methods Mol. Biol.* 2017 (1602) 159–170.
- [70] Pfaller CK, Cattaneo R, Schnell MJ. Reverse genetics of Mononegavirales: how they work, new vaccines, and new cancer therapeutics. *Virology*. 2015;479–480:331–44.
- [71] V.E. Volchkov, V.A. Volchkova, E. Mühlberger, L.V. Kolesnikova, M. Weik, O. Dolnik, et al., Recovery of infectious Ebola virus from complementary DNA: RNA editing of the GP gene and viral cytotoxicity, *Science*. 291 (2001) 1965–1969.
- [72] S. Enterlein, V. Volchkov, M. Weik, L. Kolesnikova, V. Volchkova, H.D. Klenk, et al., Rescue of recombinant Marburg virus from cDNA is dependent on nucleocapsid protein VP30, *J. Virol.* 80 (2006) 1038–1043.
- [73] A. Groseth, A. Marzi, T. Hoenen, A. Herwig, D. Gardner, S. Becker, et al., The Ebola virus glycoprotein contributes to but is not sufficient for virulence in vivo, *PLoS Pathog.* 8 (2012), e1002847.
- [74] T. Hoenen, J. Brandt, Y. Cai, J.H. Kuhn, C. Finch, Reverse genetics of filoviruses, *Curr. Top. Microbiol. Immunol.* 411 (2017) 421–445.
- [75] M. Weik, S. Enterlein, K. Schlenz, E. Mühlberger, The Ebola virus genomic replication promoter is bipartite and follows the rule of six, *J. Virol.* 79 (2005) 10660–10671.
- [76] S. Enterlein, K.M. Schmidt, M. Schumann, D. Conrad, V. Krähling, J. Olejnik, et al., The marburg virus 3' noncoding region structurally and functionally differs from that of ebola virus, *J. Virol.* 83 (2009) 4508–4519.
- [77] T. Hoenen, S. Jung, A. Herwig, A. Groseth, S. Becker, Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription, *Virology*. 403 (2010) 56–66.
- [78] M.E. Olsen, T.N. Cressey, E. Mühlberger, J.H. Connor, Differential mechanisms for the involvement of polyamines and hypusinated eIF5A in Ebola virus gene expression, *J. Virol.* 92 (2018).
- [79] P. Calain, M.C. Monroe, S.T. Nichol, Ebola virus defective interfering particles and persistent infection, *Virology*. 262 (1999) 114–128.
- [80] B. Morin, A.A. Rahmeh, S.P. Whelan, Mechanism of RNA synthesis initiation by the vesicular stomatitis virus polymerase, *EMBO J.* 31 (2012) 1320–1329.
- [81] S.L. Noton, L.R. DeFlube, C.Z. Tremaglio, R. Fearn, The respiratory syncytial virus polymerase has multiple RNA synthesis activities at the promoter, *PLoS Pathog.* 8 (2012), e1002980.
- [82] B. Morin, B. Liang, E. Gardner, R.A. Ross, S.P.J. Whelan, An in vitro RNA synthesis assay for rabies virus defines ribonucleoprotein interactions critical for polymerase activity, *J. Virol.* 91 (2017).
- [83] P.C. Jordan, C. Liu, P. Raynaud, M.K. Lo, C.F. Spiropoulou, J.A. Symons, et al., Initiation, extension, and termination of RNA synthesis by a paramyxovirus polymerase, *PLoS Pathog.* 14 (2018), e1006889.
- [84] S.M. Crary, J.S. Towner, J.E. Honig, T.R. Shoemaker, S.T. Nichol, Analysis of the role of predicted RNA

- secondary structures in Ebola virus replication, *Virology*. 306 (2003) 210–218.
- [85] J. Sztuba-Solinska, L. Diaz, M.R. Kumar, G. Kolb, M.R. Wiley, L. Jozwick, et al., A small stem-loop structure of the Ebola virus trailer is essential for replication and interacts with heat-shock protein A8, *Nucleic Acids Res.* 44 (20) (2016) 9831–9846.
