

Copyright

by

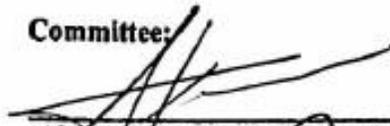
Adam James Dean Ronk

2020

The Dissertation Committee for Adam James Dean Ronk Certifies that this is the approved version of the following dissertation:

Description of interactions of Ebola virus with a putative reservoir species using next-generation sequencing

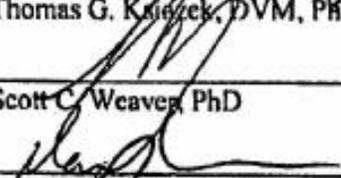
Committee:



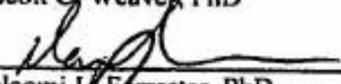
Alexander Bukreyev, PhD, M.S.
Supervisor, Chair



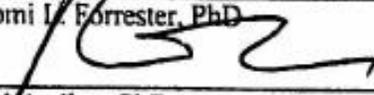
Thomas G. Kuznetsov, DVM, PhD



Scott C. Weaver, PhD



Naomi L. Forrester, PhD



Raul Andino, PhD

**Description of interactions of Ebola virus with a putative reservoir
species using next-generation sequencing**

by

Adam James Dean Ronk, B.A.

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas Medical Branch
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

**The University of Texas Medical Branch
March, 2020**

Acknowledgements

I would like to thank my mentor, Dr. Alexander Bukreyev and the remainder of my dissertation committee, along with the members of the Bukreyev and Andino laboratories, and in particular Drs. Abhishek Prasad, Ivan Kuzmin, Zach Whitfield, and Philipp Ilinykh, with whom the publications this dissertation was based upon were co-authored. Acknowledgement must also be given to Drs. Bryan Cullen, Ingo Jordan, Christian Drosten, and Gregory Ebel, who provided reagents or computational resources for this work. The work described here was funded by the National Science Foundation (MCB-151668, awarded to Raul Andino and Alexander Bukreyev), and the Defense Threat Reduction Agency (HDTRA1-14-1-001, awarded to Dr. Christopher Basler, Georgia State University and Alexander Bukreyev).

Description of interactions of Ebola virus with a putative reservoir species using next-generation sequencing

Publication No. _____

Adam James Dean Ronk, PhD

The University of Texas Medical Branch, 2020

Supervisor: Alexander Bukreyev

Ebola virus (EBOV) causes a severe, often fatal disease in humans and nonhuman primates. Recently, EBOV has caused two very large outbreaks, one of which is ongoing in the Democratic Republic of the Congo. Bats are the likely reservoir of EBOV, but little is known of their relationship with the virus. Next-generation sequencing has become an extremely powerful and flexible tool in virology over the past decade as new library preparation techniques have been developed that permit the selective sequencing of small RNAs, and the characterization of entire viral populations at incredible levels of detail. For this work, I exploited this technology to explore two aspects of the bat/virus nexus; namely the small RNA profile of infection, and the evolution of the virus in bat cells. The biology of the virus in human cells was used for comparison. Here I describe a new class of small noncoding RNAs produced by EBOV during infection of bat and human cells that resemble microRNAs, but are not associated with the microRNA machinery, and lack any discernable RNAi function. I also describe the evolution of EBOV in an experimental passage series in bat and human cells. This work led to the discovery of a potential role for host RNA editing enzymes in the evolution of EBOV in bats, and identified loci within the viral genome that appear to be associated with adaptation to human cells.

TABLE OF CONTENTS

List of Tables	xi
List of Figures	xii
List of Illustrations	xv
List of Abbreviations	xvi
INTRODUCTION.....	17
Ebola virus: Virology, epidemiology, ecology, and evolution	17
Basic virology	17
Virion and genome structure.....	17
Gene function.....	19
Attachment, entry, and uncoating	21
Replication and life cycle.....	22
History and epidemiology.....	26
A brief history of Ebola virus, 1976-2020	26
Epidemiology	29
Medical aspects of Ebola virus infection.....	33
Course of infection and pathology	33
Clinical presentation and treatment	37
Ecology and evolution	41
Ebola virus ecology: An ongoing mystery.....	41
Evolution and evolutionary history.....	45
DESCRIPTION OF THE INTERACTIONS OF EBOLA VIRUS WITH A PUTATIVE RESERVOIR SPECIES USING NEXT-GENERATION SEQUENCING.....	48
Rationale	48
Ebola virus produces discrete small non-coding RNAs independent of the host microRNA pathway and which lack RNA interference activity in bat and human cells	49
Introduction.....	49
Micro-RNA biology.....	49

RNA interference and its role in viral biology.....	51
Ebola virus and small RNAs.....	52
Materials and Methods.....	54
Cell lines and maintenance	54
Viruses	54
Experimental infections	55
Virus Inactivation and RNA Extraction with Trizol.....	56
Small RNA Library Preparation	56
Bioinformatics.....	57
miRNA-specific RT-qPCR	58
Detection of viral vRNA/cRNA/mRNA by TaqMan (hydrolysis probe) RT-qPCR	60
RNA-binding protein Immunoprecipitation (RIP) Assay.....	60
siRNA-mediated knockdown.....	61
Western blot.....	61
Focus-forming/plaque immunostaining assays.....	62
Luciferase reporter assays.....	62
Results.....	63
Virus-derived sRNA profiles of filovirus-infected bat and human cell lines	63
EBOV vncRNA production does not require host miRNA machinery ...	80
EBOV vncRNAs are not associated with Argonaute/RISC, and do not have an effect on virus replication.....	85
EBOV vncRNAs lack the ability to effectively suppress a reporter transcript with multiple miRNA-binding elements	88
Discussion.....	91
Species-specific evolution of Ebola virus during replication in human and bat cells.....	106
Introduction.....	106
Viral evolution in context	106
Viral evolution as a multilevel phenomenon	107
The origins of diversity in viral evolution	108
Mechanisms of viral evolution: Drift and selection.....	108

Genetic drift	108
Natural selection	109
Comparing the roles of drift and selection in viral evolution	110
Selection above the level of the individual: Hypotheses and basic mechanisms.....	111
Group selection and kin selection	111
Intergenomic interactions	111
High mutation rates and their consequences.....	112
The error-prone nature of RNA virus replication	112
Intrahost genetic diversity as an evolutionary strategy.....	114
Quasispecies theory	116
Host effects on EBOV evolution	118
Host RNA editing enzymes and their effects on viral evolution	119
Prior work	120
Materials and methods.....	121
Experimental passages	121
Sequencing and processing.....	122
Determination of population mutation frequency.....	122
Identification of ‘ADAR’ motif.....	123
Average number of mutations per read.....	123
Fitness estimation of variants.....	124
Visualization of PDB files and determination of $\Delta\Delta G$ value	124
Calculation of average Shannon entropy	124
ADAR RT-qPCR	125
Rescue of EBOV mutants	126
Phenotypic characterization of mutants	127
Results.....	129
Experimental evolution through serial passaging of EBOV in human and bat cells	129
131	
EBOV takes distinct evolutionary paths in human-derived and bat-derived cell lines	131

EBOV populations passaged in bat-derived cells exhibit a ‘spike’ of high frequency mutations consistent with ADAR activity in the glycan cap/mucin-like domain region of the GP protein.....	134
Human and bat cell passage produced viruses have distinct population structures	138
Human and bat cell-derived mutants displayed cell-specific fitness patterns.....	140
Discussion.....	146
CONCLUSIONS AND FUTURE DIRECTIONS	156
Conclusions.....	156
The unique biology of bats	156
Study 1: Viral small RNAs	156
Study 2: Experimental evolution	157
Thoughts on methodology	157
Future directions	158
Study 1: Viral small RNAs	158
Identification of the biogenic origin of EBOV vncRNAs	158
Characterization of the MARV vncRNAs	159
Study 2: Experimental Evolution.....	159
Experimental evolution in other bat cell lines	159
Further validation of ADAR editing activity	160
Validation of escape mutant generation via ADAR hypermutation	161
Appendix A – Detailed sequencing data for vncRNA chapter	162
Quality data.....	162
Aligned reads data	164
Appendix B – Significance results for vncRNA chapter figures	166
Appendix C – Additional reagents for vncRNA chapter	173
Tag-based strand-specific EBOV qPCR.....	173
miRNA-specific qPCR	173
Antibodies	174

siRNAs.....	174
pmIRGLO Inserts	175
Appendix D – Novel filovirus 5’ mRNA structures	176
Predicted sequences of Xīlǎng virus (XILV) (Genus <i>Striavirus</i>) transcription start sites	176
Example predicted XILV and BOMV 5’ mRNA secondary structures	177
XILV VP40 analogue	177
BOMV VP40	177
References.....	178
Vita 234	
Copyright/Intellectual Property disclaimer.....	236

List of Tables

Table 1:	GS sequence-derived ncRNA sequences and their features.....	79
Table 2:	Comparison of EBOV vncRNAs to genuine miRNAs	99
Table 3:	ADAR RT-qPCR primers	126
Table 4:	Competition assay PCR and sequencing (indicated by *) primers	129

List of Figures

Figure 1:	Viral titers and RNA abundance from human and bat cell lines following infection with EBOV and MARV	65
Figure 2:	Schematic of bioinformatics analyses of sequencing data.....	66
Figure 3:	Length distribution profiles of EBOV and MARV vsRNAs in human and bat cells	68
Figure 4:	vsRNAs decrease in abundance over time and are predominately derived from positive-strand RNA.....	69
Figure 5:	Filovirus vsRNAs are primarily derived from structured elements in the 5'-UTR of viral mRNAs	70
Figure 6:	Filovirus vncRNAs are a substantial portion of the total vsRNA population.....	72
Figure 7:	vsRNA read frequency plots for rEBOV-infected and rMARV-infected cells 12 and 24 hpi.....	74
Figure 8:	Filovirus vncRNAs exhibit cell-line dependent differential abundance.....	76
Figure 9:	Absolute quantitation of EBOV GP vncRNA by RT-qPCR.....	78
Figure 10:	Filovirus vncRNA abundance changes over time	81
Figure 11:	Length distribution profiles of EBOV and MARV vncRNAs	82

Figure 12: miRNA RT-qPCR detection of EBOV GP vncRNAs in NHP liver tissue.....	83
Figure 13: EBOV vncRNAs are produced in a Dicer-independent manner....	86
Figure 14: Validation of siRNA knockdown by western blot and RT-qPCR..	88
Figure 15: EBOV vncRNAs are not associated with RISC, and do not positively or negatively affect virus replication.....	89
Figure 16: Dual-luciferase reporter assay for EBOV vncRNA function.....	90
Figure 17: Schematic representation of experimental design	130
Figure 18: Titer of clones across passage series.....	131
Figure 19: Population-level characterization of passaged EBOV populations	132
Figure 20: Average variant frequency among replicates.....	133
Figure 21: Average Shannon entropy of passaged virus	134
Figure 22: Bat cell-passaged virus accumulates high-frequency A-to-G mutations	135
Figure 23: High-frequency A-to-G mutations in EpoNi/22.1 are likely the result of ADAR activity	137
Figure 24: Frequency trajectory of selected variants identified within variant clusters identified in human and bat cells	139

Figure 25: Trajectories of select high-fitness mutations identified in human cell passaged virus	140
Figure 26: Clusters of variants detected in specific protein domains	141
Figure 27: Predicted impact of selected mutations on protein stability	142
Figure 28: Phenotypic effects of selected mutants	143
Figure 29: Competition and complementation assays	145
Figure 30: Nucleotide composition of the EBOV genome	150

List of Illustrations

Illustration 1: Electron micrograph of a typical EBOV virion (~70,000X magnification).....	18
Illustration 2: EBOV genome organization.....	19
Illustration 3: Electron micrographs of typical EBOV inclusion bodies (human cells).....	25
Illustration 4: Mapped index cases of major EBOV outbreaks	27
Illustration 5: RNA interference mechanisms.....	50
Illustration 6: Hypothetical multispecies ecology of EBOV in bats	160

List of Abbreviations

ADAR	Adenosine deaminase acting on double stranded RNA
BDBV	Bundibugyo virus
BOMV	Bombali virus
cRNA	Complementary RNA
DIC	Disseminated intravascular coagulation
DRC	Democratic Republic of the Congo
dsRNA	Double stranded RNA
EBOV	Ebola virus
EVD	Ebola virus disease
GC	Glycan cap
GE	Gene end
GS	Gene start
GP	Ebolavirus spike glycoprotein
IAV	Influenza A virus
IFN	Interferon
MARV	Marburg virus
MBE	miRNA binding element
miRNA	microRNA
MLD	Mucin-like domain
mRNA	Messenger RNA
ncRNA	Noncoding RNA
ORF	Open reading frame
PS	Phosphatidylserine
RdRp	RNA-dependent RNA polymerase
rEBOV	Recombinant Ebola virus
RESTV	Reston virus
RISC	RNA-induced silencing complex
rMARV	Recombinant Marburg virus
RNAi	RNA interference
rRNA	Ribosomal RNA
sGP	Secreted glycoprotein
siRNA	Small interfering RNA
sRNA	Small RNA
ssGP	Small secreted glycoprotein
ssRNA	Single stranded RNA
SUDV	Sudan virus
TAFV	Tai Forest virus
vncRNA	Viral noncoding RNA
vRNA	Viral RNA
VSR	Viral suppressor of RNAi
XILV	Xīlǎng virus

INTRODUCTION

Ebola virus: Virology, epidemiology, ecology, and evolution

Ebola virus (EBOV) is a member of the genus *Ebolavirus*, which is shared with its close relatives, Sudan virus (SUDV), Reston virus (RESTV), Bundibugyo virus (BDBV), Tai Forest virus (TAFV), and Bombali virus (BOMV) (1-3). All of these other than BOMV are known to cause a severe febrile illness with or without hemorrhagic manifestations in humans (excepting RESTV) and nonhuman primates (4). The threat posed by these viruses to Africa and to international public health has been demonstrated as a result of the large West African epidemic (2014-2016) (5, 6), and a number of smaller outbreaks that have occurred subsequently. This, in addition to serious and sustained concerns regarding biological warfare and bioterrorism, makes Ebola virus a unique and ongoing threat to international health and security (4, 7). The development and deployment of an effective vaccine has reduced this threat, but as the failure (due in large part to social and political issues) of a concerted vaccination effort to adequately contain the EBOV outbreak in eastern DRC that began in August 2018 has demonstrated, the risk is not by any means controlled (8).

BASIC VIROLOGY

Virion and genome structure

Ebolaviruses, along with the related marburgviruses, cuevaviruses, and dianloviruses are members of the family Filoviridae, a group of non-segmented negative strand viruses in the order *Mononegavirales* (1, 4, 9). Like all filoviruses, morphology is variable but filamentous (**III. 1**). Many of the more exotic morphologies observed when virion preparations are viewed via electron microscopy (EM) are likely the result of

centrifugation and other manipulations required to prepare the virions for EM (4, 10-13). Virions are approximately 80 nm in diameter, but virion length is inconsistent, as approximately half of all virions are polyploid to some degree (10, 14, 15).

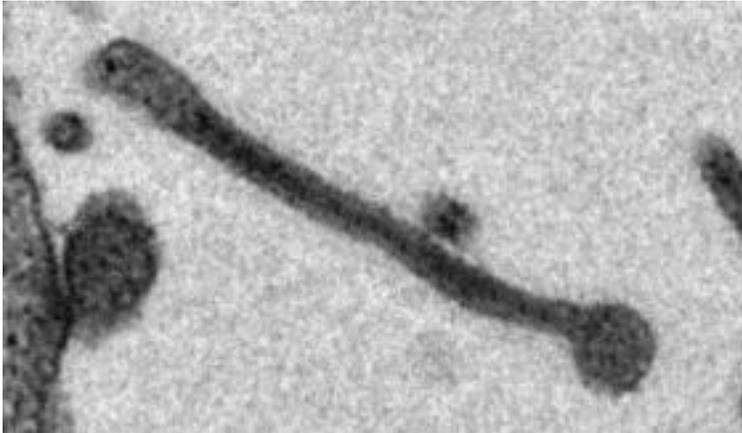


Illustration 1: Electron micrograph of a typical EBOV virion (~70,000X magnification)

Monoploid virions are approximately 970 nm (11-13), but polyploid virions of up to 14 microns in length have been documented in tissue culture (7). Monoploid particles and virions of low-order polyploidy dominate *in vivo*, however (10, 14). Under EM, the spike glycoprotein is observed to project 7-10 nm from the envelope, and individual spike glycoproteins are spaced approximately 10 nm apart (13). Underlying the lipid envelope is a protein matrix that serves to link it to the helical ribonucleocapsid (10, 13). The 19 kb EBOV genome contains seven genes in the following order: 3' genomic leader, NP (viral nucleoprotein), VP35 (polymerase cofactor), VP40 (matrix protein), GP (spike glycoprotein), VP30 (transcription initiation factor), VP24 (minor matrix protein), L (polymerase) 5' genomic trailer (4, 16) (III. 2).

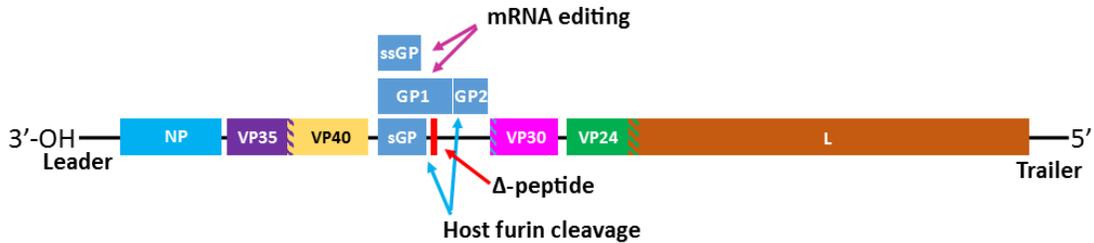


Illustration 2: EBOV genome organization

The GP gene produces four products through a combination of mRNA editing and posttranslational cleavage (17-20). Of these, only the full length glycoprotein and secreted glycoprotein (sGP) are definitively known to have a role in pathogenesis (4). The others, the “super-small secreted glycoprotein” and delta-peptide have not been conclusively associated with EBOV pathogenesis (4), though delta-peptide has been described as a viroporin (21), and may inhibit viral entry to some extent (22).

Gene function

NP, the viral nucleoprotein, forms a 70 nm helical ribonucleocapsid in complex with viral RNA, with each NP monomer binding seven nucleotides of RNA (23-25). In the absence of viral RNA or other viral proteins, NP will self-assemble into pseudocapsids (26). The other components of the replication complex (VP35, VP30, and L) are found in complex with the ribonucleocapsid in a mature virion (26-28). VP35, a phosphoprotein homologous to the “P” protein of other mononegaviruses, functions as a polymerase cofactor (29-32). Additionally, it is an interferon antagonist that acts by inhibiting IRF-3 signaling, disguising viral dsRNA from RIG-I via dsRNA binding, and sequestering PACT (33-37). This latter function enables it to act as a viral suppressor of RNA interference (VSR) (38, 39). A basic region in the C-terminal portion of the protein is responsible for these functions. VP30 is an RNA (40) and zinc-binding (41) phosphoprotein that mediates the transition between replication and transcription (42-45). There is evidence that VP30 also possesses VSR activity, which appears to be linked to inhibition of Dicer (39). L, as

the major component of the replication complex, exhibits a significant degree of homology with the polymerases of other negative strand viruses. It serves as the RNA-dependent RNA polymerase, and caps nascent viral mRNAs (16, 46, 47). Relatively little is known about the structure and function of L, as it was not possible to express and purify the intact protein until quite recently (48). VP40 serves as the major matrix protein, and is heavily involved in the budding process (15, 26, 31, 49). This matrix, found underlying the lipid envelope, interacts with the ribonucleocapsid to form a remarkably flexible filamentous virion (10, 50). VP40 is capable of budding to form virus-like particles in the absence of other viral proteins (49). VP24, a minor matrix protein without any direct homolog outside of the filoviruses, is found in small quantities in the matrix (4, 15). VP24 is not required for formation of virus-like particles in transfection-based systems, but is required for capsid formation and budding in actual infection, suggesting that its role in budding is likely related to functions of the viral lifecycle that are not replicated in these simplified systems (49, 51, 52). VP24 is known to be an interferon antagonist, preventing STAT-1 translocation to the nucleus by binding karyopherin- α (53, 54). As such, it acts largely downstream of the IFN blockade of VP35. Some evidence suggests that VP24 may have a role in the regulation of viral transcription, but other work has suggested that this could, to some extent, be an artifact of the systems in which the initial experiments were conducted (55, 56). The GP gene by default produces the secreted glycoprotein (sGP), a separate protein largely identical to the N-terminal end of GP (57, 58). Production of sGP occurs due to the fact that the second half of the sGP mRNA is one base out of frame for translation of full-length GP (57, 58). Secondary structure ahead of a stretch of seven uracil residues (the editing site) induces a polymerase stutter that includes an additional uracil at the editing site, which puts the remainder of the mRNA in frame to produce GP (59). This process ensures that infected cells produce large amounts of sGP, which is known to have immunomodulatory effects, and appears to serve as an antibody decoy (60, 61). Addition of another uracil during transcription leads to the production of the small secreted

glycoprotein (ssGP) (20, 59). The peptide product of translation of full-length GP mRNAs must be cleaved by host furin to produce two products, GP1 and GP2 (62). Host furin cleavage of the pre-sGP product produced by non-edited transcription of GP mRNAs produces sGP and delta-peptide (18, 19). GP1 includes the surface receptor binding portions of GP, and the receptor binding domain required to bind NPC1, the internal receptor EBOV uses to escape the endosome (63-65). Outside of the late endosome, this receptor binding domain is normally covered by the heavily glycosylated mucin-like domain (MLD) and glycan cap (65). GP2 includes the fusion loop and transmembrane domains (63, 65). GP1 monomers bind GP2 monomers to form a heterodimer that in turn binds with two additional GP1/GP2 heterodimers to form the trimeric surface glycoprotein (65). The trimeric spike glycoprotein is found anchored in the envelope, and interacts with VP40 (10, 15, 65). Transmembrane-anchored GP can be liberated from the cell membrane by TACE metalloprotease cleavage at a site within GP2 (66). This shed GP is known to have immunomodulatory effects, and contributes to endothelial dysfunction (67).

Attachment, entry, and uncoating

The manner in which EBOV attaches to target cells is not definitively known, and it is likely that the virus is not dependent upon a single mechanism. Reported attachment mechanisms include GP binding to C-type lectins such as DC-SIGN, and TIM-1/HAVCR and TAM binding to phosphatidylserine (PS) in the virion envelope (apoptotic mimicry) (4, 68-70). By contrast, the process of internalization and release into the cytoplasm is far better understood. After attachment, the virion is internalized via macropinocytosis, becoming incorporated into an early endosome (71, 72). The virion is then trafficked to the late endosomal state in a Rab5/Rab7 GTPase dependent manner (72). Acidification of the late endosome activates a number of low-pH dependent proteases that remove the MLD and glycan cap, exposing the receptor binding domain, permitting NPC1 binding (73). Although it has been suggested that cathepsins B and L act as the primary mediators of this

proteolytic processing (73-75), some evidence exists to suggest that although these may be involved, they are not required (76). Although fusion requires NPC1 binding, it is unknown if NPC1 binding event alone triggers membrane fusion, and there is evidence that additional processing may be required (63, 77-79). Fusion itself appears to follow the pattern established for similar viral fusion proteins, with exposure of the fusion loop leading to insertion of a hydrophobic structure into the target membrane, followed by formation of a fusion pore and release of the ribonucleocapsid into the cytosol (80).

Replication and life cycle

After release from the endosome, replication broadly follows the same pattern as the other mononegaviruses, with replication occurring exclusively in the cytoplasm (32, 81). Although replication begins immediately, significant accumulation of viral proteins begins to occur between 12 and 18 hours post-infection (81). Filovirus genome replication begins when the replication complex (composed of L homo-oligomers in complex with VP35 homo-oligomers and oligomerized VP30) anchors to the 3' leader sequence (44, 46, 82-85). A second promoter located within the 5' UTR of the NP gene appears to be required to initiate transcription, along with the 3' genomic leader (86). Interaction of the replication complex with the rest of the ribonucleocapsid is mediated by VP35, which serves as a bridge between L polymerase and the NP monomers that coat the RNA (29, 30, 46). Beginning at the leader, antigenomic complementary RNAs (cRNAs) are transcribed. Nascent cRNAs are immediately encapsidated by NP monomers (87, 88), facilitating the next step in replication; here the cRNAs are used as templates to produce viral genomic RNAs (vRNAs), starting this time at the genomic trailer sequence (5' in genomic sense, 3' in antigenomic sense) (32). Uniquely amongst the nonsegmented negative strand viruses, EBOV L polymerase adds a non-templated 3' nucleotide during genome replication, potentially as a strategy to avoid certain aspects of the cellular innate immune response (89).