- [86] M.P. Kiley, J. Wilusz, J.B. McCormick, J.D. Keene, Conservation of the 3' terminal nucleotide sequences of Ebola and Marburg virus, *Virology*. 149 (1986) 251–254.
- [87] M. Schlee, A. Roth, V. Hornung, C.A. Hagmann, V. Wimmenauer, W. Barchet, et al., Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus, *Immunity*. 31 (2009) 25–34.
- [88] Gerlier D, Lyles DS. Interplay between innate immunity and negative-strand RNA viruses: towards a rational model. *Microbiol Mol Biol Rev.* 2011;75:468–90, second page of table of contents.
- [89] P. Calain, L. Roux, The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA, *J. Virol.* 67 (1993) 4822–4830.
- [90] C. Tapparel, D. Maurice, L. Roux, The activity of Sendai virus genomic and antigenomic promoters requires a second element past the leader template regions: a motif (GNNNNN)₃ is essential for replication, *J. Virol.* 72 (1998) 3117–3128.
- [91] D. Kolakofsky, Paramyxovirus RNA synthesis, mRNA editing, and genome hexamer phase: a review, *Virology*. 498 (2016) 94–98.
- [92] P. le Mercier, D. Kolakofsky, Bipartite promoters and RNA editing of paramyxoviruses and filoviruses, *RNA*. 25 (2019) 279–285.
- [93] E. Mühlberger, A. Sanchez, A. Randolph, C. Will, M.P. Kiley, H.D. Klenk, et al., The nucleotide sequence of the L gene of Marburg virus, a filovirus: homologies with paramyxoviruses and rhabdoviruses, *Virology*. 187 (1992) 534–547.
- [94] V.E. Volchkov, V.A. Volchkova, A.A. Chepurinov, V.M. Blinov, O. Dolnik, S.V. Netesov, et al., Characterization of the L gene and 5' trailer region of Ebola virus, *J Gen Virol.* 80 (1999) 355–362.
- [95] P. Le Mercier, D. Garcin, E. Garcia, D. Kolakofsky, Competition between the Sendai virus N mRNA start site and the genome 3'-end promoter for viral RNA polymerase, *J. Virol.* 77 (2003) 9147–9155.
- [96] D.R. McGivern, P.L. Collins, R. Fearn, Identification of internal sequences in the 3' leader region of human respiratory syncytial virus that enhance transcription and confer replication processivity, *J. Virol.* 79 (2005) 2449–2460.
- [97] R. Fearn, P.L. Collins, M.E. Peeples, Functional analysis of the genomic and antigenomic promoters of human respiratory syncytial virus, *J. Virol.* 74 (2000) 6006–6014.
- [98] S. Finke, K.K. Conzelmann, Ambisense gene expression from recombinant rabies virus: random packaging of positive- and negative-strand ribonucleoprotein complexes into rabies virions, *J. Virol.* 71 (1997) 7281–7288.
- [99] E. Mühlberger, S. Trommer, C. Funke, V. Volchkov, H.-D. Klenk, S. Becker, Termini of all mRNA species of Marburg virus: sequence and secondary structure, *Virology*. 223 (1996) 376–380.
- [100] R.S. Shabman, O.J. Jabado, C.E. Mire, T.B. Stockwell, M. Edwards, M. Mahajan, et al., Deep sequencing identifies noncanonical editing of Ebola and Marburg virus RNAs in infected cells, *MBio*. 5 (2014), e02011.
- [101] A. Sanchez, M.P. Kiley, B.P. Holloway, D.D. Auperin, Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus, *Virus Res.* 29 (1993) 215–240.
- [102] L.A. Ball, G.W. Wertz, VSV RNA synthesis: how can you be positive? *Cell*. 26 (1981) 143–144.
- [103] R.S. Shabman, T. Hoenen, A. Groseth, O. Jabado, J.M. Binning, G.K. Amarasinghe, et al., An upstream open reading frame modulates ebola virus polymerase translation and virus replication, *PLoS Pathog.* 9 (2013), e1003147.
- [104] L.H. Elliott, M.P. Kiley, J.B. McCormick, Descriptive analysis of Ebola virus proteins, *Virology*. 147 (1985) 169–176.
- [105] S. Becker, C. Rinne, U. Hofsäss, H.-D. Klenk, E. Mühlberger, Interactions of Marburg virus nucleocapsid proteins, *Virology*. 249 (1998) 406–417.
- [106] J. Modrof, C. Moritz, L. Kolesnikova, T. Konakova, B. Hartlieb, A. Randolph, et al., Phosphorylation of Marburg virus VP30 at serines 40 and 42 is critical for its interaction with NP inclusions, *Virology*. 287 (2001) 171–182.
- [107] B. Hartlieb, T. Muziol, W. Weissenhorn, S. Becker, Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 624–629.
- [108] C.F. Basler, N.J. Krogan, D.W. Leung, G.K. Amarasinghe, Virus and host interactions critical for filoviral RNA synthesis as therapeutic targets, *Antivir. Res.* 162 (2018) 90–100.
- [109] E.P. Tchesnokov, P. Raesimakiani, M. Ngure, D. Marchant, M. Gotte, Recombinant RNA-dependent RNA polymerase complex of Ebola virus, *Sci. Rep.* 8 (2018) 3970.
- [110] T.J. Green, R. Cox, J. Tsao, M. Rowse, S. Qiu, M. Luo, Common mechanism for RNA encapsidation by negative-strand RNA viruses, *J. Virol.* 88 (2014) 3766–3775.
- [111] Y. Sun, Y. Guo, Z. Lou, A versatile building block: the structures and functions of negative-sense single-stranded RNA virus nucleocapsid proteins, *Protein & cell*. 3 (2012) 893–902.
- [112] S. Watanabe, T. Noda, Y. Kawaoka, Functional mapping of the nucleoprotein of Ebola virus, *J. Virol.* 80 (2006) 3743–3751.
- [113] T. Noda, H. Ebihara, Y. Muramoto, K. Fujii, A. Takada, H. Sagara, et al., Assembly and budding of Ebolavirus, *PLoS Pathog.* 2 (2006), e99.
- [114] T. Noda, K. Hagiwara, H. Sagara, Y. Kawaoka, Characterization of the Ebola virus nucleoprotein–RNA complex, *J Gen Virol.* 91 (2010) 1478–1483.
- [115] A. DiCarlo, P. Moller, A. Lander, L. Kolesnikova, S. Becker, Nucleocapsid formation and RNA synthesis of Marburg virus is dependent on two coiled coil motifs in the nucleoprotein, *Virol. J.* 4 (2007) 105.
- [116] B. Liu, S. Dong, G. Li, W. Wang, X. Liu, Y. Wang, et al., Structural insight into nucleoprotein conformation change chaperoned by VP35 peptide in Marburg virus, *J. Virol.* 91 (2017).
- [117] T. Zhu, H. Song, R. Peng, Y. Shi, J. Qi, G.F. Gao, Crystal structure of the Marburg virus nucleoprotein core domain chaperoned by a VP35 peptide reveals a conserved drug target for filovirus, *J. Virol.* 91 (2017).
- [118] Z. Su, C. Wu, L. Shi, P. Luthra, G.D. Pintilie, B. Johnson, et al., Electron cryo-microscopy structure of Ebola virus nucleoprotein reveals a mechanism for nucleocapsid-like assembly, *Cell*. 172 (2018) 966–978, e12.
- [119] R.N. Kirchdoerfer, D.M. Abelson, S. Li, M.R. Wood, E.O. Saphire, Assembly of the Ebola virus nucleoprotein from a chaperoned VP35 complex, *Cell Rep.* 12 (2015) 140–149.

- [120] D.W. Leung, D. Borek, P. Luthra, J.M. Binning, M. Anantpadma, G. Liu, et al., An intrinsically disordered peptide from Ebola virus VP35 controls viral RNA synthesis by modulating nucleoprotein–RNA interactions, *Cell Rep.* 11 (2015) 376–389.
- [121] S. Dong, P. Yang, G. Li, B. Liu, W. Wang, X. Liu, et al., Insight into the Ebola virus nucleocapsid assembly mechanism: crystal structure of Ebola virus nucleoprotein core domain at 1.8 Å resolution, *Protein & Cell.* 6 (2015) 351–362.