EBOV transcription is dependent upon VP30, which acts as a transcription initiation factor (42). The precise nature of the role of VP30 in regulating the transition between transcription and replication is largely unique to filoviruses (44). This switch is dependent upon the phosphorylation state of a number of amino acid residues within VP30. These are located primarily but not exclusively within two clusters near the N-terminus (42, 44, 45). When dephosphorylated, VP30 is incorporated into the replication complex and permits transcription (90). In the absence of dephosphorylated VP30, the replication complex is restricted to genome replication, and cannot transcribe mRNAs (42). This regulation appears to be mediated by a stem-loop structure formed by the NP gene start signal; when this is ablated, transcription occurs in a VP30-independent manner (91). The kinase(s) responsible for phosphorylation of VP30 have not been identified, but dephosphorylation of VP30 is dependent upon the host phosphatases protein phosphatase 1 (PP1), and protein phosphatase 2 (PP2) (44, 45). The responsible host enzymes are recruited to EBOV inclusion bodies by NP (92), supporting NP-mediated dynamic phosphorylation of VP30 within the inclusion bodies (93, 94). Transcription begins much like replication. The polymerase complex, this time with dephosphorylated VP30, anchors to the genomic leader (27, 44, 45, 91). However, instead of beginning replication at the leader, the polymerase complex then shifts to the beginning to the NP gene, which, like all filovirus genes is demarcated with a short and highly conserved 12 nt gene start sequence (16). Along with additional downstream sequence, this signal also forms a stem-loop structure in the mRNA product of transcription (16). EBOV has two such signals, one, which is used by the NP and L genes, and another which is used by all other genes. These differ by only a single base (16). L polymerase recognizes this signal, and initiates transcription (32, 91, 95, 96). Nascent mRNAs are capped by L polymerase shortly after production begins. Although the nature of EBOV capping is not well understood, in the prototypical mononegavirus, vesicular stomatitis virus, this tends to occur relatively quickly, within the first 40 nucleotides (97). Transcription is typically terminated when the

polymerase complex encounters the 11-12 nt gene end signal, which much like the gene start signal, is highly conserved, and contains a poly(U) stretch that initiates tailing of mRNAs (95). Termination is not perfectly efficient however, and bicistronic mRNAs are produced when termination fails. Bicistronic EBOV mRNAs exhibit significantly reduced efficiency of translation for the second open reading frame (95, 96). After termination, the polymerase may become disassociated with the genome, in which case it must start from the beginning of the genome, or it may proceed to the next gene start site and re-initiate transcription of the next gene (32, 95). As the probability that the polymerase complex will disassociate from the genome is higher than that of it continuing transcription, this means that each gene has a successively lower probability of being reached by any given “run” of a polymerase complex. This results in a transcriptional gradient in which the genes at the 3’ end of the genome are transcribed in far higher abundance than those at the 5’ end of the genome (32, 98, 99). This is one of the major mechanisms for regulation of transcription of individual genes (95). Like all filoviruses, EBOV genomes contain a mixture of overlapping and clearly demarcated genes. While the NP/VP35, VP40/GP, and VP30/VP24 gene junctions contain defined 5 nt intergenic regions, the VP35/VP40 and GP/VP30 gene junctions contain a 5 nt overlap of the gene stop signal of the previous gene with the gene start signal of the following gene (95, 96). In EBOV, but not the other members of the genus Ebolavirus, the VP24/L gene junction is unique, as it contains two gene stop signals for VP24, one of which overlaps with the gene start signal for the L gene, while the other is separated from the L gene start signal by a 5 nt intergenic region (16, 95, 96). In contrast to the short NP/VP35 and VP40/GP intergenic regions, the VP30/VP24 intergenic region is remarkably long, at 144 nt. This is likely a means of reducing transcription of VP24 and L (which must be present in only relatively small quantities for efficient viral replication), as particularly long intergenic regions have been shown to reduce the transcription of the downstream genes (95, 96). A number of mechanisms reduce

the production of L polymerase, such as the inclusion of a “false” start codon in the 5’ UTR of the gene (100).

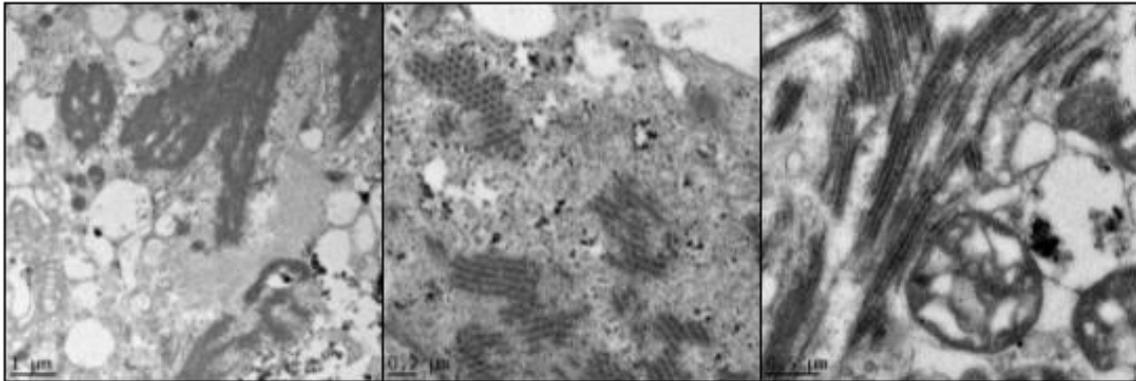


Illustration 3: Electron micrographs of typical EBOV inclusion bodies (human cells)

During replication, large inclusion bodies are formed within the cytoplasm, and are clearly visible when examining infected cells with transmitting electron microscopy (11, 13, 101). Within these inclusion bodies, helical filamentous nucleocapsid structures are clearly visible (13, 101)(**III. 3**). These inclusion bodies are organized by NP, and are the primary site of viral replication and protein production (81, 93, 102). After proteins are produced however, their ultimate destinations within the cell vary. NP, VP35, VP30, and L are found almost exclusively within inclusion bodies, whereas GP is translated at the rough endoplasmic reticulum (ER) and trafficked through the ER/Golgi pathway to the plasma membrane, where it becomes incorporated into lipid rafts in preparation for budding (81, 103). VP40 is initially found within discrete cytoplasmic inclusions at 18 hours post-infection, but becomes far more diffuse by 24 hours, eventually localizing to the inner leaflet of the plasma membrane, where it associates with the transmembrane domain of GP in preparation for viral budding (81). VP24 follows a pattern that is broadly similar to that of VP40 (81). Assembled nucleocapsids are translocated to the plasma membrane via actin-dependent transport, where VP40 and VP24 mediate budding through GP-rich lipid rafts via host ESCRT (4, 26, 104). During budding, phosphatidylserine (PS)

is incorporated into the envelope, a process that requires the host protein TMEM-16F (105). Multiple genomes can be incorporated into a single envelope, as previously discussed, with either discrete genomic RNAs, or, in some cases, linked genomic RNAs (10). After budding, virions require no additional external or internal enzymatic processing to become infectious (26).

HISTORY AND EPIDEMIOLOGY

A brief history of Ebola virus, 1976-2020

The first documented Ebola virus outbreak began in August of 1976, when a schoolteacher presented to the Yambuku mission hospital in northwestern Zaire (now Democratic Republic of the Congo, DRC) with symptoms that looked very much like malaria to the Belgian nuns who served as nursing staff (106, 107). Within days, however, it became apparent that malaria was not responsible for the schoolteacher's illness. The disease rapidly spread through the community, aided by the reuse of needles and other medical equipment at the hospital (108). Case fatality was extremely high, and the clinical presentation of the disease was dramatic, often including marked hemorrhagic symptoms (108-110). Eventually, a number of nuns were infected, of whom two were transported to the capital, Kinshasa, for treatment. There, a young nurse, Mayinga N'seka was infected when caring for one of these nuns. Nurse Mayinga eventually died on October 19th, and shortly afterward, samples of her blood were taken and sent for laboratory analysis at multiple centers in Europe and the United States, along with samples from the two nuns, for identification of the infectious agent (4, 108).

Initially, the agent responsible was thought to be Marburg virus, a filovirus related to Ebola virus that had caused a small outbreak of severe hemorrhagic fever in workers handling non-human primate tissues in Germany and Yugoslavia in 1967 (111), and which had re-emerged from Zimbabwe (then Rhodesia) in 1975 (4). Further efforts found that although the agent responsible was identical to Marburg virus by electron microscopy, it

was antigenically distinct. However, it was related to the virus that was responsible for a nearly simultaneous outbreak in southern Sudan (this virus is now known as Sudan virus, another member of the genus Ebolavirus) (112, 113). Subsequently, a number of natural outbreaks of varying size have occurred, including the recent very large epidemics in West Africa and eastern DRC. Most outbreaks have occurred within the Congo Basin, with outbreaks in 1995, 2003, 2007, 2014, 2017, and two in 2018 (4, 114-119) (III. 4). The overwhelming majority of outbreaks have been related to a single spillover event, though the 1994-1996 outbreaks in Gabon have been linked to multiple exposure events (117-119). A small number of laboratory exposures have also been reported, some resulting in the death of the exposed individual (4).

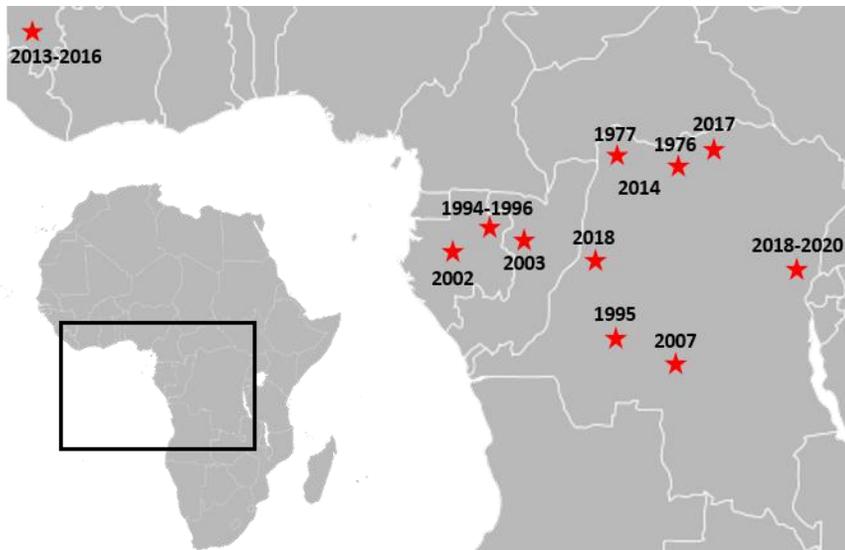


Illustration 4: Mapped index cases of major EBOV outbreaks

The two largest EBOV outbreaks have both occurred within the past decade, the first, the 2013-2016 West African epidemic had more than 28,000 confirmed cases with over 11,000 deaths (6). The second largest, the ongoing outbreak in the far Eastern region of the DRC, began in August, 2018, with nearly 4,000 cases and 3,000 deaths reported as of this writing, according to publicly available DRC Ministry of Health statistics. Due to

the unique status of the West African epidemic, a brief history will be provided here to illustrate the unique threat posed by EBOV to large urban centers in developing nations.

Prior to the 2013-2016 epidemic, and excluding a single imported case in South Africa associated with the 1994-1997 epidemic in Gabon (120), EBOV had not been associated with human outbreaks outside of the Congo Basin, although phylogenetic analysis suggests that it had been present in West Africa since the early 2000s (121). The primary case of the West African epidemic was identified as a one year old male in the village of Méliandou, Guéckédou Prefecture, Guinea. Anecdotal evidence indicates that he may have had some form of contact with bats¹, though the tree near which the contact is said to have occurred has since been damaged by fire (122). The primary case died in December 2013, suggesting that the spillover event occurred in December or late November of that year (122). Initially, the outbreak progressed slowly, and did not come to the attention of health authorities or the international community until March, 2014. As with many filovirus outbreaks, the first cases were initially dismissed as being yellow fever, malaria, or any of the other febrile diseases endemic to Africa. This was doubly the case in West Africa, where there was little reason to suspect EBOV, and far more reason to suspect Lassa virus, which is well-established in the region (123). However, after diagnostic confirmation, Guinea's Ministry of Health reported the first Ebola virus disease cases to the World Health Organization in late March. Simultaneously, reports of imported cases in Sierra Leone and Liberia surfaced. By April, the virus was well-entrenched in Sierra Leone and Liberia, establishing chains of infection in major cities in the latter, including Monrovia, the capital. However, growth was still relatively slow, though the number of cases did surpass the 1995 Kikwit outbreak, previously the largest, in late July (6). Also in late July, the first cases were reported in Freetown, the capital of Sierra Leone. August and September saw the rapid exponential growth of the outbreak, as contact tracing failed,

¹ *Mops condylurus* (Angolan freetail bats).

alongside efforts to disrupt chains of transmission in urban areas. Concomitantly, cases were exported to neighboring African countries, Europe, and the United States (5). Although local chains of transmission were generally not established, two healthcare workers were infected in the United States while caring for the imported case. Control efforts were hampered in West Africa by persistent non-compliance issues and violence directed against healthcare workers, facilities, and epidemiological staff (124). The outbreak began to wane in late 2014 and early 2015, as control efforts began to succeed (5). Sporadic cases continued to appear into December. By January 2016, transmission of the virus had largely ceased (6). “Flare up” cases, some associated with sexual transmission from survivors, continued to appear until April, when the last case in Liberia recovered (6). After the requisite 42 days (two incubation periods), the outbreak was declared over in June, 2016 (6). Near the end of the epidemic, the first field trials of the VSV-EBOV vaccine were conducted, which would eventually lead to the approval of the vaccine (125).

Epidemiology

When actually identified, the primary case in natural EBOV outbreaks has typically had some contact with wildlife, usually nonhuman primates or bat. In most cases, however, including the relatively large 1995 Kikwit outbreak where the primary case was a forest worker, no single source could be identified (115, 126). Typically, exposure occurs as a result of hunting, butchering, eating, or otherwise handling said wildlife. Given that the infectious dose is calculated as <1 plaque forming unit (PFU), in the case of an infected non-human primate this may result in an exposure to excess of tens of thousands of 0% infectious doses (ID₅₀s) (7). In at least one outbreak, contact with fruit bats was reported (116).

EBOV transmission is largely directly from person to person as a blood-borne pathogen, but infection via fomites does occur (127, 128). Mucous membrane contact with infected fluids, which includes large droplets as might be produced when a patient coughs violently or vomits, presents the most significant means of transmission in many outbreaks

(128-130). However, reuse of needles and sharps injuries represent another significant means of transmission, particularly in the context of medical care in Africa (4, 128). The spread of the 1976 Yambuku outbreak was greatly aided by the extensive reuse of needles by medical staff at the mission hospital, and given that parenteral exposure is known to lead to far more severe disease than low dose mucosal exposure, it is probable that this contributed to the extremely high (88%) case fatality associated with the outbreak (130). True aerosol transmission of EBOV is a controversial topic (131), but the general consensus in the literature is that it does not occur readily under natural conditions (4). The epidemiology of the West African epidemic appears to support this (128, 132). Another question raised prior to the West African epidemic that appears to have been at least partially resolved is that of asymptomatic infection. Prior to 2014, a number of serosurveys found EBOV seroprevalence upwards of 10% in certain African populations, with some suggesting that this may indicate that EBOV has a pattern of asymptomatic infection resembling that of Lassa virus (133). However, while asymptomatic infection and minimally symptomatic infection clearly occurred during the West African epidemic, the evidence does not appear to support the very high rates of asymptomatic infection that would be suggested by a 10% rate of seroprevalence in an area that had not recently experienced an EBOV outbreak (134, 135). Notably, the reliability and consistency of seroprevalence studies in the context of EBOV has been questioned (136).

EBOV transmission during an active outbreak generally stems from one of four sources: 1) Social contact with a mildly ill individual in the early stages of the disease, 2) Providing direct care to an acutely ill individual (including professional healthcare workers), 3) Participation in traditional healing rituals/use of traditional healers, 4) Handling and burial of infected corpses, to include participation in traditional funeral rites (4, 137). Sexual transmission is a far more infrequent mechanism, but became particularly prominent near the end of the 2013-2016 West African epidemic (128, 138). As a general rule, casual social contact is a relatively infrequent means of transmission, as cases are only

contagious when they are actively symptomatic (4, 7). However, individuals just beginning to show signs of disease may still engage in types of social contact that carry a risk of transmission. Unquestionably, however, individuals providing direct care to infected individuals are at the greatest risk of infection, as they are inevitably in direct contact with highly infectious fluids and other materials (128, 137). Tradition in many parts of sub-Saharan Africa dictates that family members provide direct care to ill family, and in the early stages of an outbreak, this often leads to clusters of cases within a single immediate family (4, 137). For this reason, one of the most effective interventions during an EBOV outbreak is the establishment of specialized treatment units where trained staff can provide care using appropriate personal protective equipment, thereby reducing the incidence of caregiver infection (139). In many of the parts of Africa in which EBOV is endemic, funeral rites include significant direct contact with the body of the deceased, generally in the form of washing and wrapping the corpse for burial, but also other forms of direct contact during the funeral itself, usually involving most of the attendees (137). This practice is obviously extremely high-risk in the context of an EBOV outbreak, and traditional funerals are often found to be the common origin of large clusters of cases within an extended family (137). Furthermore, many of the traditional healing practices in these areas involve piercing or cutting of the skin, generally without subsequent disinfection of the implements used (137). The role of these cultural factors in the spread of the virus during an outbreak is well-understood, and control measures invariably include either education regarding the risks of these practices, warnings against continuing them for the duration of an outbreak, or even outright prohibition of certain particularly high-risk customs (137, 139). Fear and cultural factors, particularly those related to mistrust of western governments and medicine have been substantial complicating factors in epidemic response since the 1976 outbreak, and represented a major roadblock to control in West Africa (5, 124, 137).

Experience has demonstrated that the most effective interventions during an EBOV outbreak are rooted in basic epidemiology (137). Cases must be identified, isolated in treatment units, and contacts must be followed up and potentially isolated themselves until the end of a 21 day period representative of the longest reasonable incubation period for the virus (7). As discussed previously, education as to the risks of certain cultural factors and properly managed safe burial are also significant interventions that have proven crucial to curtailing outbreaks (5, 137, 139). Given the generally remote nature of most outbreaks, this led to the outbreak eventually “burning out” once chains of transmission are broken (137). Outbreaks are considered resolved when 42 days (two incubation periods) have elapsed following the recovery of the final case (4). Although these strategies were effective in containing the (relatively) small outbreaks that occurred from 1976-2014, the rapid expansion of the West African epidemic during the late 2014 phase of the epidemic overwhelmed this approach, due in part to under-resourcing and underestimation of the potential of the virus to spread within the urban environment it was introduced to in late 2014 (4, 6, 140). This resulted in an outbreak of unprecedented size, which was eventually only controlled with the application of significant economic and personnel resources (5, 6). With the introduction of the recently licensed² VSV-EBOV vaccine, contact tracing now includes vaccination of contacts and contacts of contacts, essentially utilizing a ring vaccination strategy to contain the spread of the virus (141-143). Individuals who are engaged in particularly high-risk activities, such as medical and burial workers, and those who are engaged in epidemiological work should also be vaccinated (4). The knowledge that contacts and contacts of contacts will be offered the vaccine can serve as an incentive to these individuals, thereby improving the success rate of basic epidemiological measures (141, 142, 144). This vaccine-assisted control strategy was remarkably effective during the 2018 western DRC outbreak, bringing an outbreak which, due to its proximity to a major

² Licensed as “Ervebo” by FDA and European regulators in late 2019.

urban center presented a profound risk to regional and international public health, into control rapidly (145). However, even the vaccine-assisted control strategy requires effective basic epidemiology work, and as the outbreak that began in eastern DRC in August 2018 has demonstrated, mass vaccination campaigns are only minimally effective in the absence of effective and thorough case follow-up and contact tracing (8, 146). Additionally, evidence of vaccine misdirection exists (147). Finally, and perhaps most significantly, the requirement for a -70° C cold chain presents a significant and ongoing challenge to vaccine deployment in remote locations in Africa (8, 146).

MEDICAL ASPECTS OF EBOLA VIRUS INFECTION

Course of infection and pathology

In humans and nonhuman primates, initial infection following exposure is localized to the tissue exposed to the virus, with the virus entering systemic circulation via transport by dendritic cells and macrophages from the periphery to lymph nodes, where infection of differentiated myeloid lineage cells provides the primary round of amplification required to initiate parenchymal infection of solid organs (148). In addition to this early role in infection, dendritic cells and macrophages are sustained targets of viral infection throughout the course of the disease (148, 149). Curiously, undifferentiated monocytes display limited susceptibility (relative to other cell types) to infection *in vitro*, as do dendritic cells (150). Dendritic cells also exhibit an infection phenotype markedly different from that observed in macrophages (149). Infected dendritic cells fail to activate (151, 152), apparently as a result of the interferon antagonist activity of VP35 and VP24 (153). By contrast, infected macrophages produce large quantities of proinflammatory signaling mediators, contributing to cytokine storm, disseminated intravascular coagulation (DIC), and endothelial dysfunction (148, 149, 154). A particularly remarkable aspect of EBOV infection is the phenomenon of bystander lymphocyte apoptosis (155, 156). T lymphocytes are only abortively infected with EBOV (157), yet cell death still occurs. This appears to

be the result of a number of factors, including said abortive infection (157). GP has been found to induce apoptosis of T lymphocytes via binding of TLR4 on the cell surface, which acts through a number of signaling pathways to induce both apoptotic and necrotic cell death (158). This TLR4 binding behavior of GP appears to induce a sepsis-like state that explains a great deal of the observed pathophysiology of the disease (158, 159). Importantly, sGP does not appear to induce lymphocyte apoptosis or changes in endothelial adhesion (160). In combination with the interferon inhibitory effects of VP35 and VP24 on dendritic cell function and cellular innate immunity, and the direct killing of myeloid lineage immune cells, this GP-mediated lymphoid depletion induces a state of “immune paralysis” during infection that makes viral clearance difficult (153, 159). This depletion affects all lymphocyte classes (161). Translating this to the clinical world, survival has been found to be correlated with both an effective cell-mediated immune response and control of this lymphoid depletion (7, 159, 162).

After an initial round of amplification in the lymphatic system and dissemination, infection of solid organs occurs (148). Although most tissue and cell type can be infected, primary targets include the parenchymal cells of the liver, the spleen, and the kidneys (4). Epithelial cells in particular are favored sites of replication, partially explaining the gastrointestinal symptomatology observed clinically (163). EBOV infection is associated with a prolonged viremia that persists until death or viral clearance. Viral load frequently reaches 10^8 pfu/mL of serum (148). In humans and nonhuman primates, immune privileged sites are infected, including the eyes and testes (128, 138, 148, 164, 165). Virus is nearly always found in the brain in necropsy of infected nonhuman primates in laboratory challenge studies (148). Despite the substantial involvement of the endothelium in pathophysiology, endothelial cells themselves may not be heavily infected until the later stages of the disease, though antigen staining has been observed. Endothelial dysregulation is instead mediated largely by cytokine storm and shed GP (67, 166, 167).

Gross pathologic and histopathologic findings are effectively identical in humans and nonhuman primates (4). Leukocytosis is a common finding, with neutrophil numbers increasing, and monocyte and lymphocyte numbers decrease, particularly later in the course in infection, when lymphopenia is marked (168, 169). Thrombocytopenia is also notable, again worsening as the infection progresses (148). Erythrocyte counts are also noted to decline, and again, this is most notable late in the course of infection (148). Serum chemistries exhibit little change early in infection, but ALT and AST (correlates of hepatic function), along with BUN and creatinine (correlates of renal function) rise sharply later in infection (148, 168, 169). These findings are typical of multiple organ dysfunction, and represent significant hepatic and renal distress resulting from viral infection and immunopathology. Coagulopathy is nearly universal, with increased clotting time and decreased levels of coagulation factors first becoming noticeable concomitantly with elevation in ALT and AST (148, 170). The etiology of coagulopathy in Ebola virus infection is multifactorial, and includes thrombocytopenia and coagulation factor loss subsequent to both DIC and decreased production in the liver, the result of extensive hepatic pathology (148, 170).