- [122] K.C. Prins, J.M. Binning, R.S. Shabman, D.W. Leung, G.K. Amarasinghe, C.F. Basler, Basic residues within the ebolavirus VP35 protein are required for its viral polymerase cofactor function, *J. Virol.* 84 (2010) 10581–10591.
- [123] S.P. Reid, W.B. Cardenas, C.F. Basler, Homo-oligomerization facilitates the interferon-antagonist activity of the ebolavirus VP35 protein, *Virology.* 341 (2005) 179–189.
- [124] P. Luthra, D.S. Jordan, D.W. Leung, G.K. Amarasinghe, C.F. Basler, Ebola virus VP35 interaction with dynein LC8 regulates viral RNA synthesis, *J. Virol.* 89 (2015) 5148–5153.
- [125] M.R. Edwards, G. Liu, C.E. Mire, S. Sureshchandra, P. Luthra, B. Yen, et al., Differential regulation of interferon responses by Ebola and Marburg virus VP35 proteins, *Cell Rep.* 14 (2016) 1632–1640.
- [126] J.F. Bruhn, R.N. Kirchdoerfer, S.M. Urata, S. Li, I.J. Tickle, G. Bricogne, et al., Crystal structure of the Marburg virus VP35 oligomerization domain, *J. Virol.* 91 (2017).
- [127] P. Möller, N. Pariante, H.D. Klenk, S. Becker, Homo-oligomerization of Marburgvirus VP35 is essential for its function in replication and transcription, *J. Virol.* 79 (2005) 14876–14886.
- [128] M. Trunschke, D. Conrad, S. Enterlein, J. Olejnik, K. Brauburger, E. Mühlberger, The L–VP35 and L–L interaction domains reside in the amino terminus of the Ebola virus L protein and are potential targets for antivirals, *Virology.* 441 (2013) 135–145.
- [129] A. Groseth, J.E. Charton, M. Sauerborn, F. Feldmann, S.M. Jones, T. Hoenen, et al., The Ebola virus ribonucleoprotein complex: a novel VP30–L interaction identified, *Virus Res.* 140 (2009) 8–14.
- [130] O. Poch, B.M. Blumberg, L. Bougueleret, N. Tordo, Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains, *J Gen Virol.* 71 (Pt 5) (1990) 1153–1162.
- [131] M.L. Schmidt, T. Hoenen, Characterization of the catalytic center of the Ebola virus L polymerase, *PLoS Negl. Trop. Dis.* 11 (2017), e0005996.
- [132] B. Martin, B. Coutard, T. Guez, G.C. Paesen, B. Canard, F. Debart, et al., The methyltransferase domain of the Sudan ebolavirus L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure, *Nucleic Acids Res.* 46 (2018) 7902–7912.
- [133] C.G. Albarino, L.S. Uebelhoer, J.P. Vincent, M.L. Khristova, A.K. Chakrabarti, A. McElroy, et al., Development of a reverse genetics system to generate recombinant Marburg virus derived from a bat isolate, *Virology.* 446 (2013) 230–237.
- [134] D.A. Marston, L.M. McElhinney, N. Johnson, T. Muller, K.K. Conzelmann, N. Tordo, et al., Comparative analysis of the full genome sequence of European bat lyssaviruses type 1 and type 2 with other lyssaviruses and evidence for a conserved transcription termination and polyadenylation motif in the G-L 3′ non-translated region, *J Gen Virol.* 88 (2007) 1302–1314.
- [135] B. Morin, P.J. Kranzusch, A.A. Rahmeh, S.P. Whelan, The polymerase of negative-stranded RNA viruses, *Curr Opin Virol.* 3 (2013) 103–110.
- [136] S. Bamberg, L. Kolesnikova, P. Möller, H.D. Klenk, S. Becker, VP24 of Marburg virus influences formation of infectious particles, *J. Virol.* 79 (2005) 13421–13433.
- [137] T. Noda, P. Halfmann, H. Sagara, Y. Kawaoka, Regions in Ebola virus VP24 that are important for nucleocapsid formation, *J. Infect. Dis.* 196 (Suppl. 2) (2007) S247–S250.
- [138] L. Banadyga, T. Hoenen, X. Ambroggio, E. Dunham, A. Groseth, H. Ebihara, Ebola virus VP24 interacts with NP to facilitate nucleocapsid assembly and genome packaging, *Sci. Rep.* 7 (2017) 7698.
- [139] S. Watanabe, T. Noda, P. Halfmann, L. Jasenosky, Y. Kawaoka, Ebola virus (EBOV) VP24 inhibits transcription and replication of the EBOV genome, *J. Infect. Dis.* 196 (Suppl. 2) (2007) S284–S290.
- [140] Z.A. Bornholdt, T. Noda, D.M. Abelson, P. Halfmann, M.R. Wood, Y. Kawaoka, et al., Structural rearrangement of ebola virus VP40 begets multiple functions in the virus life cycle, *Cell.* 154 (2013) 763–774.
- [141] L. Kolesnikova, E. Mühlberger, E. Ryabchikova, S. Becker, Ultrastructural organization of recombinant Marburg virus nucleoprotein: comparison with Marburg virus inclusions, *J. Virol.* 74 (2000) 3899–3904.
- [142] J. Olejnik, E. Ryabchikova, R.B. Corley, E. Mühlberger, Intracellular events and cell fate in filovirus infection, *Viruses.* 3 (2011) 1501–1531.
- [143] J. Nikolic, R. Le Bars, Z. Lama, N. Scrima, C. Lagaudriere-Gesbert, Y. Gaudin, et al., Negri bodies are viral factories with properties of liquid organelles, *Nat. Commun.* 8 (2017) 58.
- [144] B.S. Heinrich, Z. Maliga, D.A. Stein, A.A. Hyman, S.P.J. Whelan, Phase transitions drive the formation of vesicular stomatitis virus replication compartments, *MBio.* 9 (2018).
- [145] K. Das, E. Arnold, Negative-strand RNA virus L proteins: one machine, many activities, *Cell.* 162 (2015) 239–241.
- [146] T. Ogino, S.P. Yadav, A.K. Banerjee, Histidine-mediated RNA transfer to GDP for unique mRNA capping by vesicular stomatitis virus RNA polymerase, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 3463–3468.
- [147] D. Testa, A.K. Banerjee, Two methyltransferase activities in the purified virions of vesicular stomatitis virus, *J. Virol.* 24 (1977) 786–793.
- [148] J. Li, J.T. Wang, S.P. Whelan, A unique strategy for mRNA cap methylation used by vesicular stomatitis virus, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 8493–8498.
- [149] G. Tekes, A.A. Rahmeh, S.P. Whelan, A freeze frame view of vesicular stomatitis virus transcription defines a minimal length of RNA for 5′ processing, *PLoS Pathog.* 7 (2011), e1002073.
- [150] V.E. Volchkov, S. Becker, V.A. Volchkova, V.A. Ternovoj, A.N. Kotov, S.V. Netesov, et al., GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases, *Virology.* 214 (1995) 421–430.
- [151] A. Sanchez, S.G. Trappier, B.W. Mahy, C.J. Peters, S.T. Nichol, The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing, *Proc. Natl. Acad. Sci. U. S. A.* 93 (1996) 3602–3607.
- [152] M. Mehedi, D. Falzarano, J. Seebach, X. Hu, M.S. Carpenter, H.J. Schnittler, et al., A new Ebola virus nonstructural glycoprotein expressed through RNA editing, *J Virol* 85 (11) (2011) 5406–5414.

- [153] K.J. Alfson, L.E. Avena, M.W. Beadles, H. Menzie, J.L. Patterson, R. Carrion Jr., et al., Genetic changes at the glycoprotein editing site associated with serial passage of Sudan virus, *J. Infect. Dis.* 212 (Suppl. 2) (2015) S295–S304.