One of the most notable aspects of EBOV pathology is the induction of disseminated intravascular coagulation (DIC) (149, 171). DIC is defined as a pathologic state in which the balance between coagulation and fibrinolysis is dysregulated (172). This leads to abnormal coagulation, and the formation of microthrombi in the small vessels and capillaries. This state of deranged coagulation rapidly consumes available plasma coagulation factors and platelets, leading to extravasation of blood in the form of oozing hemorrhage (172). At the gross pathologic level, this is observed as congestion and disseminated minor to marked hemorrhage in multiple tissues. Microthrombi in vessels and extravasated erythrocytes are observed in tissue sections under microscopy (148). A number of external factors can trigger DIC, including endotoxin, cytokines (particularly IL-1), and TNF- α . These induce the release of tissue factor (TF), a transmembrane

glycoprotein found on the surface of a variety of cell types, including endothelial cells and myeloid lineage cells (172). Excessive release of TF initiates hyper-activation of the extrinsic coagulation cascade which in turn consumes coagulation inhibitors and ultimately results in the pathologic state of DIC (172). The mechanism by which EBOV induces DIC is believed to be twofold. Firstly, the cytokine storm triggered by infection of macrophages etc. makes a significant contribution, as excessive IL-1 and TNF- α alone can trigger DIC (149, 172). However, GP and soluble GP are thought to make significant contributions as well by inducing endothelial dysregulation/sepsis-like state by a number of mechanisms, including binding TLR4 (158, 159).

On necropsy/autopsy, notable findings typically include lymphadenopathy, hepatomegaly and friability and reticulation of hepatic tissues. Splenomegaly and other splenic pathology is often noted, continuing a general theme of lymphatic system involvement (148). Congestion (accumulation of blood/fluid as a result of impaired outflow) is commonly noted in the small intestine, with particularly marked congestion frequently noted at the pylorus/duodenum junction (148). Congestion is commonly noted in mucosal-associated lymphoid tissues throughout the lower GI tract and in a number of glandular tissues, including the adrenal glands. Hemorrhagic signs include petechiae and ecchymoses observed on multiple organs, including the skin, gastrointestinal tract, the organs of the urogenital tract, and lungs (148). Significant amounts of frank blood are not commonly found (7, 173).

Histopathology in EBOV infection is most remarkable in the liver, where widespread hepatocellular necrosis is a common finding, along with mononuclear cell infiltration of the portal ducts (148). Small droplet steatosis is frequently observed, and inflammatory effect is apparent. Additional characteristic histopathologic findings in EBOV infection include significant necrosis and lymphoid cell depletion in the spleen and lymph nodes, with signs of lymphoid apoptosis apparent with TUNEL staining of these tissues (148, 156, 174). Kidneys display tubular necrosis, with localized evidence of

apoptotic cell death (148). Significant viral antigen load is observed in all of these tissues. In the GI tract, mononuclear infiltration is apparent, as is necrotic cell death. The most remarkable finding is normally significant antigen load, correlating well with the large amounts of virus typically shed via the GI tract (148). Lung tissue typically shows signs of intra-alveolar congestion, edema, and hemorrhage, with significant antigen staining (148). By contrast, cardiac tissue exhibits significant observable antigen staining, but typically displays comparatively little pathologic lesion (148). No CNS histopathology is available for EBOV, but in Marburg virus infection, panencephalitis and leukocyte infiltration have been observed (4, 175).

Clinical presentation and treatment

Clinically, EBOV infection presents as Ebola virus disease (EVD), a complex viral syndrome typified by multiple organ dysfunction, endothelial dysregulation, and sepsis-like presentation later in the course of the disease (7). Despite the previous name “Ebola hemorrhagic fever”, actual hemorrhagic manifestations do not present in all cases, though they are common in instances of severe disease (7). Signs and symptoms appear after an incubation period of 2-21 days, though the majority of cases will show symptoms within 4-10 days (7). Incubation periods of longer than 21 days may occur in up to 5% of cases according to some modeling (176). Onset of symptoms is sudden, beginning with headache and various myalgias (4, 7, 173, 177). Progression to an influenza-like illness consisting of fever, myalgia, arthralgia, prostration, and pharyngitis is rapid (4, 7, 177). At this point, the differential diagnosis is extremely broad. With disease progression, the gastrointestinal symptoms that characterize the later stages of the disease appear, and include nausea, abdominal pain, emesis, and diarrhea (7, 163). Emesis and diarrhea increase in severity as the case progresses, and these fluids are highly infectious (7, 163). Without adequate fluid replacement, patients become dehydrated quickly (173). Maculopapular rash is common (upwards of 50% of cases), appearing within 5-7 days of onset of signs and symptoms

(177). Additional signs and symptoms include coughing, shortness of breath, edema, and in males, swelling of the urogenital tract (169, 173, 177). Neurological symptoms in severe cases include confusion, seizures, and, in fatal cases, coma. Curiously, patients with severe disease occasionally develop intractable hiccups (177). Hemorrhagic manifestations are far more common in EVD cases associated with EBOV than in those associated with the other ebolaviruses (7, 173). Nearly all patients display some degree of coagulopathy (173). Other mild manifestations include petechiae and ecchymoses on the skin (most apparent on light-skinned persons) and mucous membranes and formation of hematomas. Injection of blood into the sclera occurs occasionally (7, 173). A significant percentage of cases will present with bleeding from mucous membranes and oozing of blood from needle puncture sites. Gastrointestinal bleeding, including frank hemorrhage from the rectum, is somewhat common. As a result, patients may exhibit hematemesis and hematochezia (7, 173). Significant hemorrhage into the GI tract can occur, but is relatively uncommon. Although hemoptysis may be the result of hemorrhage into the alveolar spaces, it may also result from GI origin blood entering the upper respiratory tract (7, 173). Hematuria is also frequently observed. Serious hemorrhage is very rare, and exsanguination is not a typical cause of death (7, 173). Instead, death typically occurs as the result of the septic shock-like state induced by infection, with hypovolemia typically being the proximal cause of mortality (7, 173). In the later stages of disease, viremia can be profound (7). Perimortem patients are typically comatose (7, 177). The presence of significant hemorrhagic manifestations is a sign of severe and likely fatal disease (178). Patients who succumb to the disease normally do so within 6-10 days of onset of symptoms (7). Case fatality for EVD caused by EBOV varies, but 50-70% is a reasonable estimate of the range, with most outbreaks falling close to 60% (4). Patients who recover typically begin to do so within one to two weeks of onset of symptoms (178, 179). Recovery is slow, and marked by continued myalgia, arthralgia, generalized weakness and prostration, hepatitis, and anorexia (179). Patients may experience long-term to permanent hearing and vision loss to

varying degrees (180). At least once case of recrudescent EBOV meningoencephalitis has been described (181). Total viral clearance may not occur immediately in some immunologically privileged sites, including the testes and eyes (138, 165, 180). Persistence in the testes has been definitively linked to transmission of the virus to sexual partners, and this persistence has been observed out to several months following clearance of viremia (138).

The clinical presentation of EVD is difficult to distinguish from a number of other infectious diseases common in Africa, particularly in the early phase of the disease (182). Malaria in particular is a common misdiagnosis (183). Other differential diagnoses include dengue, yellow fever, a number of bacterial diseases, and in West Africa, Lassa fever (123, 182). The severity of symptoms can eliminate many of these possibilities as the disease progresses, but laboratory diagnostics are ultimately required for definitive diagnosis, as Marburg virus disease is clinically indistinguishable from EVD (4, 182). A high index of suspicion should be maintained for patients in endemic regions presenting with malaria-like symptoms in the absence of the classic malarial cyclic fever (183). Definitive diagnosis requires RT-PCR detection of EBOV RNA in the serum of an acutely ill patient (4). In the context of an outbreak, many diagnoses are clinical, due to the need to conserve resources and the difficulty of establishing laboratory infrastructure in some parts of Africa (184, 185). However, as highly atypical presentations are known, testing should be considered even for patients in an outbreak zone with apparent infectious disease that does not conform to the EVD clinical criteria (4, 184, 185).

No specific treatment existed for EVD until very recently, with supportive care being the only available modality of treatment (186). As a general rule, treatment remains largely supportive, particularly in Africa, and even the most effective treatment regimens still utilize specific treatments as adjuncts (169, 179). Imperative in the treatment of EVD is maintenance of fluid and electrolyte balance. Hypovolemic shock is the most common immediate cause of death, so this must be a management priority (169, 179). Although

intravenous fluid management is preferable oral fluid management is acceptable. Palliative measures are recommended, and include management of pain, nausea, and anxiety (169, 179). Certain NSAID analgesics, including aspirin and ibuprofen, should be avoided due to their antiplatelet effect (169, 179). With proper management, case fatality can be minimized. In an ICU setting in the developed world, with intensive care measures such as renal dialysis and total respiratory support, the vast majority of patients have survived (179).

A number of small-molecule and biologic therapies have emerged for EVD, particularly during and in the aftermath of the 2013-2016 West African epidemic (187). The practical field efficacy of these therapies has been quite variable, ranging from TKM-Ebola, which failed in field trials (188), to a number of monoclonal antibodies preparations and small molecule antiviral agents that have shown varying degrees of promise in both nonhuman primates and field trials (146, 189-192). Although specific immunotherapies have found success, the question of convalescent serum/plasma should be addressed. Although these have been previously suggested as therapeutic options, trials have largely found them to be largely ineffective (193-197). Despite this, reports of the use of convalescent serum, plasma, and even whole blood were common in the early stages of the West African epidemic (5, 6, 198). Although this activity is of dubious value, and there is some long-term risk associated with the practice, it could be argued that the risk of HIV infection is outweighed by the immediate concern of treating an EVD patient (5, 6).

The degree of patient isolation and staff personal protective equipment (PPE) required have been somewhat controversial topics in the field, though experience from the epidemic in West Africa has been extremely instructive (6, 169, 199). Patients should be placed in isolation, though cohorting with other EVD cases is acceptable (200). EVD patients produce copious amounts of highly infectious material, and although true small aerosol transmission does not appear to be a feature of natural outbreaks, large droplets are a transmission hazard (128). Where possible, negative pressure isolation rooms should be

used (200). Staff engaged in patient care must wear a full protective ensemble, generally consisting of a wraparound protective smock (often including an apron), head and neck covering, overshoes, double gloves, and eye protection (if a facemask is not a component of the respiratory protection used) (199). Respiratory protection is required, and an N95 respirator represents the minimum that should be used in most cases, with a PAPR being preferable (199). The process of donning and doffing PPE should be supervised and assisted by a second trained individual to minimize the risk of contamination during PPE removal, and to ensure that PPE is used properly (199). Previous guidelines issued by CDC recommended only basic contact precautions (surgical mask, gown, double gloves), but these have been found to be inadequate (169). Guidelines exist for home care of filovirus-infected individuals for situations in which family members refuse admission of a patient to a treatment unit, and these have been moderately successful in reducing transmission to family caregivers (8, 139, 140, 144). Treating an EVD patient outside of a dedicated treatment unit remains an extremely high-risk activity, and can be deleterious to public health efforts more broadly (128).

ECOLOGY AND EVOLUTION

Ebola virus ecology: An ongoing mystery

To date, the ecology of EBOV remains, for the most part, poorly described. Given the pathogenicity and lethality of the virus in primates, and the fact that long-term sustained transmission does not occur in these animals, it is clear that they do not serve as anything other than incidental hosts. Apes are known to be naturally infected, and somewhat regular epizootic events occur in the great ape populations of Central Africa (119, 201). Duikers, a group of small forest-dwelling antelopes, have frequently been discussed within the context of the ecology of the virus, but quality evidence for this assertion is limited to a single RT-PCR detection of EBOV RNA in one animal, with no prior, contemporary, or subsequent serological or other evidence (201). Starting early in the history of EBOV

research, and particularly after the 1995 Kikwit outbreak, significant efforts have been made to identify the reservoir host of the virus (202). Although these have been largely unsuccessful to date, there is a large body of evidence supporting the hypothesis that fruit bats, and in particular epomophorine bats, serve as a reservoir for the virus (116, 202-207). At least one EBOV outbreak was closely linked to exposure to fruit bats (116). A smaller, but still significant body of evidence suggests that some microbats, particularly those of the families *Molossidae* and *Miniopteridae* may also play a role in the ecology of EBOV and/or other members of the genus *Ebolavirus* (207). Although no ebolavirus has ever been isolated from bats, EBOV RNA has been isolated from three species of megachiropteran bat: *Hypsignathus monstrosus*, *Epomops franqueti*, and *Myonycteris torquata* (203). The geographic distribution of these bats is known to overlap with the distribution of EBOV in Africa. Recently, a single report of detection of EBOV RNA and anti-EBOV antibodies in the microbat *Miniopterus schreibersii* in West Africa surfaced (208). However, as of this writing, the data have not been formally published. Importantly, Bombali virus, a related ebolavirus was discovered in association with the microbat *Mops condylurus* (3). Another filovirus, Marburg virus, has been isolated on multiple occasions from *Rousettus aegyptiacus*, and experimental infection studies have demonstrated that the animals are productively infected without clinical disease (209, 210). As such, the consensus in the field is that *Rousettus aegyptiacus* is the reservoir host for MARV (205). However, EBOV does not productively infect *Rousettus aegyptiacus* (211). Experimental infection of bats with EBOV has only been successful in two species, *Epomophorous whalbergi* and *Mops condylurus*. In both cases, infection did not produce clinical disease (202). However, these infections were performed with very small group sizes, and no effort has been made to duplicate the results. This apparent lack of clinical disease in bats has not been definitely associated with a single factor or specific set of factors. One popular hypothesis, that bats have dramatically different innate immune responses to filovirus infections, has not been borne out by *in vitro* studies (212). However, it appears that a number of mechanisms may

contribute to the more general resistance of bats to viral infections, including efficient systems for avoiding and dampening excessive inflammatory responses, truly remarkable tissue damage repair mechanisms, and an extremely diverse VDJ gene set (at least 5 times more diverse than that of most other mammals) that may permit bats to produce a wide array of neutralizing antibodies (213).

Owing to the limited number of successful experimental infections that have been performed, little to nothing is known about the behavior of EBOV in its likely bat reservoir. In the aforementioned experimental infections in *Epomophorous whalbergi* and *Mops condylurus*, the virus replicated to a relatively high viremia (about seven logs in one case) in the absence of clinical signs (202). This is a significantly higher viremia than that observed in *Rousettus aegyptiacus* (210), and may represent a mechanism for increasing the probability of transmission. This would be required in the case of EBOV due to the differences in natural history and social dynamics between *Rousettus* bats and many of the bats thought to act as reservoirs for EBOV. *Rousettus* bats are primarily cave-dwelling, and form densely-populated colonies. With the exception of the microbats, all of the bats that have been identified as likely EBOV reservoirs are forest-dwelling, and tend to form smaller, less dense social groups (214, 215). Some forest dwelling bats do live in relatively dense social groups, and the most commonly cited example of a forest-dwelling bat that forms large and dense colonies is *Eidolon helvum*. However, this species is one of the few that can be definitively excluded as a potential EBOV reservoir, as their NPC1 is resistant to EBOV GP binding (214-216). *Hypsignathus monstrosus* in particular forms very small groups, and bats only congregate in significant numbers for mating (217). Regarding pathology, the only reference point is MARV infection of *Rousettus aegyptiacus*, which is associated with little to no gross pathology, and no significant histopathology (210). Viral antigen staining can be identified in a number of organs, but tends to be limited to relatively small foci, within which staining is often observed to be diffuse (210). The liver appears to be a primary location of replication, much as in primates. The salivary glands, bladder, and

large intestines are also productively infected, suggesting multiple potential routes of transmission (210, 218). Oral shedding has been demonstrated (218). Given the already noted proclivity of EBOV for immunologically-privileged sites in humans, it is noteworthy that the testes are indeed infected during MARV infection of *Rousettus* (210).

The epidemiology and ecology of EBOV in its bat hosts is poorly described, and given that the bats that have been implicated as possible EBOV reservoirs share relatively little of their natural histories with *R. aegyptiacus* (214, 215), MARV's relationship with that species is less informative than might otherwise be suspected. Of the bat species with the most compelling evidence linking them to EBOV, most are semi-gregarious forest-dwelling bats. Unlike *Rousettus* bats, these animals tend to form colonies of modest size, and roost primarily in trees (214, 215). This behavior increases the likelihood of contact with members of another susceptible species, versus the cave-dwelling behavior of *Rousettus aegyptiacus*. Although transmission between bats may have a maternal/infant component, intradermal inoculation via bites has consistently been a favored hypothesis (205, 207, 219). The actual mechanism of natural transmission from bats to other species is likely exposure to urine, feces, or saliva via contamination of vegetation and fruits that are subsequently consumed by susceptible species (205, 207). However, it is not unreasonable to assume that contact with bat carcasses could also constitute a significant route of exposure for non-human primates, chimpanzees in particular. Human exposure to EBOV via bats likely occurs primarily through hunting, butchering, and preparing bats for consumption as food (7). There is evidence to suggest that the biannual birthing cycle of the bat species that have been implicated as potential reservoir hosts is connected to the incidence of human outbreaks (220). The relationship between reproduction in bats and filovirus biology is unclear, however (207, 219, 220). Incidents of animal-to-human exposure, whether from primates, bats, or other species are clearly infrequent, as outbreaks have nearly always been the result of a single spillover event (4). With this in mind, and considering that the first known filovirus outbreaks occurred relatively recently and the

number and frequency of such outbreaks at least appears to be increasing (though this may be the result of more efficient detection and communication of outbreaks), it seems likely that their occurrence is caused by specific ecological changes in Africa, likely including deforestation, climate change, and growth of the human population (221).

Evolution and evolutionary history

The age of the family *Filoviridae* and its constituent members has been a matter of some controversy, in part due to the nature of molecular clock analysis (222). Depending upon the methodology used, estimates have ranged from a few thousand years using uncorrected posterior BEAST analysis of extant filovirus genomes (223), to tens of millions of years using endogenous viral elements, so-called ‘viral fossils’ (224-227). There is persuasive evidence that the low end of this range is not credible due to inherent flaws in the methodology used (228), and multiple streams of evidence appear to support the estimates on the upper end of this range. Filovirus-like endogenous viral elements have been discovered in a diverse array of mammals, including marsupials (224-227, 229). Intriguingly, the discovery of an apparent fish-specific filovirus suggests the possibility that filoviruses are in fact hundreds of millions of years old, matching the divergence of modern bony fishes from the other vertebrates (9).

There is convincing evidence to suggest that filoviruses, and the ebolaviruses in particular have had a prolonged evolutionary relationship with bats. The first line of evidence supporting this hypothesis comes from the presence of filovirus-like endogenous viral elements in the genomes of multiple bat species, particularly those in the microbat genus *Myotis* (224, 226, 229). Phylogenetic analysis has found these elements to be basal to the extant mammalian filovirus genera, indicating that they likely originated from a common ancestor of all of these viruses (229). Selective pressure on bat NPC1 genes provides a second line of evidence for the long-term co-evolution of bats and filoviruses (216).

Modern mammalian filoviruses evolve relatively slowly in an absolute sense. This is particularly true of EBOV, with recent isolates remaining approximately 95% identical to the first viruses isolated in 1976 (146). Although there is evidence for a recent bottleneck event, this remarkable apparent stability may be explained in part by the ecology of the virus (223, 228). Bats appear to provide a fairly restrictive replicative environment for filoviruses, and as such may act as a unit of strong purifying selection. This would appear to leave genetic drift as the primary means of evolutionary change for these viruses. Rate estimates are typically in the range of $5-8 \times 10^{-4}$ substitutions/site/year (146, 223). However, this is extremely variable within outbreaks, due to the shift in population growth dynamics during an outbreak (230). Outbreaks represent a highly aberrant evolutionary pattern for EBOV (230). Outbreaks have historically been initiated by a single spillover event, representing a significant genetic bottleneck, followed by sustained transmission in humans (4, 230). Though population growth in an outbreak is initially linear, it quickly becomes exponential in a large outbreak (4, 230). Like all viruses, exponential population growth tends to increase the apparent evolutionary rate somewhat, an effect observed during the 2013-2016 West African epidemic (230, 231). Perhaps due to this genetic stability, EBOV has tended to exhibit relatively little adaptive evolution during those outbreaks which have been adequately sampled (230, 232). Even during the West African epidemic, where adaptive evolution does indeed appear to have occurred, several months of human transmission were required before the first of the identified adaptive mutations appeared (233, 234). However, prolonged human passaging did induce changes in GP that reduced viral tropism for bat cells relative to human cells (233).

One of the major issues with conducting phylogenetic analysis of EBOV evolution is the limited nature of the available datasets. Unfortunately, as a general rule, available full-genome sequences have originated from human outbreaks, many of which have extensive human passage histories plus at least one in vitro passage. This presents an obvious problem when attempting to reconstruct the evolutionary history of EBOV, as the

entirety of the enzootic evolution of the virus must be inferred, and there is the potential for interference resulting from human transmission and the evolutionary dynamics of human outbreaks. This makes environmental sampling an imperative for the understanding of the evolution of the virus. Wittmann et al. conducted a field surveillance program that allowed the authors to collect samples from the carcasses of wild animals, predominantly apes, which had succumbed to EBOV (235). The authors obtained a number of sequences from the GP and NP ORFs, but unfortunately did not sequence full genomes. Analysis of the full-length GP sequences from this set by Dudas et al. found that they were very diverse, and nested within the phylogeny of EBOV (235, 236). This supports the notion that zoonotic EBOV exists as a highly diverse population of genomes. Sampling solely from outbreaks reveals only a minuscule fraction of this diversity as it enters the human population in spillover events, making it difficult to understand the evolutionary history of the virus in any meaningful way. Moreover, it complicates attempts to understand important divergence events. Finally, recent work has suggested that the increasing pace of deforestation in Africa is a major contributory factor to human outbreaks of EVD, a suggestion that has been borne out by the increasing frequency of EVD outbreaks (221).

DESCRIPTION OF THE INTERACTIONS OF EBOLA VIRUS WITH A PUTATIVE RESERVOIR SPECIES USING NEXT-GENERATION SEQUENCING

Rationale

As presumptive reservoir hosts of EBOV (207), and known reservoir hosts of MARV (209), bats have been a topic of intense interest within the field of filovirology. However, despite over a decade of interest and dozens of publications, the vast majority of the biology of the bat/virus nexus remains unexplored, particularly on the cellular level. Small RNAs play significant roles in virus/vector and virus/reservoir relationships for a number of other important human pathogens (237-240), and a small but growing body of literature indicates that EBOV produces small RNAs of unknown function (241-246). This represents an area of interest worthy of exploration. Another important factor in virus/vector and virus/reservoir relationships is the nature of the selective pressures imposed upon the virus by the vector/reservoir, which often includes host enzymes that directly edit the viral genome (247, 248). As a result, the evolution of the virus is affected by these factors (231, 249). A convenient and powerful tool for exploring both of these phenomena is next-generation sequencing, which has become more powerful and flexible over the past decade as novel library preparation techniques have been developed that permit the efficient selective capture and sequencing of small RNAs, and the characterization of entire viral populations at incredible levels of detail. This work approaches the bat/virus nexus with the intent of exploring the biology of small RNAs in EBOV infection, and the effects of bat biology on the evolution of the virus. These basic elements of the biology of EBOV in bats are crucial to the broader context of the bat/virus nexus.