- [154] V.A. Volchkova, H. Feldmann, H.D. Klenk, V.E. Volchkov, The nonstructural small glycoprotein sGP of Ebola virus is secreted as an antiparallel-orientated homodimer, *Virology*. 250 (1998) 408–414.
- [155] V.A. Volchkova, O. Dolnik, M.J. Martinez, O. Reynard, V.E. Volchkov, RNA editing of the GP gene of Ebola virus is an important pathogenicity factor, *J. Infect. Dis.* 212 (Suppl. 2) (2015) S226–S233.
- [156] V.A. Volchkova, J. Vorac, L. Repiquet-Paire, P. Lawrence, V.E. Volchkov, Ebola virus GP gene polyadenylation versus RNA editing, *J. Infect. Dis.* 212 (Suppl 2) (2015) S191–S198.
- [157] M. Mehedi, T. Hoenen, S. Robertson, S. Ricklefs, M.A. Dolan, T. Taylor, et al., Ebola virus RNA editing depends on the primary editing site sequence and an upstream secondary structure, *PLoS Pathog.* 9 (2013), e1003677.
- [158] V.A. Volchkova, O. Dolnik, M.J. Martinez, O. Reynard, V.E. Volchkov, Genomic RNA editing and its impact on Ebola virus adaptation during serial passages in cell culture and infection of guinea pigs, *J. Infect. Dis.* 204 (Suppl. 3) (2011) S941–S946.
- [159] J.R. Kugelman, M.S. Lee, C.A. Rossi, S.E. McCarthy, S.R. Radoshitzky, J.M. Dye, et al., Ebola virus genome plasticity as a marker of its passaging history: a comparison of *in vitro* passaging to non-human primate infection, *PLoS One*. 7 (2012), e50316.
- [160] M.L. Blondot, V. Dubosclard, J. Fix, S. Lassoued, M. Aumont-Nicaise, F. Bontems, et al., Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein, *PLoS Pathog.* 8 (2012), e1002734.
- [161] P.L. Collins, M.G. Hill, J. Cristina, H. Grosfeld, Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus, *Proc. Natl. Acad. Sci. U. S. A.* 93 (1996) 81–85.
- [162] P.L. Collins, J.A. Melerio, Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years, *Virus Res.* 162 (2011) 80–99.
- [163] R. Fearn, P.L. Collins, Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription, *J. Virol.* 73 (1999) 5852–5864.
- [164] J. Modrof, E. Mühlberger, H.D. Klenk, S. Becker, Phosphorylation of VP30 impairs Ebola virus transcription, *J. Biol. Chem.* 277 (2002) 33099–33104.
- [165] B. Hartlieb, J. Modrof, E. Mühlberger, H.D. Klenk, S. Becker, Oligomerization of Ebola virus VP30 is essential for viral transcription and can be inhibited by a synthetic peptide, *J. Biol. Chem.* 278 (2003) 41830–41836.
- [166] S.P. John, T. Wang, S. Steffan, S. Longhi, C.S. Schmaljohn, C.B. Jonsson, The Ebola virus VP30 is an RNA binding protein, *J. Virol.* 81 (17) (2007) 8967–8976.
- [167] N. Biedenkopf, B. Hartlieb, T. Hoenen, S. Becker, Phosphorylation of Ebola virus VP30 influences the composition of the viral nucleocapsid complex: Impact on viral transcription and replication, *J. Biol. Chem.* 288 (16) (2013) 11165–11174.
- [168] N. Biedenkopf, J. Schlereth, A. Grunweller, S. Becker, R.K. Hartmann, RNA-binding of Ebola virus VP30 is essential for activating viral transcription, *J. Virol.* 90 (16) (2016) 7481–7496.
- [169] J. Schlereth, A. Grunweller, N. Biedenkopf, S. Becker, R.K. Hartmann, RNA binding specificity of Ebola virus transcription factor VP30, *RNA Biol.* 13 (2016) 783–798.
- [170] R.N. Kirchdoerfer, C.L. Moyer, D.M. Abelson, E.O. Saphire, The Ebola virus VP30–NP interaction is a regulator of viral RNA synthesis, *PLoS Pathog.* 12 (2016), e1005937.