**Ebola virus produces discrete small non-coding RNAs independent of
the host microRNA pathway and which lack RNA interference activity
in bat and human cells**

INTRODUCTION

Micro-RNA biology

Although only discovered relatively recently, microRNAs have been found to be crucial post-transcriptional regulators of gene expression in animals and plants (250). Eukaryotic endogenous miRNAs are transcribed in the nucleus from defined miRNA loci, normally by RNA polymerase II (251), although some are transcribed by RNA polymerase III (252). This produces primary miRNAs (pri-miRNAs) (253). These consist of multiple connected hairpin structures that must be processed by the microprocessor complex, also located in the nucleus. This complex, comprised of Drosha (an RNase III enzyme) and DGCR8 processes the pri-miRNAs into individual hairpin structures, known as pre-miRNAs (254, 255). Some RNA editing of pre-miRNAs does occur in the nucleus, affecting up to 16% of pre-miRNAs (256). Following processing by the microprocessor complex, pre-miRNAs are exported to the cytoplasm in a GTP-dependent process by the karyopherin exportin-5, which recognizes a 3' 2 nt overhang left by Drosha (257). In the cytoplasm, another RNase III enzyme, Dicer, recognizes the ends of the hairpin structure and removes the "loop" of the pre-miRNA hairpin, leaving the stem of the hairpin as a pair of 22nt RNAs known as the miR/miR* duplex (258, 259). Base pairing in this duplex is normally imperfect. By convention, the RNA originating from the 5' end of the pre-miRNA is referred to as the 5p strand, while the RNA originating from the 3' end is designated as the 3p strand (253). Either strand may function as the ultimate mature miRNA. At this point, Dicer loads the miR/miR* duplex into the RNA-induced silencing complex (RISC) (260). Although either strand can serve as the mature miRNA, selection appears to be based

upon thermodynamic stability (261). The selected strand, now referred to as the guide strand, is retained in RISC while the other strand, known as the passenger strand is ejected from RISC, after which it is quickly degraded (260, 261). The mature miRNA guide strand is used by RISC to recognize cognate mRNAs, which are identified via sequence complementarity with the guide strand. Generally, if sequence complementarity is perfect, the result is degradation of the transcript by the slicing activity of Argonaute, the primary protein component of RISC (260, 262). Imperfectly complementary mRNAs are not normally degraded, but are instead regulated by translational repression, a process by which translation is blocked by preventing association of the ribosome complex with the mRNA (263). **Illustration 5** depicts these processes. Although the vast majority of mammalian miRNAs are produced via this canonical process, at least one non-canonical pathway exists that is independent of Dicer, and is responsible for the biogenesis of the human miRNA hsa-miR-451 (264).

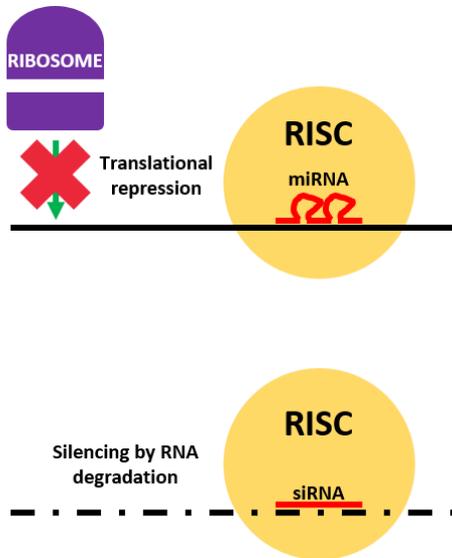


Illustration 5: RNA interference mechanisms.

As would be expected from any mechanism involved in post-transcriptional regulation of protein production, miRNAs play important roles in the maintenance of homeostasis, and in disease. Roles for miRNAs and the regulation of miRNA production

have been described in multiple disease states, including neurodegenerative disease and cancer (265, 266). Involvement of miRNAs in the pathogenesis of infectious diseases is also well-described, including a number of viral infections (237-240). At least one, hepatitis C virus, utilizes a host miRNA (hsa-miR-122) in its life cycle to enhance translation, shield the viral genome from host factor-mediated degeneration, and promote genomic stability (267-270). Several DNA viruses (herpesviruses) and some retroviruses actually produce their own miRNAs as a means of facilitating infection (238, 239, 271, 272). By contrast, non-reverse transcribing RNA viruses have been thought incapable of producing miRNAs, in part due to the small size of their genomes, among other factors (273, 274). This assertion has been tested within the past decade, particularly in the case of positive strand RNA virus, along with other findings of miRNA-like molecules produced by certain flaviviruses (275-278). The latter reports have been controversial, however (279, 280). Although not miRNAs in the conventional sense, influenza A virus (IAV) produces small viral non-coding RNAs (vncRNAs) that have roles in replication. IAV also produces a small RNA with miRNA-like qualities that has been validated functionally (281-283). Although the notion of RNA virus miRNAs is a controversial one, this growing body of literature at a minimum makes a strong case for caveating the absolute statement that RNA viruses do not produce miRNAs.

RNA interference and its role in viral biology

The miRNA pathway is one type of RNA interference (RNAi) pathway. The other commonly discussed mechanism is the small interfering RNA (siRNA) pathway (284), often described as a mechanism of antiviral defense in plants and invertebrates (285). Briefly, small interfering RNAs (siRNAs) are produced from long double stranded RNAs (dsRNAs) via Dicer processing, and loaded into RISC, after which strand selection occurs in a mechanism not unlike that previously described for miRNAs (286). However, siRNA-loaded RISC complexes nearly always degrade their targets via slicing, as base

complementarity is usually perfect (286). In plants and invertebrates, this is a potent antiviral effector mechanism (285). For this reason, viruses that infect plants and invertebrates typically possess at least one VSR (287, 288). These act by directly inhibiting the function of some part of the RNAi pathway, often Dicer or a component of RISC (287, 288). In mammals, however, the role of RNAi in antiviral immunity has been and remains controversial (289-295). This is because the general consensus in the field has been that RNAi has been largely displaced as a primary component of the innate immune response in vertebrates, including mammals, by mechanisms centered upon interferon (IFN) (296). This view is bolstered by the fact that vertebrates have lost the second Dicer enzyme used by invertebrates to process siRNAs. In invertebrates, Dicer-1 functions largely within the miRNA pathway, whilst Dicer-2's role is mostly restricted to the siRNA pathway, where it also serves as a dsRNA sensor (260). The single Dicer enzyme possessed by mammals is homologous to invertebrate Dicer-1, and biochemical studies suggest that its processing of long dsRNAs into siRNAs is very inefficient when compared to invertebrate Dicer-2 (297-299). Moreover, a number of older efforts have not been successful in producing convincing evidence for mammalian RNAi (289, 300-302). In contrast, recent work has, in fact, convincingly described a functional siRNA-mediated antiviral pathway in mammalian cells following RNA virus infection (303-306).

Ebola virus and small RNAs

As an enveloped nonsegmented negative strand virus, EBOV belongs to the order *Mononegavirales*. To date, no member of this order, which includes a large number of human pathogens, has been definitively described as producing valid and biologically functional miRNAs. Recently though, a number of reports based upon *in silico* modeling have suggested that EBOV may produce pre-miRNAs and mature miRNAs. Additional reports have validated the existence of most of these molecules via detection by sequencing and/or RT-qPCR of material obtained from EBOV-infected humans, nonhuman primates,

and rodents, along with experimentally infected human retinal epithelial cells (241-246). However, each of these reports stopped short of making an attempt to characterize the biogenesis of these molecules or determine if these form an association with RISC, much less have any biological function. In another line of evidence for some level of involvement of small RNAs in the biology of EBOV, the virus is known to possess three proteins (VP35, VP30, and VP40) with some level of reported VSR activity (38, 39, 307). Some evidence suggests that VP35 is a particularly potent VSR (38, 307). Given that most other viruses that possess such potent VSRs infect species that utilize RNAi as an antiviral mechanism (287, 288), this is curious.

For these reasons, we believed that investigation of the potential role of miRNAs and RNAi in bats and humans was warranted. To profile the small RNA response to filovirus infection in bat and human cells, we used next-generation sequencing (NGS) technology to analyze the viral small RNA (vsRNA) populations of two bat cell lines and one human cell line after infection with EBOV and MARV. Both viruses produced multiple vncRNAs in high abundance. We designed a series of experiments to explore the biological origin of the EBOV vncRNAs and any potential function that they may have as virally-derived miRNAs. We found that the EBOV vncRNAs are produced via a mechanism independent of Dicer, but which may involve the Integrator complex, and which do not appear to interact with any of the Argonaute (AGO) family of RISC catalytic proteins. As would be expected from molecules that do not interact with RISC, no evidence was found to suggest that the EBOV vncRNAs have roles in either host transcript repression or silencing. Furthermore, the EBOV vncRNAs were incapable of acting as antiviral siRNAs. Considered in total, our findings contradict some of the conclusions of prior *in silico* studies, the identified vncRNAs are almost certainly not products of host RNAi machinery, and have no role in the suppression of host transcripts or of viral replication. This work serves to emphasize the importance of thorough biological validation of *in silico* findings,

as two of the EBOV vncRNAs identified closely match those predicted by prior *in silico* work which were assumed to be functional as viral miRNAs.

MATERIALS AND METHODS

Cell lines and maintenance

RO6E/J cells were a kind gift of Dr. Ingo Jordan (CureVac AG, Germany) and were maintained in DMEM/F12 + GlutaMAX (Thermo Fisher) complete media (10% heat-inactivated [56°C/30 min] and 50 mg/L gentamicin sulfate [Cellgro] final concentration) at 37°C, 5% CO₂. EpoNi/22.1 cells were a kind gift of Dr. Christian Drosten (The Charité – Universitätsmedizin Berlin) and were maintained identically to RO6E/J cells. HepG2 cells were obtained from ATCC (HB-8065) and were maintained in DMEM low glucose (1 g/L) complete media (Thermo Fisher) at 37°C, 5% CO₂. HEK 293T/17 (herein referred to as “293T”) cells were obtained from ATCC (CRL-11268) and maintained in DMEM High Glucose (4.5 g/L) complete media (Thermo Fisher) at 37°C, 5% CO₂. 769-P cells were acquired from ATCC (CRL-1933) and were maintained in RPMI-1640 complete media (Thermo Fisher) at 37°C, 5% CO₂. Dicer knockout NoDice 4-25 and parental cell line (referred to as “293T-P” to differentiate between 293T cells obtained from ATCC) were a kind gift of Dr. Bryan Cullen (Duke University School of Medicine) and were maintained identically to 293T cells obtained from ATCC. BHK-21 C13 (CCL-10) and Vero E6 (CRL-1586) cells were obtained from ATCC and maintained identically to HepG2 cells.

Viruses

All experiments involving infectious EBOV and MARV were conducted in the Galveston National Laboratory Biosafety Level 4 (BSL-4) facility by trained staff with the appropriate U.S. government permissions and registrations for work with the EBOV and MARV full-length cDNA clones (FLC) and viruses. Recombinant EBOV (strain Mayinga)

wt and VP35 R312A mutant viruses expressing eGFP were rescued and propagated as described previously (153) using EBOV FLC provided by Drs. John Towner and Stuart Nichol (CDC) and the EBOV NP, VP35, L, VP30, and T7 polymerase plasmids provided by Drs. Yoshihiro Kawaoka (University of Wisconsin) and Heinz Feldmann (NIH). The EBOV wt FLC lacking the eGFP transgene was prepared by restriction digest of the eGFP-flanking BsiWI sites followed by re-ligation of the linearized vector. The MARV reverse genetics system used in these studies (308) was a kind gift of Drs. Jonathan Towner and Stuart Nichol (CDC). Plasmids containing the MARV FLC expressing eGFP (strain Uganda 371Bat2007, isolate 811277), rodent-cell codon optimized NP, VP30, VP35, and L genes, and T7 polymerase were transfected into BHK-21 C13 cells using the *TransIT* LT1 transfection reagent (Mirus Bio). Rescued viruses (passage 0) were further amplified by two to three passages in Vero E6 cells to generate working stocks for infection experiments. The VP35 R301A virus was generated by sub-cloning the VP35-containing AfeI-KasI fragment of the FLC into pCAGGS MCS. Site-directed mutagenesis was performed using the QuickChange kit (Agilent), after which the mutated fragment was re-inserted into the FLC and rescued as described above. Recovered viruses were titrated in Vero E6 cell monolayers. Plaques were visualized 11-12 days post-infection by crystal violet staining.

Experimental infections

To generate samples for small RNA sequencing, 90% confluent monolayers of RO6E/J, EpoNi/22.1, or HepG2 cells were inoculated with EBOV and MARV viruses at an MOI of 2 PFU/cell, or mock (complete media)-infected in biological triplicate, allowed to adsorb for 1 h, washed twice with cold PBS, and then replenished with fresh complete media. At 12 hpi and 24 hpi, cell supernatants were collected and frozen at -80°C for subsequent titration, and cell monolayers were lysed in 1.0 mL of Trizol reagent for removal from the BSL-4 and RNA isolation. Titration of supernatants was performed as

for virus stocks. For RT-qPCR comparison of EBOV miRNA abundance in Dicer-competent versus Dicer knockout cells, cells were inoculated as described above with MOI 2 PFU/cell of the EBOV wt-eGFP virus or mock infected, with cell monolayers being collected in Trizol at 20 hpi. Procedures for other experiments involving live filoviruses are described separately.

Virus Inactivation and RNA Extraction with Trizol

After removal of supernatant, monolayers of filovirus infected or mock infected cells were treated with 1 mL Trizol (Ambion), incubated for 10 min at room temperature, and removed from the BSL-4. 200 μ L of chloroform per 1 mL of Trizol used was added to each sample and the samples were incubated at room-temperature for 5 min, after which they were centrifuged at 12,000 xg for 20 min. The aqueous phase was removed to fresh nuclease-free 1.5 mL tubes, to which 15 μ g of linear acrylamide (Ambion) and 1.0 mL of 2-propanol per mL of Trizol used was added. Samples were incubated at -20°C overnight to precipitate RNA. Following incubation, the samples were centrifuged at maximum speed (16,900 xg) at 4°C for 30 min, and then washed three times with 80% ethanol (each wash followed by centrifugation at maximum speed for 30 min at 4°C). Precipitated RNA pellets were allowed to air dry for \sim 5 min before being re-suspended in nuclease-free water. Re-suspended total RNA was quantified and assessed for purity by Nanodrop (Thermo Fisher).

Small RNA Library Preparation

Total RNA integrity was assayed using an RNA 6000 Nano chip on the Agilent 2100 Bioanalyzer. All RNAs used to make libraries had an RNA integrity number (RIN) \geq 8. Small RNA libraries were made using the TruSeq Small RNA Sample Prep kit (Illumina) as per manufacturer's suggested protocol. Briefly, 1 μ g of total RNA was 3' and 5' adapter-ligated, reverse-transcribed using Superscript II (Invitrogen), and then PCR-amplified, during which unique index sequences were added to the libraries. PCR-

amplified libraries were electrophoresed on 2.5-3% agarose gels, with bands corresponding to ~18-35 bp inserts being excised. The size-selected libraries were gel purified using the Monarch Gel Extraction Kit (New England Biolabs). Purified libraries were validated for sequencing by RT-qPCR, Qubit, and Bioanalyzer, and then sequenced single-end for 50 cycles on an Illumina HiSeq 2500 on rapid run mode. Detailed information regarding total reads passing filter, read quality information, and virus alignment counts can be found in Appendix A.

Bioinformatics

For small RNA analysis, FASTQ files containing the raw sequence reads were trimmed of the 3' adapter sequence using FASTX Toolkit , size-selected to only include reads between 19-32 bp in length, and aligned to the appropriate infectious clone reference genome using Bowtie v. 0.12.8 (309) in the -a --best --strata mode, allowing for a single mismatch. SAM output files from Bowtie were fed into SAMTools (310) to generate SAM, BAM, and mpileup files. To determine vsRNA fold change from 12 hpi to 24 hpi, for each library vsRNAs were normalized by dividing the total number of 19-32 nt vsRNA reads by the total number of 19-32 nt reads to obtain a vsRNA frequency. The following equation was then used to generate a vRNA-normalized vsRNA frequency:

$$x = \left\{ \left[s_f \div \left(\frac{s_1 + s_2 + s_3}{3} \right) \right] \div \left[v_f \div \left(\frac{v_1 + v_2 + v_3}{3} \right) \right] \right\} \div \frac{r_1 + r_2 + r_3}{3}$$

where s_f is the vsRNA frequency of a single biological replicate at either 12 or 24 hpi, s_{1-3} are the vsRNA frequencies for each of three biological replicates at 12 hpi, v_f is the vRNA abundance for a single replicate at 12 or 24 hpi, v_{1-3} are vRNA abundance for each of three biological replicates at 12 hpi (this equation calculates the vRNA-adjusted vsRNA fold change), and r_{1-3} are the individual vsRNA fold change for each of three biological replicates at 12 hpi. Normalization for individual vncRNA abundance in EpoNi/22.1 and RO6E/J cells versus HepG2 cells was performed using the following equation:

$$x = \left[\left(\frac{n_i}{o_i} \times 10^6 \right) \div \left(v_i \div \frac{v_1 + v_2 + v_3}{3} \right) \right] \div \frac{r_1 + r_2 + r_3}{3}$$

where n_i is the number of 19-32 nt reads for a given vncRNA for a given library, o_i is the total number of 19-32 nt reads processed for a given library (reads per million, or RPM was calculated by multiplying the resulting value by 106), v_i is the vRNA copy number for a given library (per ng total RNA), v_{1-3} are the vRNA copy numbers (per ng total RNA) for each HepG2 biological replicate (this equation results in the vRNA-adjusted RPM), and r_{1-3} is the vRNA-adjusted RPM for each of three HepG2 biological replicates. To determine fold change from 12 hpi to 24 hpi for individual vncRNAs, we used the following equation:

$$x = \left[\left(RPM_f \div \frac{RPM_1 + RPM_2 + RPM_3}{3} \right) \div \left(v_f \div \frac{v_1 + v_2 + v_3}{3} \right) \right] \div \frac{r_1 + r_2 + r_3}{3}$$

where RPM_f is the RPM for a given vncRNA at 12 or 24 hpi from an individual biological replicate, RPM_{1-3} are the RPMs for each biological replicate at 12 hpi, v_f is the vRNA copy number (per ng total RNA) for an individual biological replicate at 12 or 24 hpi, v_{1-3} are the vRNA copy numbers (per ng total RNA) for each of three biological replicates at 12 hpi (this equation calculates the vRNA-adjusted fold change for a given vncRNA), and r_{1-3} are the individual vRNA-adjusted fold change values for each of three biological replicates at 12 hpi. Additional analysis and statistics were performed using viRome, a package for R (311), Microsoft Excel, and Graphpad Prism 6. Statistics for all experiments were performed using Graphpad Prism 6. Unless otherwise noted, p-values for all statistical analyses are two-tailed.

miRNA-specific RT-qPCR

Primer sequences and/or unique ID numbers for all RT-qPCR experiments are provided in Appendix C. SYBR Green RT-qPCR was performed using the miRCURY LNA RT and miRNA PCR System kits (Qiagen) per the manufacturer's suggested

protocol. Briefly, 10 ng input total RNA (unless otherwise specified) was poly-A tailed and then reverse transcribed with a universal miRNA RT-primer, which adds a tag sequence to the RNAs. A kit-provided synthetic miRNA spike-in (UniSp6) was used in all reactions as a control for monitoring efficiency of reverse transcription. The resulting miRNA cDNA libraries were diluted 1:60 in nuclease free water, and 3 μ L of diluted cDNA (50 pg) was used as the input for triplicate RT-qPCR reactions using EBOV miRNA-specific primers or primers to endogenous (hsa-miR-103a-3p) and exogenous (UniSp6) miRNAs. Minus reverse-transcriptase (-RT) and no template controls (NTC) were also included for all primer sets. RT-qPCR was performed on Step One Plus or QuantStudio 6 Flex Real-Time PCR systems (Thermo Fisher) using cycling settings specified in the miRCURY LNA miRNA PCR protocol, and data was collected in the associated StepOne or QuantStudio software (Thermo Fisher). The UniSp6 synthetic miRNA was used as the reference miRNA, since its “expression” was consistent between all biological and RT-qPCR technical replicates and experimental conditions, and spike-ins have been shown to be reliable references for RT-qPCR normalization (312, 313). Relative fold-change between samples was calculated in Microsoft Excel using the $2^{-\Delta\Delta C_t}$ method of approximation. For the purposes of determining fold change, targets without Ct values (i.e. undetectable) were arbitrarily assigned a Ct value of 40 (the last cycle). Thus, fold change approximations represent a minimum fold change comparison. For absolute quantitation, a synthetic single-stranded RNA oligonucleotide exhibiting 5'-monophosphate and 3'-hydroxyl moieties and homologous to the 22 nt form of the EBOV GP vncRNA (Integrated DNA Technologies, IDT) was spiked into human brain total RNA (Takara Bio) at a concentration of 0.1 ng/10 ng total RNA (corresponding to $\sim 8.35 \times 10^9$ copies of the synthetic GP vncRNA per 10 ng total RNA). 10 ng of spiked human brain total RNA was reverse transcribed as described above, diluted 1:60 in nuclease free water, and then further serially diluted in 1:10 increments. 3 μ L of each serial dilution were then used as standards for absolute quantitation.

Detection of viral vRNA/cRNA/mRNA by TaqMan (hydrolysis probe) RT-qPCR

Tag-based, strand-specific RT-qPCR assays were designed for EBOV as previously described (314). After RNA extraction, first strand cDNA synthesis was performed using the Superscript IV system (Thermo Fisher Scientific) and followed a modified version of the manufacturer's protocol. 20 ng of RNA was used as input. Primer annealing was performed in the absence of dNTPs, which were added to the reverse transcription master mix. The reverse transcription reaction was performed at 50°C. Following reverse transcription, qPCR reactions were set up with the TaqMan Universal PCR Master Mix (Thermo Fisher Scientific) using 4 ng of cDNA. Custom primer-probe sets (FAM-labeled) were obtained from Thermo Fisher Scientific. AT-rich flaps were added to the primers to improve efficiency, as previously described (315). Absolute quantitation was performed based upon DNA oligonucleotide standards obtained from Integrated DNA Technologies.

RNA-binding protein Immunoprecipitation (RIP) Assay

Near-confluent (~90%) T225 flasks of 293T cells were either infected with rEBOV wt-eGFP virus at an MOI 2 or mock-infected. At 20 hpi, cells were washed two times with cold PBS and lysed in Magna RIP Lysis Buffer with protease and RNase inhibitors added (EMD Millipore), and subjected to one freeze-thaw cycle at -80°C. A single T225 flask inoculated with rEBOV wt-eGFP was used for all target protein and control immunoprecipitations in a given independent experiment. Prior to immunoprecipitation, 10 µL volumes of the rEBOV wt-eGFP-infected and mock-infected lysates were collected to use as pre-IP input controls for qPCR. Argonaute (Ago) proteins were immunoprecipitated from lysates corresponding to ~ 1.0-2.0 x 10⁷ cells using 5 µg of each antibody (Appendix C) and the EZ-Magna RIP RNA-Binding Protein Immunoprecipitation Kit (EMD Millipore) as per manufacturers suggested protocol. IgG isotype control immunoprecipitation was performed using 5 µg of the kit supplied normal mouse IgG. Immunoprecipitations were performed overnight at 4°C with gentle agitation.

Following proteinase K digestion, 1 mL of Trizol was added to the RNA/protein eluates and pre-IP lysates for residual virus inactivation and removal from the BSL-4. Purified RNA from immunoprecipitations were too low in concentration to be quantified using the Qubit HS RNA Assay Kit (Thermo Fisher), so equal volumes of each immunoprecipitation (6.5 μ L) were used as the input for the miRCURY LNA RT Kit. For pre-IP input RNA, 10 ng of total RNA was used as the input for reverse transcription.

siRNA-mediated knockdown

A list of siRNAs utilized, their manufacturers, and their final concentrations used in our experiments can be found in Appendix C. For siRNA transfection, we used TransIT-siQuest reagent (Mirus Bio). To determine transfection efficiency, the siGLO fluorescent reporter (Dharmacon) was used in tandem with all transfections. All siRNA transfections were done in triplicate.

Western blot

Antibodies, manufacturers, and dilutions are provided in Appendix C. 100 μ g of RIPA lysate were run on 4-12% Bis-Tris PAGE gels (Invitrogen), and dry blotted onto nitrocellulose membranes using the iBlot 2 system (Invitrogen). Membranes were blocked in 5% non-fat dry milk dissolved in 1X TBS with 0.1% final concentration Tween-20 (TBS-T) for 1 h. Membranes were incubated overnight at 4°C with the primary antibody diluted to the indicated concentration in either 5% BSA dissolved in 1X TBS/0.1% Tween-20 or blocking buffer, depending on manufacturer's suggested protocol. Membranes were then washed three times in TBS-T, after which the secondary antibody was diluted in blocking buffer and allowed to incubate with the membranes for 1 h at room temperature. Following washing with TBS-T, membranes were developed using Amersham ECL Western Blotting Reagent (GE Healthcare Life Sciences), and imaged on an Odyssey Fc imaging system (Li-cor Biotechnology) using 10 min exposures in the chemiluminescent

and 700 nm channels. Band pixel density analysis was performed using Li-cor Image Studio Lite Ver 5.2.