- [171] W. Xu, P. Luthra, C. Wu, J. Batra, D.W. Leung, C.F. Basler, et al., Ebola virus VP30 and nucleoprotein interactions modulate viral RNA synthesis, *Nat. Commun.* 8 (2017) 15576.
- [172] J. Modrof, S. Becker, E. Mühlberger, Ebola virus transcription activator VP30 is a zinc-binding protein, *J. Virol.* 77 (2003) 3334–3338.
- [173] M.J. Martinez, V.A. Volchkova, H. Raoul, N. Alazard-Dany, O. Reynard, V.E. Volchkov, Role of VP30 phosphorylation in the Ebola virus replication cycle, *J. Infect. Dis.* 204 (Suppl. 3) (2011) S934–S940.
- [174] M.J. Martinez, N. Biedenkopf, V. Volchkova, B. Hartlieb, N. Alazard-Dany, O. Reynard, et al., Role of Ebola virus VP30 in transcription reinitiation, *J. Virol.* 82 (2008) 12569–12573.
- [175] P.A. Ilinykh, B. Tigabu, A. Ivanov, T. Ammosova, Y. Obukhov, T. Garron, et al., Role of protein phosphatase 1 in dephosphorylation of Ebola virus VP30 protein and its targeting for the inhibition of viral transcription, *J. Biol. Chem.* 289 (2014) 22723–22738.
- [176] N. Biedenkopf, C. Lier, S. Becker, Dynamic phosphorylation of VP30 is essential for Ebola virus life cycle, *J. Virol.* 90 (10) (2016) 4914–4925.
- [177] T. Kruse, N. Biedenkopf, E.P.T. Hertz, E. Dietzel, G. Stalman, B. Lopez-Mendez, et al., The Ebola virus nucleoprotein recruits the host PP2A-B56 phosphatase to activate transcriptional support activity of VP30, *Mol. Cell* 69 (2018) 136–145, e6.
- [178] J. Batra, J.F. Hultquist, D. Liu, O. Shtanko, J. Von Dollen, L. Satkamp, et al., Protein interaction mapping identifies RBBP6 as a negative regulator of Ebola virus replication, *Cell.* 175 (2018) 1917–1930, e13.
- [179] B. Tigabu, P. Ramanathan, A. Ivanov, X. Lin, P.A. Ilinykh, C.S. Parry, et al., Phosphorylated Vp30 of Marburg virus is a repressor of transcription, *J. Virol.* 92 (21) (2018), e00426-18.
- [180] T. Fowler, S. Bamberg, P. Möller, H.D. Klenk, T.F. Meyer, S. Becker, et al., Inhibition of Marburg virus protein expression and viral release by RNA interference, *J Gen Virol.* 86 (2005) 1181–1188.
- [181] J. Olejnik, A.J. Hume, D.W. Leung, G.K. Amarasinghe, C.F. Basler, E. Mühlberger, Filovirus strategies to escape antiviral responses, *Curr Top Microbiol Immunol* 411 (2017) 293–322.
- [182] A.L. Hartman, J.S. Towner, S.T. Nichol, A C-terminal basic amino acid motif of Zaire ebolavirus VP35 is essential for type I interferon antagonism and displays high identity with the RNA-binding domain of another interferon antagonist, the NS1 protein of influenza A virus, *Virology.* 328 (2004) 177–184.
- [183] D.W. Leung, N.D. Ginder, D.B. Fulton, J. Nix, C.F. Basler, R.B. Honzatko, et al., Structure of the Ebola VP35 interferon inhibitory domain, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 411–416.
- [184] C.R. Kimberlin, Z.A. Bornholdt, S. Li, V.L. Woods Jr., I.J. MacRae, E.O. Saphire, Ebolavirus VP35 uses a

- bimodal strategy to bind dsRNA for innate immune suppression, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 314–319.
- [185] D.W. Leung, R.S. Shabman, M. Farahbakhsh, K.C. Prins, D.M. Borek, T. Wang, et al., Structural and functional characterization of Reston Ebola virus VP35 interferon inhibitory domain, *J. Mol. Biol.* 399 (2010) 347–357.