Focus-forming/plaque immunostaining assays

Antibodies used for plaque immunostaining can be found in Appendix C. Confluent monolayers of Vero E6 were inoculated with serially diluted samples and allowed to adsorb for 1 h at 37°C, after which a 0.6% methylcellulose/2% FBS MEM overlay was added. After 5 days incubation, monolayers were fixed with 10% formalin solution and incubated overnight in the BSL-4 prior to being removed. Fixed monolayers were blocked with 5% milk in TBS-T for 1 h, washed, incubated with primary antibody for 1 h, washed, incubated with secondary antibody for 1 h, washed, and developed using 4-CN peroxidase substrate reagent (Seracare).

Luciferase reporter assays

We utilized the pMirGLO Dual Luciferase Reporter assay system (Promega), which expresses both firefly luciferase (FLuc, the “reporter”), and *Renilla* luciferase (hRLuc) as an internal control for normalization. Artificial 3'-UTRs were designed by generating a 456 nt scrambled sequence lacking homology to any known sequence in GenBank, and embedding two miRNA binding elements (MBE) within the sequence, for a total insert length of 486 nts. MBEs were designed to be perfectly complementary to the EBOV GP vncRNA, to contain mismatched base pairs within the center of the sequence (outside of the seed, “imperfect”), or to be complementary to only the first 10 nts of the EBOV GP vncRNA (all sequences provided in Appendix C). Synthetic dsDNA fragments (“gBlocks”) were synthesized by IDT. Artificial 3'-UTRs were cloned immediately downstream of the FLuc reporter. For live virus experiments, 769-P and EpoNi/22.1 cells were transfected with each of the reporter plasmids using TransIT LT-1 reagent (Mirus), incubated for 4 h, and then infected with rEBOV wt virus at MOI 2 PFU/cell. 24 hpi, cells

were lysed in passive lysis buffer, subjected to a -80°C freeze/thaw cycle, and assayed on a luminometer using the Dual Luciferase Reporter Assay System (Promega), with statistical comparisons being made to the empty vector plasmid (no 3'-UTR). A control experiment was performed by transfecting a synthetic siRNA homologous to the EBOV GP vncRNA (Dharmacon) at a final concentration of 50 nM 4 h post transfection of the reporter constructs. siRNA transfection was performed using the Trans-IT siQuest reagent (Mirus). Cells were lysed 24 h post siRNA-transfection and processed identically to cells infected with live virus. Statistical comparisons for the siRNA experiment were made to cells transfected with both the pMirGLO construct containing the perfectly complementary MBE and a scrambled siRNA. For both experiments, a 10-second integration step was used for both FLuc and hRLuc readings, and FLuc luminescence values were normalized to corresponding hRLuc values. Both experiments were performed independently twice, with each treatment being performed in quadruplicate.

RESULTS

Virus-derived sRNA profiles of filovirus-infected bat and human cell lines

We profiled the EBOV and MARV-derived vsRNA populations in a human hepatocarcinoma-derived cell line (HepG2) and two bat cell lines, RO6E/J, (*Rousettus aegyptiacus* fetal tissue-derived) (316), and EpoNi/22.1, *Epomops buettikoferi* (adult kidney-derived) (317). Both bat cell lines have previously been shown to be susceptible to EBOV and MARV infection, though the growth kinetics of these viruses differ both between each other and in comparison to Vero E6 cells (317, 318). HepG2 cells were selected despite not being tissue-matched as they yield more and better quality RNA per cell than comparably immunocompetent human renal cell lines in our hands. This is important in this context due to the need for high quality RNA for optimal sequencing library preparation. Cells were infected for 12 and 24 h with recombinant wild-type EBOV and MARV, or mutants exhibiting a point mutation in VP35 that disables the interferon

antagonism and dsRNA-binding properties of the protein (36, 319, 320); derivatives of these viruses expressing eGFP (308, 321) were used for visualization of infections. Total RNA was harvested at either 12 or 24 h post infection (hpi). Cell lines varied in their susceptibility to filovirus infections, as measured by plaque assay titration of cell culture supernatants (**Fig 1. A,E**), RT-qPCR of intracellular viral genomes and mRNA (**Fig. 1B,C,D,F**), and by visualization of eGFP signal by fluorescence microscopy (data not shown). Both EBOV and MARV wt viruses tended to outgrow VP35 mutant viruses at one or both time points, with the exception of MARV wt at 24 hpi in EpoNi/22.1 cells, which had a significantly lower titer than the VP35 mutant virus. The ratio of EBOV VP40 mRNA to vRNA, a measure of transcriptional activity also varied significantly between human and bat cell lines and between time points in both bat cell lines (**Fig. 1D**).

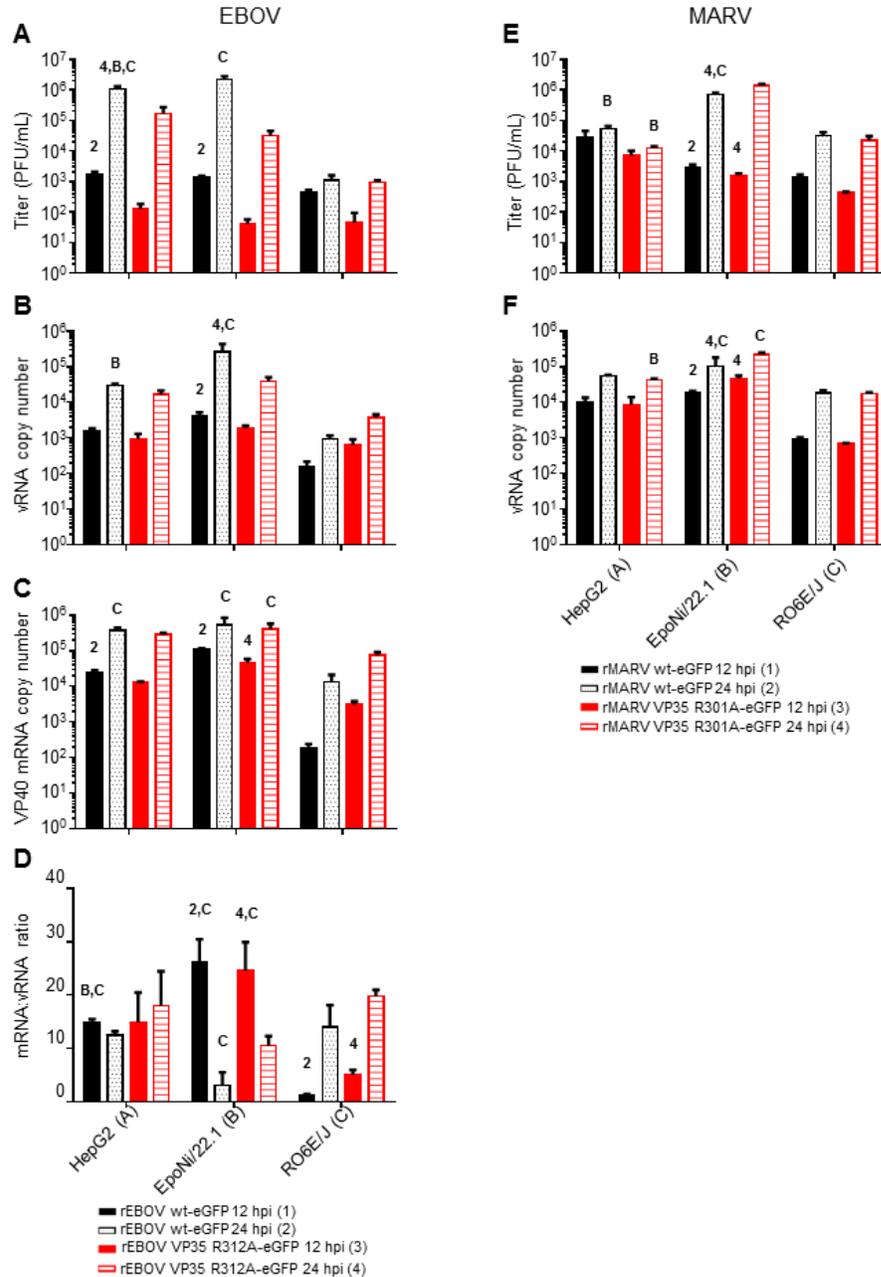


Figure 1: Viral titers and RNA abundance from human and bat cell lines following infection with EBOV and MARV

Supernatants from EBOV- (A) or MARV- (E) infected cells used to make sRNA sequencing libraries were assayed by plaque titration. Total RNA from cell monolayers were assayed by RT-qPCR for the presence of vRNA/cRNA and mRNA for EBOV (B,C) or only vRNA/cRNA for MARV (F). RNA quantities are expressed as copy number per ng total RNA. For EBOV, the ratio of mRNA to vRNA was calculated for each virus and timepoint (D). For all panels, time points were assigned numbers, and cell lines were assigned letters. Within a cell line, comparisons were made between 12 and 24 h time

points for each virus or between viruses within a time point. Comparisons between cell lines were made for each virus at a given time point. Data plotted represents the means of three biological replicates \pm SD. Numbers or letters above bars indicate significant differences between either time points (assigned with the corresponding number) or cell lines (assigned with the corresponding letter) as measured by 2-way ANOVA ($\alpha=0.05$), followed by Tukey's multiple comparisons post-test. For a comprehensive list of comparisons and p values, please refer to Appendix B.

After creating cDNA libraries and deep sequencing, the adapter-trimmed libraries were aligned against the reference genome for each virus. Bioinformatics analyses of sequencing data are summarized in **Fig. 2**.

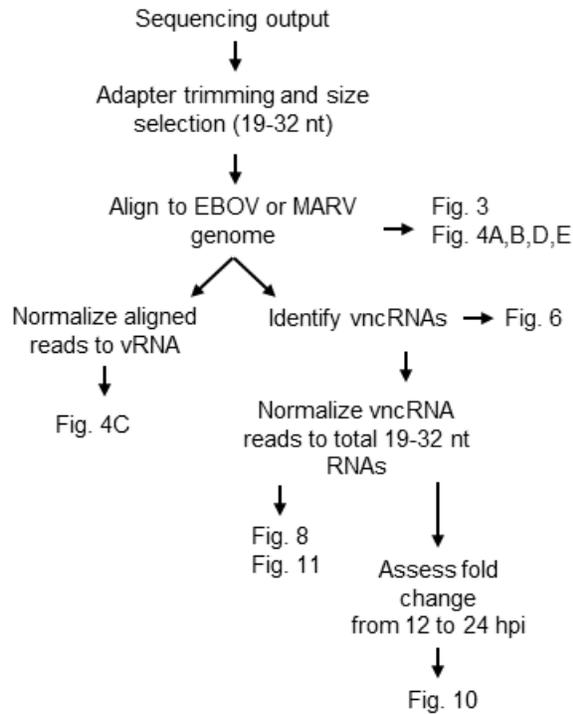


Figure 2: Schematic of bioinformatics analyses of sequencing data

At 12 hpi, vsRNAs were predominately 22 nts in length from all cell lines infected with both EBOV (**Fig. 3A-C**) and MARV (**Fig. 3D-F**). At 24 hpi, vsRNAs were more evenly distributed by length in EBOV-infected cells; however, they were almost entirely composed of 22 nt reads in MARV-infected cells. In all libraries from EBOV-infected cells, the percentage of 19-32 nt vsRNAs relative to the entire population of 19-32 nt

sequences in each library increased significantly from 12 to 24 hpi (**Fig. 4A**), though only EpoNi/22.1 and RO6E/J cells showed an increase across timepoints following infection with MARV (Fig 4B). Normalization of vsRNA abundance to viral genomic RNA (vRNA) (**Fig. 4C**) showed that the relative proportion of vsRNAs remained static from 12 to 24 h in EBOV-infected HepG2 and EpoNi/22.1 cells; however, they dramatically increased in relative proportion in RO6E/J cells. It should be noted that in the case of rEBOV wt-eGFP-infected cells, vsRNAs were barely detectable in RO6E/J cells at 12 hpi (<100 reads per biological replicate). vsRNAs derived from MARV tended to decrease in relative abundance in all cell lines, though only significantly so for rMARV wt-eGFP in HepG2, rMARV VP35 R301A-eGFP in EpoNi/22.1, and both viruses in RO6E/J. Across all cell lines, vsRNAs were predominantly derived from the positive (anti-genomic) strand of the viral RNA for all viruses, with the exception of rEBOV wt-eGFP infected RO6E/J cells at 12 hpi (**Fig 4. D,E**). This lack of strong strand bias for rEBOV wt-eGFP was likely due to the extremely limited number of reads aligning to the virus at this time point.

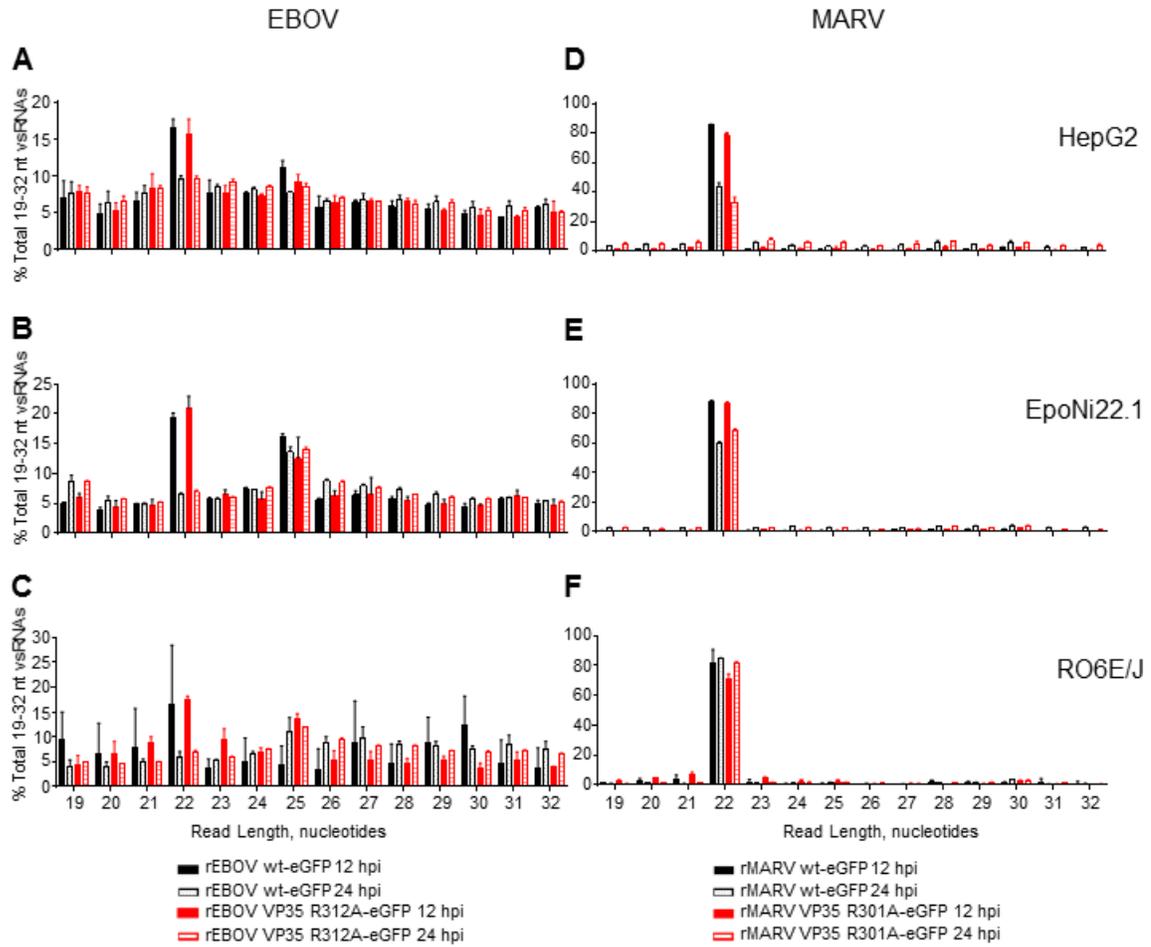


Figure 3: Length distribution profiles of EBOV and MARV vsRNAs in human and bat cells

Reads aligning to the EBOV (A-C) or MARV (D-F) genomes were plotted by length as a percentage of the total number of reads aligned to the viral genomes. The mean of three biological replicate sequencing libraries at each time point \pm SD are plotted. **A)** HepG2, **B)** EpoNi/22.1, **C)** RO6, **D)** HepG2, **E)** EpoNi/22.1, **F)** RO6.

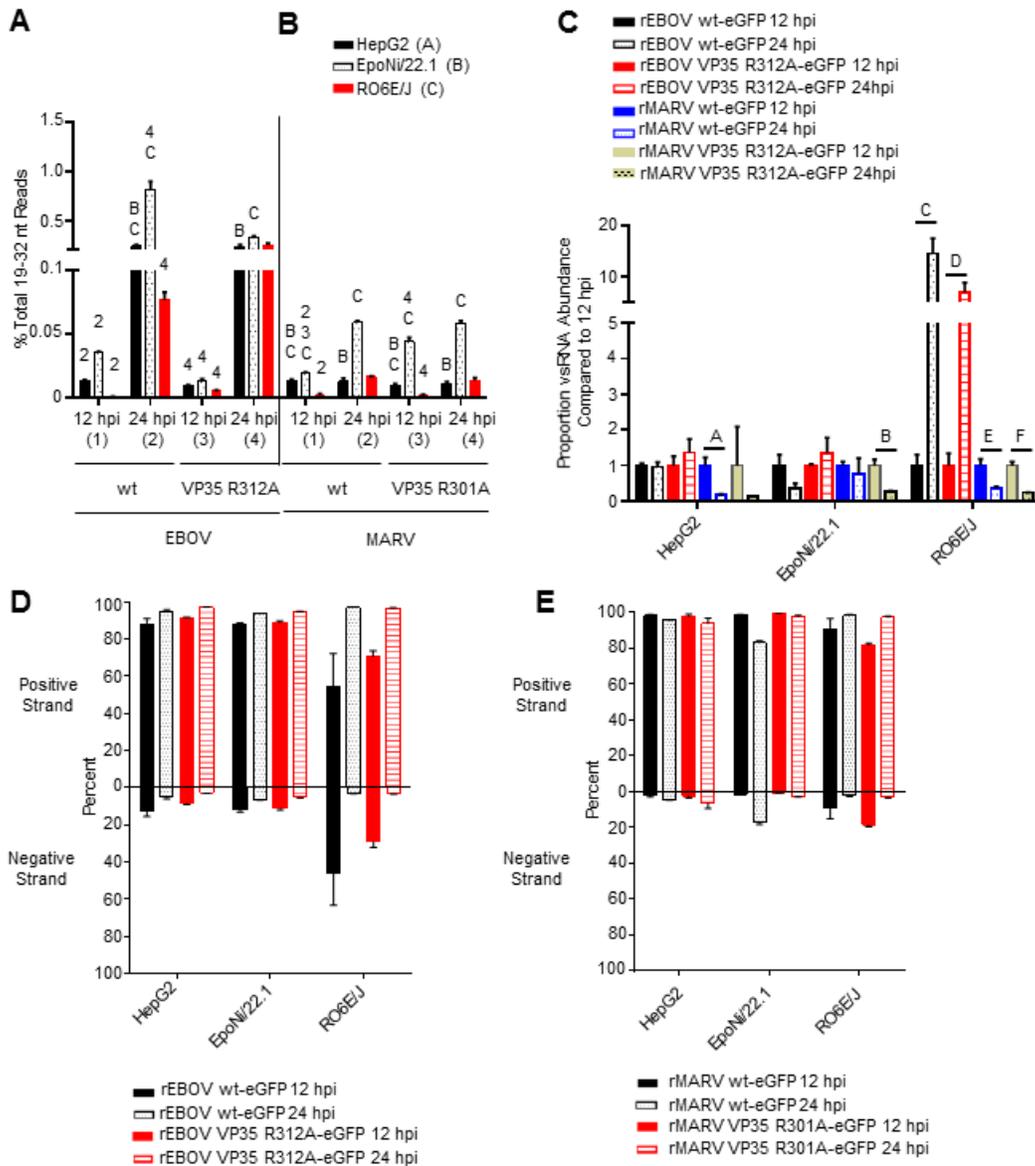


Figure 4: vsRNAs decrease in abundance over time and are predominately derived from positive-strand RNA

vsRNA reads for each virus and cell line expressed as a percentage of the total 19-32 nt RNAs sequenced from each library (**A** and **B**), or normalized to vRNA abundance (**C**). Normalization procedure is described in Materials and Methods under Bioinformatics. (**A**-**B**), For each virus, time points were assigned numbers, and cell lines were assigned letters. Within a cell line, comparisons were made between 12 and 24 h time points for each virus or between viruses within a time point. Comparisons between cell lines were made for each virus at a given time point. Numbers or letters above bars indicate significant differences between either time points (assigned with the corresponding number) or cell lines (assigned with the corresponding letter) as measured by 2-way ANOVA ($\alpha=0.05$), followed by Tukey's multiple comparisons post-test. For a comprehensive list of comparisons and p

values, please refer to **Appendix B**. (C) The data from Figure 2A and 2B were normalized to viral genomic RNA (see Bioinformatics under Materials and Methods) and the ratio of vsRNA reads sequenced at 24 hpi vs. 12 hpi plotted. Statistical comparisons were made between 12 and 24 hour timepoints for each virus within each cell line, and significance was determined by unpaired t-test with Welch's correction (A $p=0.0231$, B $p=0.0161$, C $p=0.0147$, D $p=0.0274$, E $p=0.0221$, F $p=0.0058$). The percentage of vsRNA reads aligning to either the positive or the negative strand of each virus per time point is plotted for EBOV (D) and MARV (E). For all panels, the mean of three biological replicates \pm SD are plotted.

The lack of a more symmetric distribution of reads mapping to both the positive and negative strand, particularly of 19-23 nt reads, led us to hypothesize that the majority of vsRNA reads were being produced from mRNA rather than dsRNA viral replicative intermediates. In all cell lines, a substantial proportion of EBOV and MARV-derived vsRNA reads mapped to the 5'-UTRs of several viral genes (**Fig. 5**) (16, 99, 322, 323).

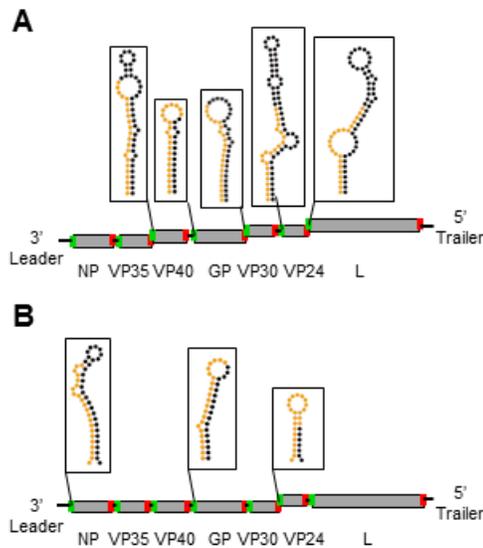


Figure 5: Filovirus vsRNAs are primarily derived from structured elements in the 5'-UTR of viral mRNAs

Schematic representation of EBOV (A) and MARV (B) genome organization, with predicted secondary structures for the 5'-termini of genes from which the highest proportion of sequenced vsRNAs were derived. Nucleotides comprising the 22 nt vncRNA derived from each gene are colored orange, and the rest of the stem-loop structure is colored black. Structures depicted for EBOV and MARV were computed in references 16 and 99, and were drawn here using the tool Forna. Transcription start and stop signals for each gene are depicted in green and red, respectively. Black connections between genes reflect

intergenic regions; overlaps between start and stop signals are staggered. For both viruses, the eGFP transgene is omitted from the schematic.

In EBOV-infected cells, these reads were primarily aligned to the gene start sites (GS) of VP40, eGFP/GP, VP30, VP24, and L (**Figs. 6A, 7A,B**). Since the first 59 nts of both eGFP and GP 5' UTRs are identical in the recombinant EBOV we used, it was not possible to differentiate the proportion of GS-derived reads from each, as the analysis software evenly split the reads between the two sites. Reads aligning to the GS of NP were only rarely observed and reads from the GS of VP35 were almost completely absent among all replicates and time points (**Fig. 6C-E**). For MARV, GS-derived reads were primarily aligned to the NP, VP24, and GP genes (**Figs. 6B, 7C,D**), with VP40 and VP35-derived reads being rarely sequenced, and a complete absence of reads aligning to VP30 and L GS (**Fig. 6F-H**). The eGFP GS-derived read from the recombinant MARV we used is unique, as it extends into the ORF of eGFP, so it was possible to determine that a minority population of reads derived from this GS was present in most libraries. Using previously predicted secondary structures for EBOV and MARV mRNA 5' termini, we found very few reads corresponding to what would be predicted to be the 3p arm of the precursor stem-loop from both EBOV and MARV in all cell lines. These data indicate that these RNA species may either have been size-selected out of the analysis or are subject to rapid degradation.