- [186] P. Ramanan, M.R. Edwards, R.S. Shabman, D.W. Leung, A.C. Endlich-Frazier, D.M. Borek, et al., Structural basis for Marburg virus VP35-mediated immune evasion mechanisms, *Proc. Natl. Acad. Sci. U. S. A.* 109 (2012) 20661–20666.
- [187] S. Bale, J.P. Julien, Z.A. Bornholdt, C.R. Kimberlin, P. Halfmann, M.A. Zandonatti, et al., Marburg virus VP35 can both fully coat the backbone and cap the ends of dsRNA for interferon antagonism, *PLoS Pathog.* 8 (2012), e1002916.
- [188] R.N. Harty, M.E. Brown, G. Wang, J. Huibregtse, F.P. Hayes, A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding, *Proc. Natl. Acad. Sci. U. S. A.* 97 (2000) 13871–13876.
- [189] O. Dolnik, L. Kolesnikova, S. Welsch, T. Strecker, G. Schudt, S. Becker, Interaction with Tsg101 is necessary for the efficient transport and release of nucleocapsids in marburg virus-infected cells, *PLoS Pathog.* 10 (2014), e1004463.
- [190] S. Urata, T. Noda, Y. Kawaoka, S. Morikawa, H. Yokosawa, J. Yasuda, Interaction of Tsg101 with Marburg virus VP40 depends on the PPPY motif, but not the PT/SAP motif as in the case of Ebola virus, and Tsg101 plays a critical role in the budding of Marburg virus-like particles induced by VP40, NP, and GP, *J. Virol.* 81 (2007) 4895–4899.
- [191] E.O. Freed, Viral late domains, *J. Virol.* 76 (2002) 4679–4687.
- [192] O. Dolnik, L. Kolesnikova, L. Stevermann, S. Becker, protein to support budding of Marburg virus-like particles, *J. Virol.* 84 (2010) 7847–7856.
- [193] S.P. Reid, L.W. Leung, A.L. Hartman, O. Martinez, M.L. Shaw, C. Carbonnelle, et al., Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation, *J. Virol.* 80 (2006) 5156–5167.
- [194] S.P. Reid, C. Valmas, O. Martinez, F.M. Sanchez, C.F. Basler, Ebola virus VP24 proteins inhibit the interaction of NPI-1 subfamily karyopherin alpha proteins with activated STAT1, *J. Virol.* 81 (2007) 13469–13477.
- [195] M. Mateo, S.P. Reid, L.W. Leung, C.F. Basler, V.E. Volchkov, Ebolavirus VP24 binding to karyopherins is required for inhibition of interferon signaling, *J. Virol.* 84 (2010) 1169–1175.
- [196] W. Xu, M.R. Edwards, D.M. Borek, A.R. Feagins, A. Mittal, J.B. Alinger, et al., Ebola virus VP24 targets a unique NLS binding site on karyopherin alpha 5 to selectively compete with nuclear import of phosphorylated STAT1, *Cell Host Microbe* 16 (2014) 187–200.
- [197] M.R. Edwards, B. Johnson, C.E. Mire, W. Xu, R.S. Shabman, L.N. Speller, et al., The Marburg virus VP24 protein interacts with Keap1 to activate the cytoprotective antioxidant response pathway, *Cell Rep.* 6 (2014) 1017–1025.
- [198] A. Page, V.A. Volchkova, S.P. Reid, M. Mateo, A. Bagnaud-Baule, K. Nemirov, et al., Marburgvirus hijacks nrf2-dependent pathway by targeting nrf2-negative regulator keap1, *Cell Rep.* 6 (2014) 1026–1036.
- [199] M.E. Olsen, C.M. Filone, D. Rozelle, C.E. Mire, K.N. Agans, L. Hensley, et al., Polyamines and hypusination are required for ebolavirus gene expression and replication, *MBio.* 7 (2016).
- [200] A.R. Feagins, C.F. Basler, Lloviu virus VP24 and VP35 proteins function as innate immune antagonists in human and bat cells, *Virology.* 485 (2015) 145–152.