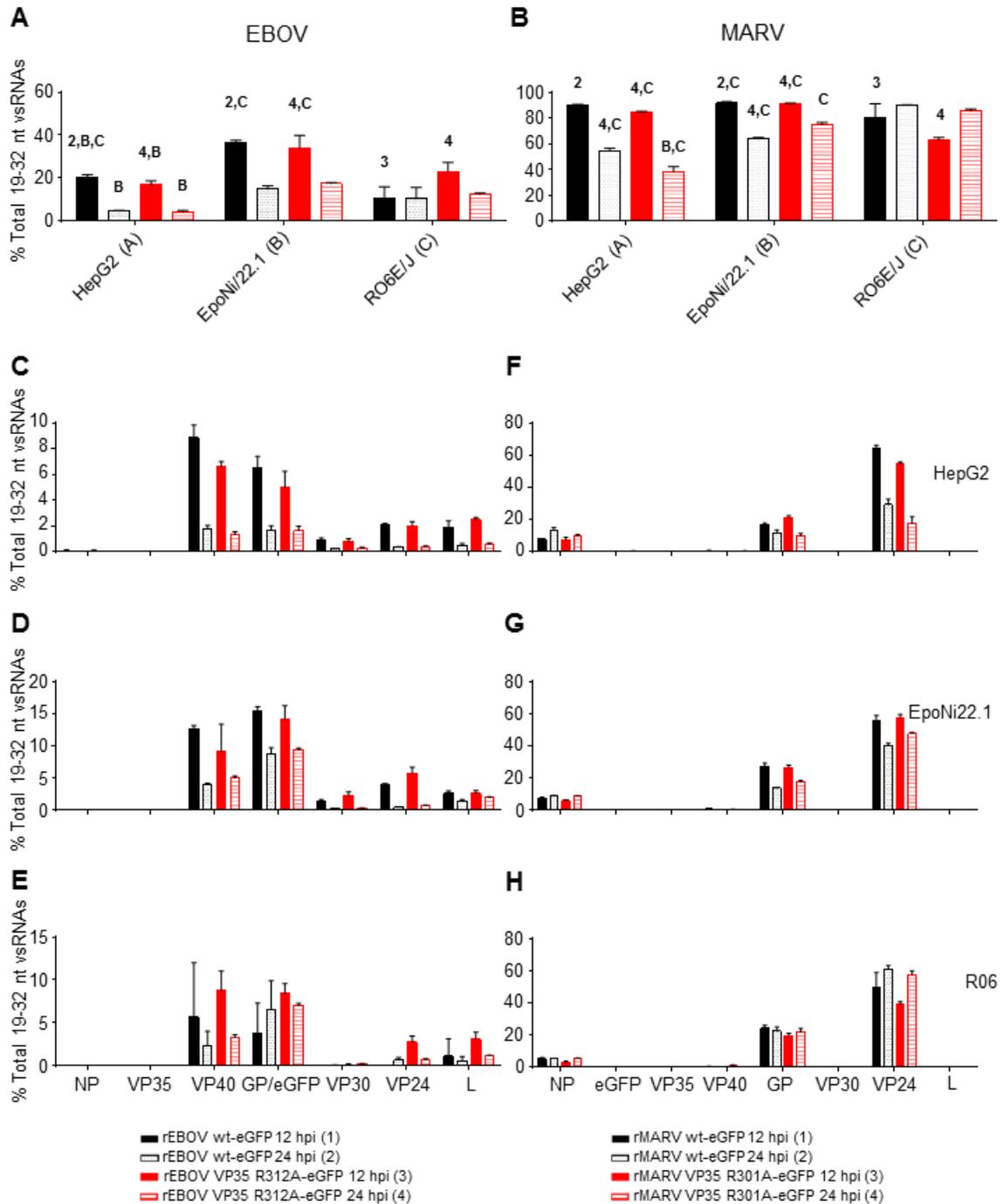


Figure 6: Filovirus vncRNAs are a substantial portion of the total vsRNA population

(A) Proportion of combined EBOV VP40, eGFP/GP, VP30, VP24, and L vncRNAs compared to the total vsRNA population in all cell lines. (B) Proportion of combined MARV NP, GP, and VP24 vncRNAs compared to the total vsRNA population in all cell lines. For each virus, time points were assigned numbers, and cell lines were assigned letters. Within a cell line, comparisons were made between 12 and 24 h time points for

each virus or between viruses within a time point. Comparisons between cell lines were made for each virus at a given time point. Numbers or letters above bars indicate significant differences between either time points (assigned with the corresponding number) or cell lines (assigned with the corresponding letter) as measured by 2-way ANOVA ($\alpha=0.05$), followed by Tukey's multiple comparisons post-test. For a comprehensive list of comparisons and *p* values, please refer to **Appendix B**. **(C-E)** Histograms plotting individual proportions of EBOV vncRNAs compared to the total vsRNA population in HepG2 **(C)**, EpoNi/22.1 **(D)**, and RO6E/J **(E)**. **(F-H)** Histograms plotting individual proportions of MARV vncRNAs compared to the total vsRNA population in HepG2 **(F)**, EpoNi/22.1 **(G)**, and RO6E/J **(H)**. For all panels, the means of three biological replicate sequencing libraries for each group \pm SD are plotted.

Comparisons of the relative abundance of vncRNAs in EpoNi/22.1 and RO6E/J cells versus HepG2 cells were performed by calculating reads per million (RPM) for each vncRNA and normalizing the RPM to vRNA abundance for each library relative to HepG2 cells **(Fig. 8)**. In addition, we quantified the absolute abundance of the EBOV GP vncRNA from each library using RT-qPCR **(Fig. 9)**. For rEBOV-infected EpoNi/22.1 cells, the abundance of each vncRNA was statistically equal to those in HepG2 cells, with the exception of the GP vncRNA derived from the rEBOV wt-eGFP virus **(Figs. 8A-D, 9A-B)**. Strikingly, at 24 hpi for both wt and VP35-mutant rEBOV, RO6E/J cells displayed a substantial increase in the relative abundance of vncRNAs as compared to HepG2 cells **(Figs. 8A-B, 9B)**. A more moderate fold increase in the relative proportion of vncRNAs was observed in rMARV-infected RO6E/J cells at 24 hpi **(Fig 8E-G)**. We next calculated fold change for vRNA-normalized RPM values from each virus from 12 to 24 hpi **(Fig 9. C-D, Fig. 10)**. For rEBOV-infected HepG2 and EpoNi/22.1 cells, individual vncRNAs tended to decrease substantially in abundance between the two timepoints **(Fig 9. C,D, 10A,B)**. Conversely, the relative abundance of vncRNAs increased by 24 hpi in RO6E/J cells infected with rEBOV VP35 R312A-eGFP, congruent with our observation that EBOV vsRNAs overall increase over time in this cell line, though this analysis was not conducted using RPM normalization for RO6E/J cells infected with rEBOV wt-eGFP virus due to the low number of reads from the 12 hpi timepoint.

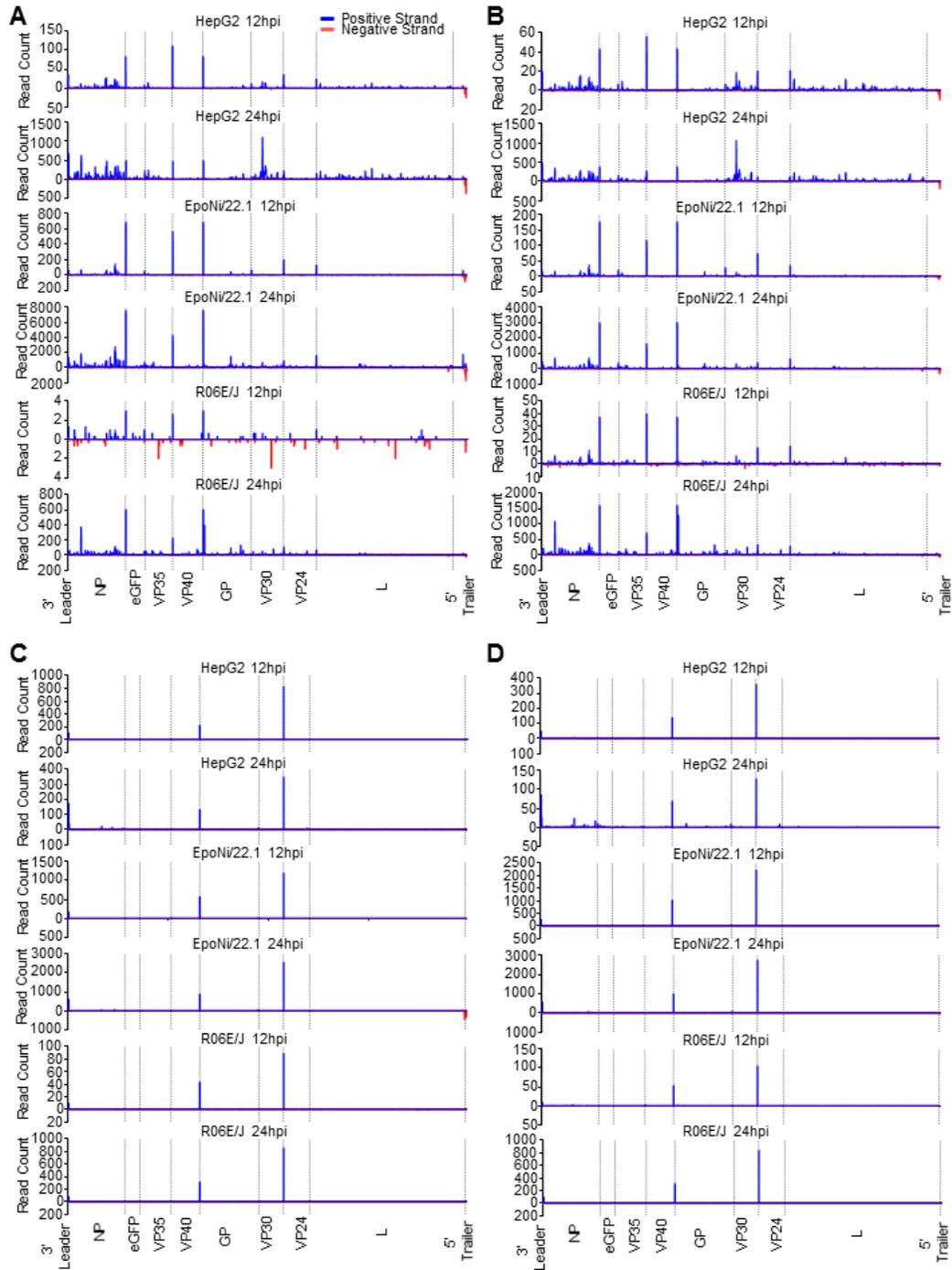


Figure 7: vsRNA read frequency plots for rEBOV-infected and rMARV-infected cells 12 and 24 hpi

The mean number of reads aligning to the virus genome from three biological replicate libraries for each cell line, virus, and timepoint are plotted. Dashed lines indicate transcription start sites, except for the last line, which denotes the border between the 3'-UTR of L and the 5'-trailer. (A) rEBOV wt-eGFP, (B) rEBOV VP35 R312A-eGFP, (C)

MARV wt-eGFP, **(D)** MARV VP35 R301A-eGFP. Note that the y-axis denotes the number of reads aligning at each position; a single read may have multiple alignments.

However, when the vRNA-normalized absolute abundance of the GP vncRNA from rEBOV wt-eGFP-infected cells at 12 hpi was compared to 24 hpi, a difference was seen, though it failed to reach significance due to the presence of an outlier identified by Grubb's test for outliers (**Fig. 9D**, not removed in Figure data). In contrast, rMARV-derived vncRNAs tended to decrease in relative abundance between 12 and 24 hpi in all cell lines, though these changes were not significant in EpoNi/22.1 cells infected with rMARV wt-eGFP, or HepG2 cells infected with rMARV VP35 R301A-eGFP (**Fig. 10C,D**).

Finally, we assessed the degree of agreement between deep-sequencing read normalization using the RPM method vs. absolute quantitation of the GP vncRNA using RT-qPCR and found a significant degree of correlation (**Fig. 9E**). Taken together, these data suggest that the efficiency of vsRNA production is at least in part cell line-dependent, and that disabling of the dsRNA-binding activity of EBOV/MARV VP35 does not dramatically affect the abundance of vsRNAs over time. Additionally, due to the impaired replicative ability of EBOV in RO6E/J cells versus HepG2 and EpoNi/22.1 cells, this analysis implicates enhanced vsRNA production as a potential restriction factor for rEBOV infection. When grouped by size, EBOV vncRNAs were primarily 22 or 25 nts in length, depending on the gene they were derived from across all cell lines (**Fig. 11**). MARV vncRNAs were almost entirely 22 nts in length, with the exception of NP, which included a substantial proportion of larger species (**Fig. 11**).

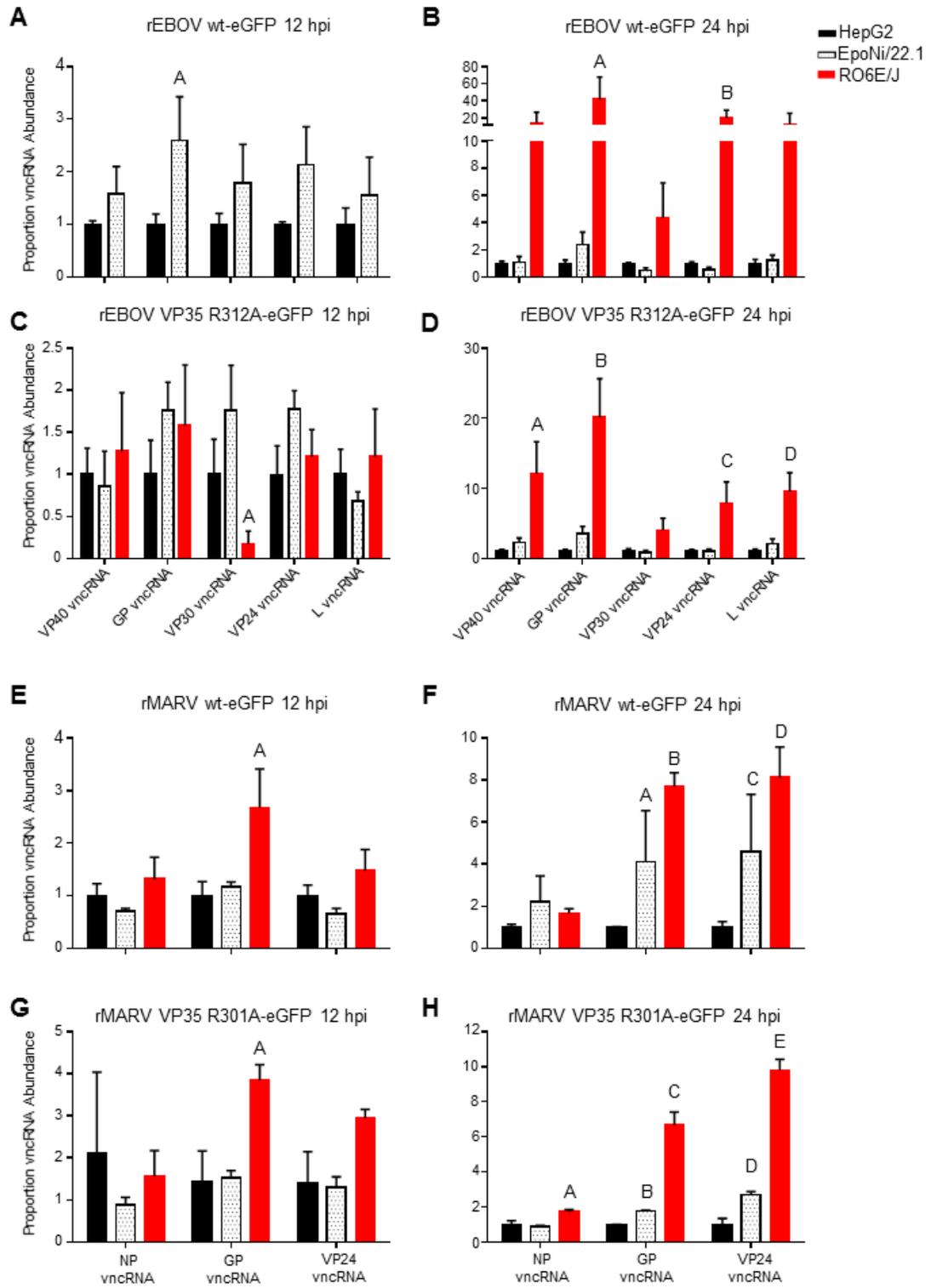


Figure 8: Filovirus vncRNAs exhibit cell-line dependent differential abundance

For each virus and timepoint, the abundance of each vncRNA relative to vRNA abundance from EpoNi/22.1 and RO6E/J libraries was compared to the abundance in HepG2 cells.

Normalization procedure is described in the Bioinformatics subsection of Materials and Methods. All statistical comparisons were made against HepG2 libraries, using 2-way ANOVA followed by Dunnett's multiple comparisons post-test, except in the case of (A), where Sidak's multiple comparisons post-test was used. Letters denote significance as denoted in each panel caption. (A) rEBOV wt-eGFP 12 hpi. A: $p=0.0067$. Note that RO6E/J cells were omitted from this analysis. (B) rEBOV wt-eGFP 24 hpi. A: $p\leq 0.0001$, B: $p=0.0172$. (C) rEBOV VP35 R312A-eGFP 12 hpi. A: $p=0.0448$. (D) rEBOV VP35 R312A-eGFP 24 hpi. A: $p\leq 0.0001$, B: $p\leq 0.0001$, C: $p=0.0012$, D: $p\leq 0.0001$. (E) rMARV wt-eGFP 12 hpi. A: $p\leq 0.0001$. (F) rMARV wt-eGFP 24 hpi. A: $p=0.0245$, B: $p\leq 0.0001$, C: $p=0.0102$, D: $p\leq 0.0001$. (G) rMARV VP35 R301A-eGFP 12 hpi. A: $p=0.0026$. (H) rMARV VP35 R301A-eGFP. A: $p=0.0472$, B: $p=0.0341$, C: $p\leq 0.0001$, D: $p\leq 0.0001$, E: $p\leq 0.0001$. For all panels, the means of three biological replicate sequencing libraries for each group \pm SD are plotted.

We further observed a remarkable abundance of a G-to-U SNP at position one (the 5' terminus) in all EBOV and MARV vncRNA sequences (**Table 1**). In EBOV-infected HepG2 and EpoNi/22.1 cells, the frequency of this SNP ranged 15-20% of all other derivations of the same sequence (including the "wt" sequence and other, more minor SNPs), particularly in the 22 nt size class (data not shown). In RO6E/J cells, the abundance of the 1U SNP was far lower, on par with the abundance of 1U variants in MARV vncRNAs from all cell lines. Another peculiar finding observed in libraries from all cell lines, and for both wt and VP35 mutant viruses was that the vast majority of 26 nt reads corresponding to the EBOV eGFP/GP vncRNA contained a SNP at the 3' terminus (G to U). In some cases, the proportion of this read to the wt sequence exceeded 100:1 (data not shown). We analyzed the individual base quality scores across randomly selected libraries for each of the vncRNA sequences using BamView v.1.2.11, and found the maximum expected error rate at these positions to be 0.1-1%; far below the frequency observed.

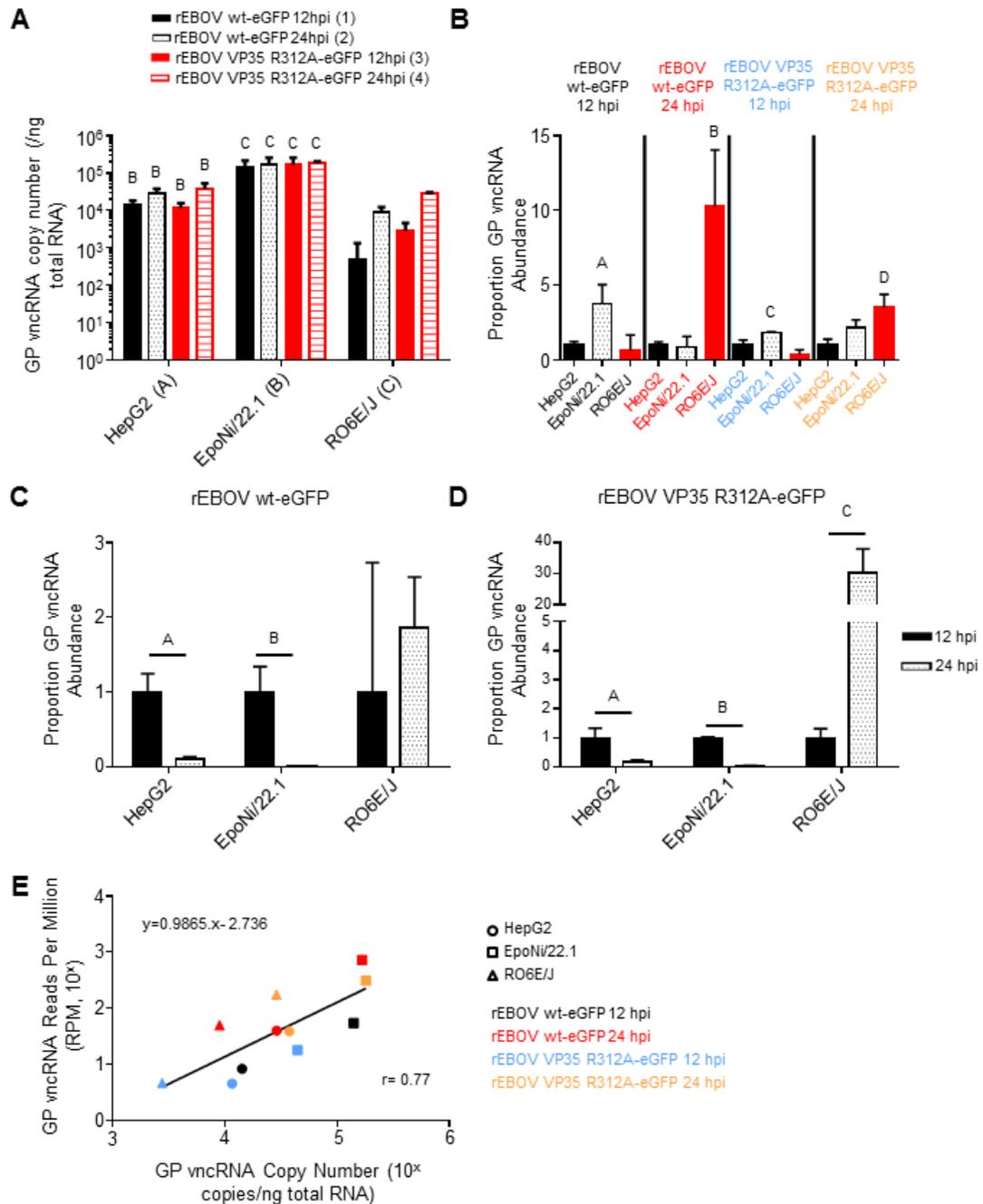


Figure 9: Absolute quantitation of EBOV GP vncRNA by RT-qPCR

Total RNA from the same samples used to generate the rEBOV wt-eGFP and rEBOV VP35 R312A-eGFP small RNA sequencing libraries was used for absolute quantitation of the EBOV GP vncRNA using RT-qPCR. (A) EBOV GP vncRNA copy number per ng of total RNA. For statistical comparison, time points were assigned numbers, and cell lines were assigned letters. Within a cell line, comparisons were made between 12 and 24 h time points for each virus or between viruses within a time point. Comparisons between cell

lines were made for each virus at a given time point. Numbers or letters above bars indicate significant differences between either time points (assigned with the corresponding number) or cell lines (assigned with the corresponding letter) as measured by 2-way ANOVA ($\alpha=0.05$), followed by Tukey's multiple comparisons post-test. For a comprehensive list of comparisons and p values, please refer to **Appendix B (B)** For each virus and timepoint, the abundance of the GP vncRNA relative to vRNA abundance from EpoNi/22.1 and RO6E/J was compared to the abundance in HepG2 cells (set to 1). Normalization procedure is described in Materials and Methods under Bioinformatics. All statistical comparisons were made against HepG2, using 1-way ANOVA followed by Dunnett's multiple comparisons post-test. Specific p values are as follows: A: $p= 0.0230$, B: $p= 0.0039$, C: $p= 0.0133$, D: $p= 0.0052$. For either rEBOV wt-eGFP (**C**), or rEBOV VP35 R312A-eGFP (**D**) viruses, the proportion of the GP-vncRNA relative to vRNA abundance at 24 hpi was compared the abundance at 12 hpi (set to 1). Normalization procedure is described in Materials and Methods under Bioinformatics. Statistical significance was computed using independent unpaired t-tests ($\alpha= 0.05$) with the Holm-Sidak correction for multiple comparisons. For panels (**A-D**), data plotted represents the means of three biological replicates \pm SD. (**E**) The degree of agreement between RPM normalization and absolute quantitation of the GP vncRNA methods was assessed by plotting the mean RPM of three biological replicates from each sample group against the mean GP vncRNA copy number (per ng total RNA) from each sample group. The coefficient of correlation was calculated by the Pearson method ($p= 0.0054$), and a best-fit line drawn using linear regression.

Table 1: GS sequence-derived ncRNA sequences and their features

Virus	Sequence (positive sense)	Gene	hsa-miR homolog	Predicted targets in miRDb
EBOV	<u>GATGAAGATTAAGAAAAACCTA</u>	VP40	hsa-miR-3142	197 for hsa-miR-3142, 673 for VP40
	<u>GATGAAGATTAAGCCGACAGTG</u>	GP	None	632
	<u>GATGAAGATTAATGCGGAGGTC</u>	VP24	hsa-miR-3679-3p	517 for hsa-miR-3679-3p, 632 for VP24
	<u>GATGAAGATTAAGAAAAAGGTA</u>	VP30	None	673
	<u>GAGGAAGATTAAGAAAACTGC</u>	L	None	1150
MARV	<u>GAAGAATATTAACATTGACATT</u>	NP	None	766
	<u>GAAGAACATTAATTGCTGGATG</u>	GP	hsa-miR-424-5p	77 for hsa-miR-424-5p, 917 for GP
	<u>GAAGAACATTAAGAAAAAGGAT</u>	VP24	None	981

We next asked whether EBOV vncRNAs were present in the tissue of animals following infection with EBOV. Total RNA was extracted from archived liver tissue from a rhesus macaque vaccinated against EBOV, that survived lethal challenge, and a control animal that succumbed to infection 8 days post challenge. These RNAs were subjected to miRNA-specific RT-qPCR. The EBOV GP vncRNA was detectable in the liver from the control macaque; however, no detectable GP vncRNA was present in tissue from the vaccinated animal (**Fig. 12A**). Melt curve analysis revealed a single amplification product which was nearly identical when comparing RNA extracted from the liver of the control animal and RNA extracted from EpoNi/22.1 cells 24 hpi with rEBOV wt-eGFP (**Fig. 12B**). The control animal exhibited high circulating viremia upon euthanasia, while the vaccinated animal had no detectable viremia when euthanized 28 days post-challenge. This demonstrates that EBOV vncRNAs are produced *in vivo* during the course of infection.

EBOV vncRNA production does not require host miRNA machinery

We next asked whether biogenesis of the EBOV vncRNAs we identified was dependent upon processing by miRNA pathway-associated endoribonucleases (i.e. Dicer and Drosha). We infected a Dicer-knockout cell line (NoDice 4-25) and the Dicer-competent parental line (293T-P) (324) with rEBOV wt-eGFP, and probed for the presence of EBOV vncRNAs by miRNA-specific RT-qPCR. At 20 hpi, there was approximately a 4- and 10-fold increase in the production of the VP40 and GP vncRNAs, respectively, in 293T-P versus NoDice 4-25 cells (**Fig. 13A**). The L vncRNA was omitted from this analysis because it could not reliably be detected in all biological and technical replicates in NoDice 4-25 cells.

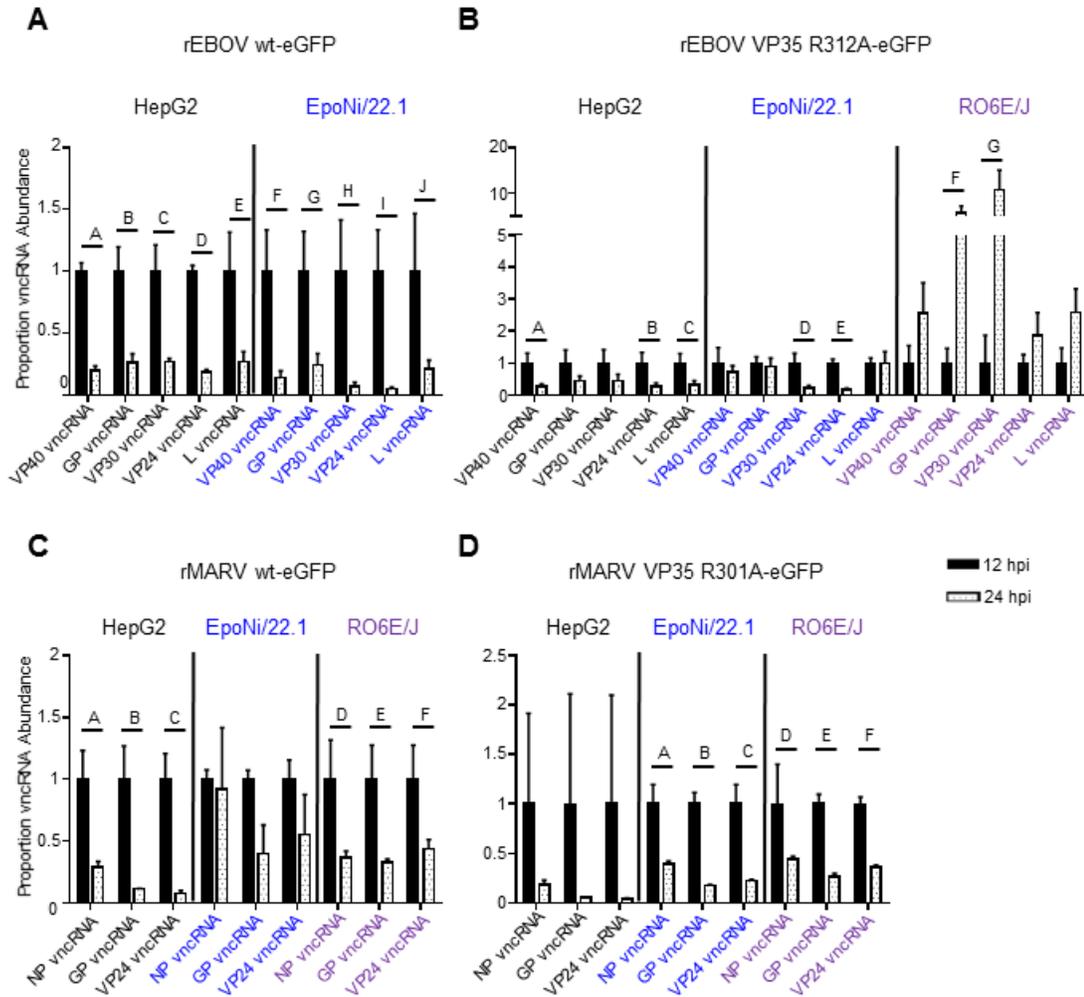


Figure 10: Filovirus vncRNA abundance changes over time

For each virus, the proportion of each vncRNA relative to vRNA abundance at 24 hpi was compared to the abundance at 12 hpi (set to 1). Normalization procedure is described in the Bioinformatics subsection of Materials and Methods. All statistical comparisons were made using 2-way ANOVA followed by Sidak's multiple comparisons post-test. Letters denote significance as denoted in each panel caption. **(A)** rEBOV wt-eGFP. A: $p \leq 0.0001$, B: $p \leq 0.0001$, C: $p \leq 0.0001$, D: $p \leq 0.0001$, E: $p \leq 0.0001$, F: $p = 0.0044$, G: $p = 0.0134$, H: $p = 0.0022$, I: $p = 0.0017$, J: $p = 0.0098$. Note that RO6E/J cells were omitted from this analysis. **(B)** rEBOV VP35 R312A-eGFP. A: $p = 0.0175$, B: $p = 0.0205$, C: $p = 0.0322$, D: $p = 0.0065$, E: $p = 0.0037$, F: $p = -0.0079$, G: $p \leq 0.0001$. **(C)** rMARV wt-eGFP. A: $p = 0.0007$, B: $p \leq 0.0001$, C: $p \leq 0.0001$, D: $p = 0.0091$, E: $p = 0.0056$, F: $p = 0.0182$. **(D)** rMARV VP35 R312A-eGFP. A: $p = 0.0002$, B: $p \leq 0.0001$, C: $p \leq 0.0001$, D: $p = 0.0054$, E: $p = 0.0007$, F: $p = 0.0020$. For all panels, the means of three biological replicate sequencing libraries for each group \pm SD are plotted.

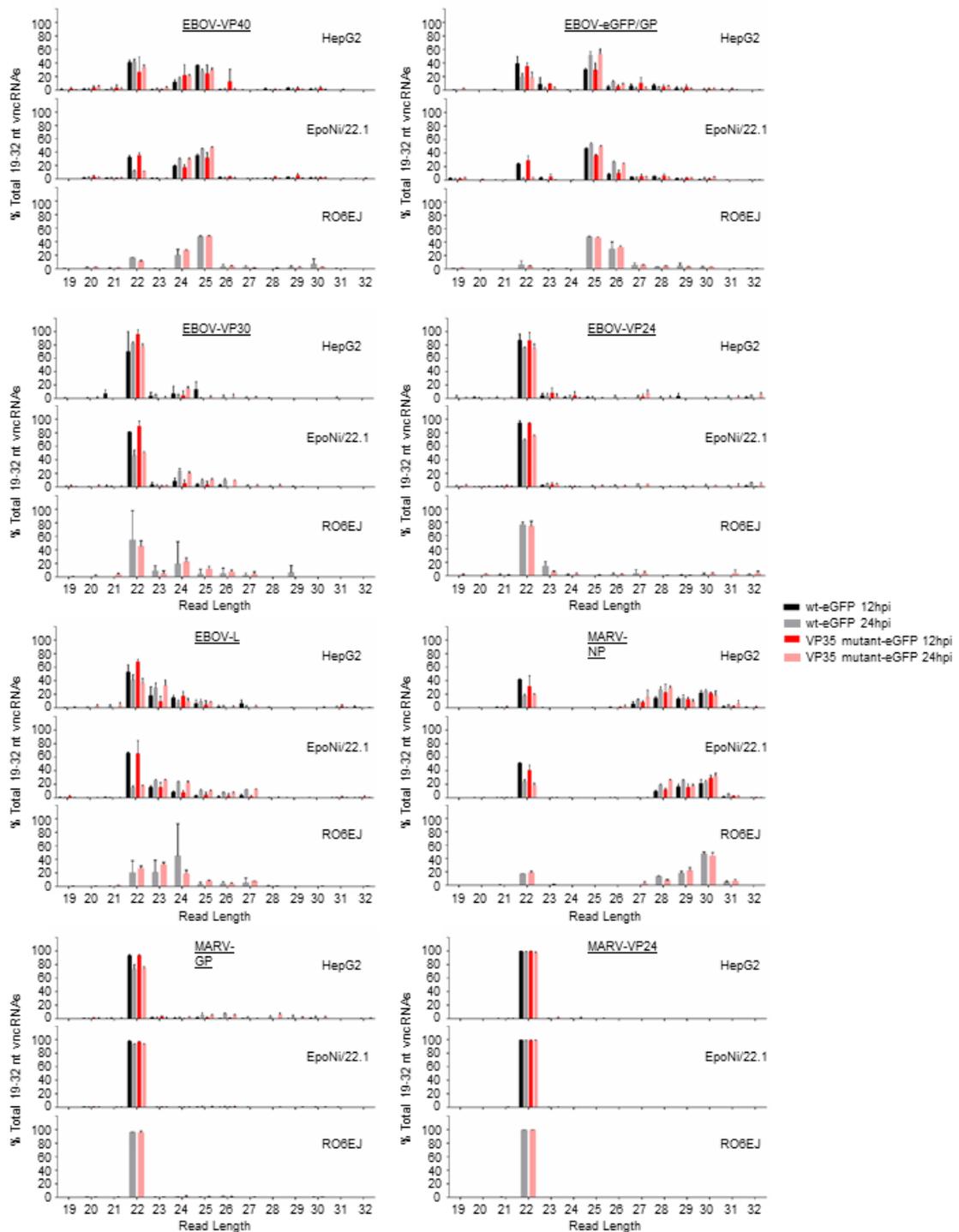


Figure 11: Length distribution profiles of EBOV and MARV vncRNAs

vncRNA sequences from 19-32 nts, including reads with SNPs, were plotted as a percentage of the total number of reads derived from each TSS. The mean of three biological replicate sequencing libraries for each group is plotted \pm SD.

The apparent difference in relative abundance of viral vncRNAs between cell lines was negligible when accounting for viral load as measured by RT-qPCR detection of viral genomic RNA (vRNA) and mRNA, which were approximately 6- and 10-fold respectively, in 293T-P versus NoDice 4-25 (**Fig. 13B**). By contrast, in NoDice 4-25 cells production of the endogenous mature miRNAs hsa-miR103a-3p and hsa-let-7a-5p was impaired by approximately 2,000 and 200-fold, respectively, compared to the parental cell line. However, the former may be impacted by factors other than the lack of Dicer (279). There was no difference in the relative abundance of the endogenous U6 snRNA, which is not dependent on Dicer for its biogenesis, between cell lines.

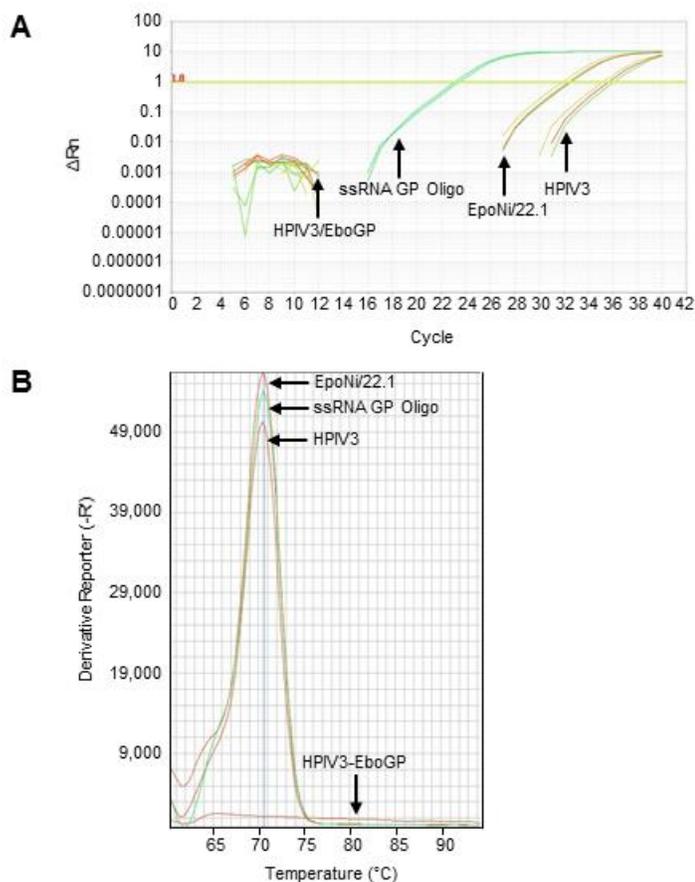


Figure 12: miRNA RT-qPCR detection of EBOV GP vncRNAs in NHP liver tissue

Total RNA was extracted from archived liver tissue from EBOV-vaccinated and unvaccinated rhesus macaques and subjected to miRNA-specific RT-qPCR. RT-qPCR reactions for unknowns were performed in triplicate; RT-qPCR reactions for serially diluted standards were performed in duplicates. Total RNA extracted from EpoNi/22.1

cells 24 hpi with rEBOV wt-eGFP virus was used as a positive control. A synthetic single-stranded RNA oligo homologous to the GP vncRNA was spiked into total human brain RNA at a concentration of 0.1 ng/uL (total human brain RNA concentration: 10 ng/ μ L) and used as a standard for absolute quantitation (undiluted – 10^{-7}). **(A)** Amplification curves: human parainfluenza virus type 3 (HPIV3) empty vector-vaccinated control rhesus macaque (moribund animal euthanized 8 days post-challenge, extrapolated GP vncRNA copy number = 3.06×10^3 copies/ng total RNA); HPIV3/EboGP-vaccinated rhesus macaque (surviving animal euthanized 28 days post-challenge at study endpoint, extrapolated GP vncRNA copy number = 0 copies/ng total RNA); EpoNi/22.1 rEBOV wt-eGFP 24 hpi (extrapolated GP vncRNA copy number = 2.78×10^4 copies/ng total RNA); EBOV GP vncRNA standard (10^{-3} dilution; 8.4×10^6 copies/ng total RNA). **(B)** Melt curve analysis of GP vncRNA RT-qPCR amplicons. For clarity, the melt curve for only a single technical replicate is shown for each sample, data shown for each sample is representative of all technical replicates for that sample. The average amplicon T_m of three technical replicates for each sample (two for the standard) was: HPIV3-vaccinated control rhesus macaque, $T_m = 70.45$ °C; HPIV3-EboGP-vaccinated rhesus macaque, $T_m = 65.19$ °C; EpoNi/22.1 total RNA, rEBOV wt-eGFP 24 hpi, $T_m = 70.50$ °C; EBOV GP vncRNA standard (10^{-3} dilution; $\sim 8.4 \times 10^5$ copies/ng total RNA), $T_m = 70.54$ °C.

We next used siRNA-mediated knock-down of a panel of host proteins with ribonuclease activity, including RNase L, DIS3, CPSF3L, Drosha and AGO, to determine their effect, if any, on production of EBOV vncRNAs. The degree of knock-down was verified by western blot, or in the case of AGO2 and Drosha (which we were not able to reliably detect by western blot), RT-qPCR detection of miRNA products and mRNA transcripts (**Fig. 13D, 14C**, respectively). Cells were infected with rEBOV wt virus 48 h post-transfection (hpt), and harvested 24 hpi (72 hpt). Gene-specific knockdown of all nucleases in the panel except for CPSF3L resulted in significantly higher levels of EBOV vRNA and mRNA compared to the scrambled control (**Fig. 13C**). For each siRNA-treated biological replicate, including the scrambled siRNA, fold abundance of the GP vncRNA relative to mock transfected cells was normalized separately to relative fold abundance of EBOV vRNA and mRNA-to-Generate a vncRNA:vRNA/mRNA ratio. For cells normalized to viral mRNA, only CPSF3L-siRNA transfected cells displayed a significant reduction of the GP vncRNA:mRNA as compared to the scrambled control (**Fig. 13D**). When normalized to vRNA, knockdown of RNase L showed a significant increase in the relative abundance of the GP vncRNA, likely explained by the importance of RNase L as

an antiviral effector. While the fold abundance of the GP vncRNA remained decreased in CPSF3L siRNA-treated cells, it was not significant when normalizing to vRNA. Relative abundance of hsa-miR103a-3p, which is not dependent on CPSF3L for processing, export, or maturation, remained unchanged. When standardizing the fold change for each gene-specific siRNA knockdown to the scrambled control, the ratios of fold change in vncRNA to fold change in mRNA or vRNA were effectively 1:1 (**Fig. 13E-F**), supporting our hypothesis that viral mRNA is the biogenic substrate for EBOV vncRNAs. In contrast, no such association was observed between fold change in abundance of hsa-miR-103a-3p and either viral mRNA or vRNA (data not shown). Taken together, these data suggest that EBOV vncRNAs are produced in a Dicer-independent manner, but still involve interaction with host nucleases.

EBOV vncRNAs are not associated with Argonaute/RISC, and do not have an effect on virus replication

A critical requirement of miRNA functionality is the ability to be loaded and selected for in RISC. To determine whether EBOV vncRNAs were associated with host Argonaute proteins, we performed RNA-binding protein immunoprecipitation (RIP) on rEBOV wt-eGFP infected 293T cells 20 hpi, using antibodies individually targeting human AGO 1-4, as well a pan-AGO antibody. Following isolation of Ago-associated RNAs, we probed for the presence of four of the EBOV vncRNAs (VP40, GP, VP24, L) using miRNA-specific RT-qPCR. Out of six independent experimental repetitions, EBOV GP vncRNA was enriched in two experiments, and EBOV L vncRNA was enriched in one (data not shown). Subsequent experiments failed to enrich for any of the four EBOV vncRNAs profiled (**Fig. 15A**, only data for EBOV GP vncRNA shown). Endogenous miRNAs (hsa-let-7a-5p and/or hsa-miR-103a-3p) were used as positive controls and were highly enriched for AGO2 and pan-AGO precipitations, but were not associated with AGO1, AGO3 or AGO4.

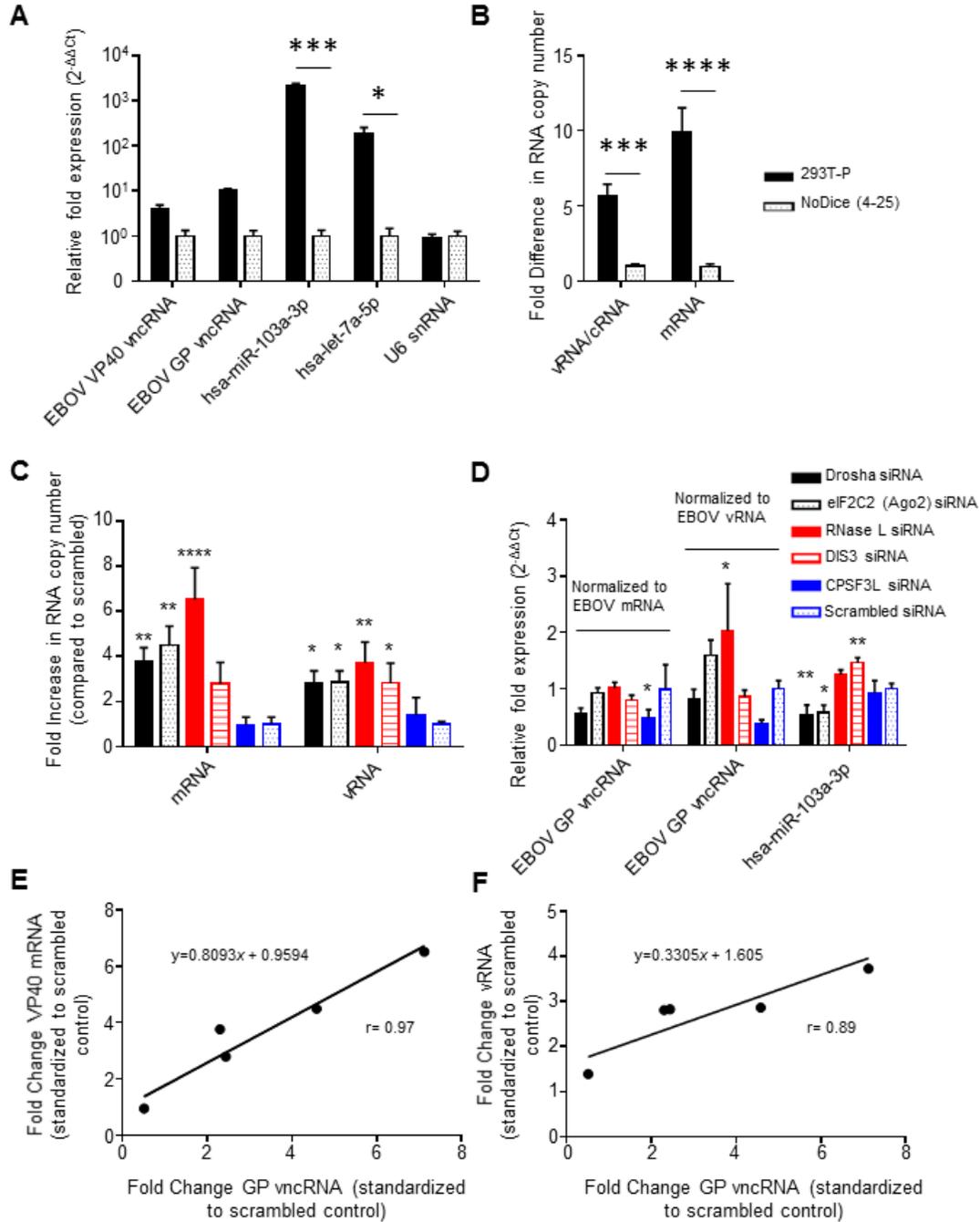


Figure 13: EBOV vncRNAs are produced in a Dicer-independent manner

293T-P or Dicer-null NoDice 4-25 cells were infected at an MOI 2 PFU/cell with rEBOV wt virus and lysed in Trizol 20 hpi. (A) miRNA-specific RT-qPCR analysis of the VP40 and GP vncRNAs as well as the endogenous cellular miRNAs hsa-miR-103a-3p and hsa-let-7a-5p. The Dicer-independent U6 snRNA was assayed as a negative control. Relative fold change was performed using the $2^{-\Delta\Delta C_t}$ method of approximation, with an exogenous

spike-in RNA used as the reference. Asterisks denote significance as determined by 2-way ANOVA ($\alpha=0.05$) followed by Sidak's multiple comparisons post-test. * $p<0.05$, *** $p<0.001$. **(B)** Tag-based strand-specific RT-qPCR was performed on the same samples to determine viral RNA copy number. RNA equivalents for vRNA and mRNA were determined using a standard curve obtained by serial dilutions of strand-specific standards. Data is expressed as fold change compared to NoDice 4-25 cells. Asterisk denotes significance ($p<0.05$) compared to NoDice 4-25 cells as measured 2-way ANOVA ($\alpha=0.05$) followed by Sidak's multiple comparisons test. *** $p=0.0005$, **** $p<0.0001$. **(C)** Tag-based strand-specific RT-qPCR was performed on total RNA from Droscha, Ago2, RNase L, DIS3, CPSF3L, and scrambled siRNA-transfected cells, as well as mock transfected cells to determine viral RNA copy numbers. Data was first normalized to mock transfected cells, and is expressed as fold change compared to the scrambled siRNA control. Asterisks denote significance ($p<0.05$) as measured by one-way ANOVA ($\alpha=0.05$, calculated separately for mRNA and vRNA) followed by Dunnett's multiple comparisons test, with all comparisons made to the scrambled control. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$. **(D)** miRNA-specific RT-qPCR of EBOV GP vncRNA. The relative fold-change (as measured by the $2^{-\Delta\Delta C_t}$ method of approximation) was first normalized to the mock for all samples, and then to either mock-normalized viral mRNA fold change values or mock-normalized vRNA fold change values, as indicated. Fold change values of the endogenous hsa-miR-103a-3p miRNA were not normalized to viral titer as we assumed its levels were not directly affected by viral RNA levels. Asterisks denote significance ($p<0.05$) as determined by one-way ANOVA ($\alpha=0.05$, calculated separately for mRNA-normalized GP vncRNA, vRNA-normalized GP vncRNA, and hsa-miR-103a-3p) followed by Dunnett's multiple comparisons test, with all comparisons made to the scrambled control. * $p<0.05$, ** $p<0.01$. For panels **A-D**, the means of three biological replicates \pm SD are plotted. The degree of correlation between the GP vncRNA and either viral mRNA **(E)** or vRNA **(F)** was plotted by taking the mean fold change of the GP vncRNA for each siRNA transfection normalized to the scrambled siRNA and plotting it against corresponding values of the viral RNA. Deming regression was performed to fit a line through the data points, and the coefficient of correlation (r) was calculated using the Pearson method.

To further investigate the potential for EBOV vncRNAs to act as either proviral or antiviral elements, we identified two potential binding sites, one each in the 5'-UTRs of GP and VP30 mRNAs, for the VP40/GP/VP30/VP24-derived vncRNAs, and single binding site in the ORF of GP and the 5'-UTR of L mRNAs for the L-derived vncRNA, and transfected 293T-P cells with synthetic siRNA duplexes with the guide strands designed to be homologous to the VP40 and L vncRNAs, followed by infection with rEBOV-wt virus 4 hpt. At 48 hpi, we observed no significant difference in the virus titers in the supernatants of cells treated with VP40, L, or combined siRNAs compared to the scrambled negative control siRNA or mock transfected cells at either 25 or 50 nM

concentrations (**Fig. 15B**). Taken together, these experiments indicate EBOV vncRNAs are not associated with host miRNA machinery, and do not positively or negatively affect virus replication. Moreover, we demonstrate that EBOV effectively resists suppression by siRNAs complementary to putative binding sites in several mRNAs.

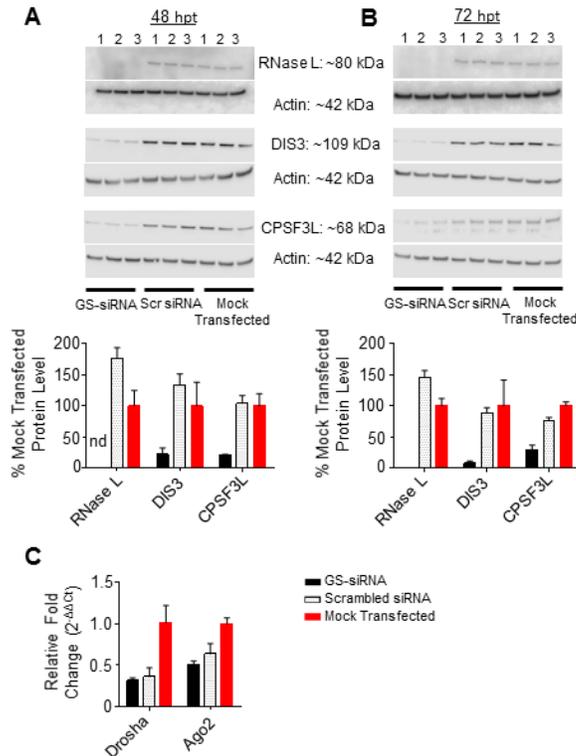


Figure 14: Validation of siRNA knockdown by western blot and RT-qPCR

A, B top: Western blots and band pixel density analysis of RNase L, DIS3, and CPSF3L in siRNA transfected cells 48 hours (**A**) and 72 hours (**B**) post transfection. Bottom: quantitative data on the levels of siRNA derived from western blots and normalized to mock-transfected cells. For both (**A**) and (**B**), siRNA knockdowns were performed in triplicate (denoted as the numbers above the blot images). **C**) RT-qPCR of Drosha and Ago2 mRNA transcripts. For all panels, bars represent the mean of three biological replicates \pm SD. GS = gene specific siRNA, Scr = scrambled (non-specific) siRNA.

EBOV vncRNAs lack the ability to effectively suppress a reporter transcript with multiple miRNA-binding elements

We next asked whether EBOV vncRNAs could suppress the expression of a luciferase reporter with miRNA-binding elements (MBE) in the 3'-UTR of the reporter. We designed three separate synthetic 3'-UTRs containing EBOV GP vncRNA MBEs and

cloned them into the pMirGLO Dual-Luciferase vector (Promega). MBEs were designed to be either perfectly complementary to the EBOV GP vncRNA, partially complementary (containing mismatched nucleotides at positions 11-13 of the MBE, “imperfect”), or complementary to only the first 10 nts of the EBOV GP vncRNA.

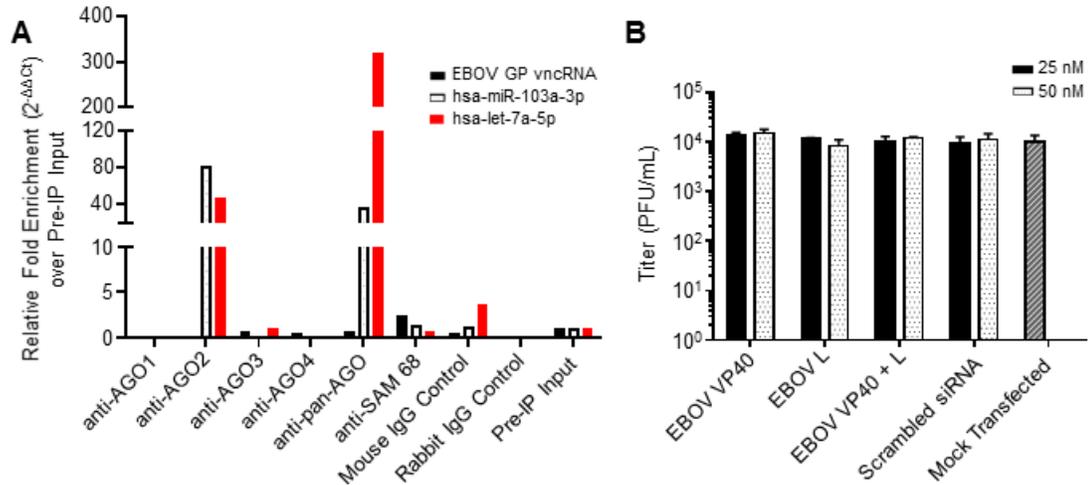


Figure 15: EBOV vncRNAs are not associated with RISC, and do not positively or negatively affect virus replication

(A) RT-qPCR analysis of EBOV vncRNAs following RNA-binding protein immunoprecipitation (RIP) of AGO 1-4 with individual AGO and pan-AGO antibodies. Data shown is the combined data of two independent experiments, data for pan-AGO RIP is representative of 5 separate experiments. SAM-68 is used as an irrelevant RNA-binding protein control. Of the four EBOV vncRNAs profiled, only data for the EBOV GP vncRNA is shown. (B) EBOV titers in supernatants of 293T-P cells 48 hpi. Prior to infection, cells were transfected with the indicated total concentration of siRNAs homologous to the VP40 and L vncRNAs. siRNAs were transfected individually, or combined (combined concentrations 12.5 nM each or 25 nM each). Scrambled control siRNA or mock transfected cells were used as controls for comparison. Significance was tested by 2-way ANOVA ($\alpha=0.05$) followed by Dunnett’s multiple comparisons post-test, with separate independent comparisons being made to the mock infected and scrambled siRNA-transfected control. For panel (B), the means of three biological replicates \pm SD are plotted.

Following transfection with each of the plasmids, 769-P (human renal adenocarcinoma) and EpoNi/22.1 cells were infected with rEBOV wt at MOI 2 PFU/cell. Cell lysates were collected 24 hpi, and assessed for firefly luciferase (FLuc) expression, with normalization to Renilla luciferase. In rEBOV-infected 769-P cells, FLuc expression was significantly reduced in cells transfected with either perfect or the 10 nt MBE

constructs compared to empty vector transfected cells; however, cells transfected with the perfect MBE construct and subsequently mock infected showed similar reduction in FLuc signal (**Fig. 16A**). In rEBOV-infected EpoNi/22.1 cells, FLuc expression was significantly increased in cells transfected with the perfect MBE construct in comparison to empty vector transfected cells (**Fig. 16B**). As a control experiment, we transfected a synthetic siRNA homologous to the EBOV GP vncRNA 4 h post transfection with each of the reporter plasmids. In both 769-P and EpoNi/22.1 cells, transfection of the GP vncRNA reduced expression of the FLuc reporter compared to cells transfected with a scrambled siRNA, except in EpoNi/22.1 cells transfected with the 10 nt MBE construct (**Fig. 16C,D**). These data suggest that EBOV vncRNAs do not demonstrate silencing activity in either bat or human cells.

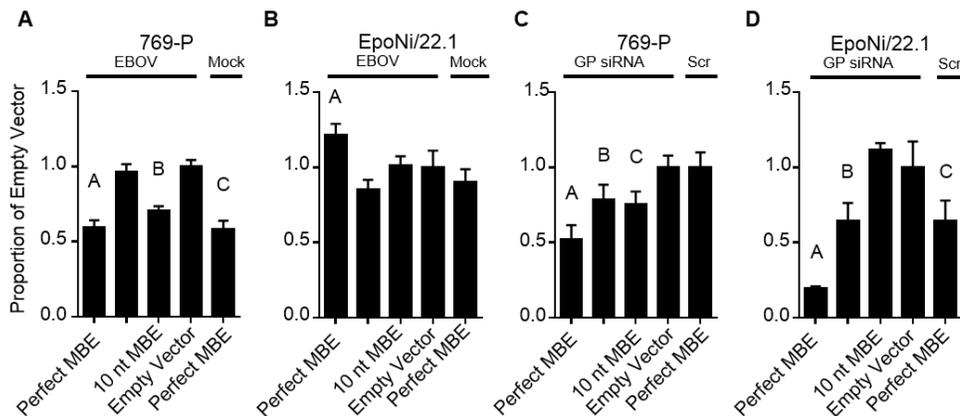


Figure 16: Dual-luciferase reporter assay for EBOV vncRNA function

769-P (**A,C**) or EpoNi/22.1 (**B, D**) cells were transfected with the indicated pMirGLO Dual-Luciferase reporter constructs and either infected with rEBOV wt (**A-B**) or subsequently transfected with a synthetic siRNA homologous to the EBOV GP vncRNA at 50 nM final concentration (**C-D**). For each assay, FLuc values were normalized to corresponding RLuc values to generate a ratio; the data is expressed as the mean FLuc/RLuc ratio \pm SD proportional to the empty vector control construct. All assays were performed in quadruplicate biological replicates (n=4). One-way ANOVA followed by Dunnett's multiple comparisons post-test was used to assess statistical significance, with all comparisons made to the empty vector transfected group. Multiplicity-adjusted p-values are reported. (**A**) A: p<0.0001, B: p<0.0001, C: p<0.0001, (**B**) A: p=0.0065, (**C**) A: p<0.0001, B: p=0.0145, C: p=0.0064, (**D**) A: p<0.0001, B: p=0.0017, C: 0.0018. Representative data of two independent experiments.

DISCUSSION

This work addresses two controversial issues within RNA virology, namely mammalian antiviral RNAi, and RNA virus miRNAs. Although it is possible to engineer RNA viruses to produce functional miRNAs (325-328), there has been relatively little published on the subject of naturally produced RNA virus miRNAs. What has been published, primarily involving flaviviruses, has been exceptionally controversial in the field, and has received a great degree of critical attention (273, 280). Despite this intense opposition to the notion of RNA virus miRNAs, multiple *in silico* modeling efforts have identified potential pre-miRNAs and mature miRNAs as products of the genomes of a number of RNA viruses (329). Most notably for the purposes of this work, numerous relatively recent publications have identified EBOV in particular as a candidate for production of genuine viral miRNAs (241, 242, 244, 246). In the area of antiviral RNAi, a similar division in the field separates those who assert that mammals are capable of using some form of antiviral RNAi from those who point to voluminous evidence of its absence, including the aforementioned differences in the physiology of the RNAi pathways in invertebrates and mammals (289, 300-302). Again, however, recent work, this time somewhat less controversial, has demonstrated that mammals do indeed retain some ability to utilize siRNAs in an antiviral role (303-306). Given the unique physiology and immunology of bats (330), we asked whether either of these phenomena have any role in the natural history of filovirus infection of bats. First, we asked if bats utilize antiviral RNAi to control EBOV infection. Second, we asked if EBOV produces genuine and functional viral miRNAs via a canonical or non-canonical pathway.

Next-generation sequencing has opened new horizons in RNA biology, evolutionary biology, and virology. In particular, specialized sequencing technologies like small RNA sequencing have made it possible to explore entirely new aspects of the interactions of viruses with their hosts. In this study, we applied next-generation

sequencing to answer basic questions about the nature of EBOV infection of bats and to make comparisons to human infection. We infected bat and human cells with EBOV and MARV, including mutants that impair the IFN antagonism of these viruses. The classic signatures of an antiviral siRNA response (331), such as equal strand distribution were not detected. In fact, vsRNAs were overwhelmingly derived from the positive strand. However, this is not unsurprising for a number of reasons. Despite the fact that mammalian antiviral RNAi has been reported recently, the findings of the most influential study indicate that the reason this process has not been observed previously largely come down to masking/suppression of antiviral RNAi by VSR activity, as the effect was only observed when the VSR of the virus used was disabled (306). Given that EBOV possesses multiple VSRS of apparently varying potency, it is quite possible, even likely, that any antiviral RNAi activity would be quickly suppressed once expression of these proteins reaches a sufficient level. Furthermore, the viruses used to assess the potential for mammalian antiviral RNAi activity were (+)ssRNA viruses, which produce long and unshielded dsRNA replicative intermediates (306). (-)ssRNA viruses like EBOV produce very short dsRNA intermediates that are well-shielded from cellular dsRNA sensors, making access to these intermediates very difficult (36).

Despite the fact that we did not identify any signatures suggestive of an siRNA response, we did identify a large number of vsRNAs. The overwhelming majority of these were a group of vncRNAs derived from the GS sequences of multiple viral genes. In EBOV, these were VP40, GP, VP30, VP24, and L. These RNAs varied in length depending on the gene, but generally fell within a range of 22 to 25 nt. MARV produced vncRNAs from NP, GP, and VP24, which were nearly uniform in length, at 22nt. This narrow size range matches that of cellular miRNAs. Many of these vncRNAs have homologs in human miRNAs (**Fig. 6A,B, Table 1**). There was a great deal of consistency in the size of the vncRNAs between cell lines, which would seem to indicate that some non-stochastic mechanism is responsible for their biogenesis. The positive strand bias we observed is also

important, as it effectively eliminates the possibility that these vncRNAs are the product of digestion of dsRNA replicative intermediates, meaning that they must be the product of digestion of the positive sense replicative strand (cRNA), or viral mRNAs. EBOV cRNAs are fully encapsidated, shielding them from nucleolytic digestion. This encapsidation process occurs simultaneously with production of cRNAs, further limiting exposure (27, 29). EBOV mRNAs, which are known to contain strong secondary structure including pre-miRNA-like hairpins, are not shielded, which would leave them far more vulnerable to digestion by Dicer or other nucleases (16).

Analysis of the vncRNAs revealed a curious finding of unclear significance. An unusually large fraction of detected EBOV (and to a lesser degree MARV) vncRNAs incorporated a G-to-U SNP at position 1. Although there is no clear mechanism for this phenomenon, the profound degree to which filovirus GS sequences are conserved makes a genomic origin extremely unlikely, as such a mutation would likely have profound deleterious effects on viral transcription (16) (32). Misincorporation during mRNA transcription is a far more plausible mechanism, perhaps as a result of template misalignment or some similar process (332). We were unable to distinguish between the wt and SNP vncRNA variants via RT-qPCR, largely due to technical limitations. Consultation with Qiagen/Exiqon technical support led us to believe that it was unlikely that this problem was resolvable. As such we were unable to explore this finding further. An additional G-to-U SNP was identified in the 26 nt variant of the EBOV eGFP/GP vncRNA, comprising the overwhelming majority of reads. Although this may appear to be a PCR-biased artifact, it was consistently identified as the majority sequence in all EBOV libraries from all cell lines. As such, it is likely that this is better explained as another example of transcriptional misincorporation.

In order for miRNAs and siRNAs to function, they must be incorporated into RISC (260). This, therefore, is a defining characteristic of this class of small RNAs. Although recent work has found that a number of previously *in silico* predicted EBOV

“miRNAs” are indeed present in the serum of infected humans, nonhuman primates, and mice, no effort was made in this study or indeed any previous study to demonstrate that the vncRNAs in question are in fact present in RISC (333). One report did imply that two of the predicted miRNAs were Dicer-dependent, but this was also not thoroughly validated (241). This lack of rigorous validation of the biogenic origin and function of so-called “viral miRNAs” has left a significant gap in knowledge. To address this, it was important to investigate both the biogenic origin and potential function(s) of the vncRNAs we detected, particularly with regards to the involvement of host miRNA pathway components. A stable Dicer knockout cell line was used to demonstrate that production of EBOV vncRNAs is Dicer-independent (**Fig 13A,B**). However, pre-miRNA processing independent of Dicer, but dependent upon Argonaute 2 (AGO2) is known, and has been reported as a mechanism for the production of a viral miRNA-like molecule in H5N1 IAV (264, 334, 335). In addition to this, some other nucleolytic enzymes have been implicated in the production of miRNA-like molecules during viral infections. According to some reports, RNase L is involved in atypical processing of viral RNAs into siRNA-like small RNAs (336). The endosome is an interesting contender, as its 3'-5' exoribonuclease activity could produce the observed vncRNAs if the complex stalled at the hairpin structure formed by the GS sequence at the 5' end of a mature EBOV mRNA. The catalytic component of the endosome, DIS3, is largely nuclear, but appears to be present in the cytoplasm to some degree according to multiple reports (337). Drosha and the integrator complex (of which CPSF3L is a component) have been shown to process viral RNAs into small RNAs, some of which would be expected to resemble miRNAs (338, 339). Drosha processing likely occurs after translocation from the nucleus, an event known to occur in response to infection with a number of RNA viruses, to include the mononegavirus vesicular stomatitis virus (VSV) (340). Cytoplasmic Drosha isoforms are also known to exist (341). However, the biochemistry of Drosha cleavage of RNAs would make the production of the observed vncRNAs unlikely, as Drosha uses a molecular ruler to cut

approximately 11 bps up the hairpin from its base, then approximately 22 bps from the junction of the loop of the hairpin with the stem (253, 342, 343). A casual inspection of the predicted secondary structures of EBOV GS sequences makes it clear that this would not produce the observed vncRNAs (16). The remaining contender, the integrator complex, and CPSF3L specifically, has been shown to produce snRNAs and miRNAs during infection with *Herpesvirus samiri* via cleavage of stem-loop structures in virally transcribed precursor RNAs (338, 339). Despite the fact that in this case, the process is nuclear, and requires Dicer to complete the maturation of the miRNAs (effectively replacing Drosha's pri-miRNA processing role), CPSF3L is also found in the cytoplasm (344), and may well recognize the secondary structure present in filovirus mRNA 5'-UTRs.

With this target list in hand, we performed siRNA-mediated knockdowns of AGO2, Drosha, RNase L, and components of the endosome and integrator complex in human cells. Unfortunately, the lack of sequence data for *Epomops* bats, and the difficulties associated with transfecting bat cells precluded experiments in either of the bat cell lines used in the initial experiment. Knockdown of AGO2, RNase L, and DIS3 (endosome component) were not deleterious to the production of EBOV vncRNAs. Knockdown of Drosha and CPSF3L (integrator complex) did have an effect on the production of the GP vncRNA, with significance achieved only with CPSF3L knockdown. This would suggest that CPSF3L has at least some role in the production of EBOV vncRNAs, though the extent of involvement cannot be determined from our data. Combining our loss of function data with the fact that CPSF3L is known to be involved in the production of *Herpesvirus samiri* miRNAs, this is not an implausible mechanism (338, 339). What is conclusively evident from this data, however, is that production of the observed EBOV vncRNAs is entirely independent of Dicer, and likely independent of the entire host miRNA processing system. Furthermore, the strongly linear relationship between viral mRNA abundance and vncRNA abundance across our knockdown panel strengthens the conclusion that the substrate for production of filovirus vncRNAs is indeed viral mRNA.

The second component of the validation of the EBOV vncRNAs was an investigation of the association of EBOV vncRNAs with host Argonaute proteins. Argonaute proteins are the catalytic component of RISC, and are the component with which miRNAs have the most direct interaction (286). As such, it would be expected that a functional miRNA will be found in complex with at least one of the four Argonaute proteins. As a general lack of reagents for use with bats prevented us from performing experiments in bat cells, we performed our Argonaute immunoprecipitation (AGO-RIP) experiments in human 293T cells. Although modest but significant enrichment of the L vncRNA and GP vncRNA was found after AGO-RIP in one (L vncRNA) and two (GP vncRNA) of the six independent experiments conducted, the remaining four experiments failed to reproduce these results, despite far more robust enrichment of endogenous host miRNAs. The same antibody, which precipitates all four isoforms of Argonaute, was used for all experiments. It is therefore unclear why we were able to detect enrichment in the first two experiments. However, given that detection was not consistent, the most reasonable conclusion is that any association between the EBOV vncRNAs and Argonaute is of a largely nonspecific and transient character, if any such association occurs at all. This would not be expected to be of any biological significance. This finding has implications beyond this report, as two of the previously reported “EBOV viral miRNAs” are strikingly similar to those identified in our sequencing data, and the authors of these and subsequent reports have assumed that these molecules are functional RNAs. EBV-miR-T2-5p is nearly identical to the sequenced VP40 vncRNA, only being shifted in the 3' direction by a single nucleotide (242). The other, ZEBOV-miR-1-5p was slightly less similar, but was still largely identical, only being shifted from the sequenced VP24 vncRNA by 7 bases (244). Given that the sequences of the EBOV vncRNAs described here are so similar to these RNAs, but do not associate with RISC in any significant fashion, it is unlikely that these molecules form any biologically meaningful association with RISC. This directly contradicts the assumptions made in these previous reports, and provides a useful

cautionary note with regards to the interpretation of *in silico* predictions and unvalidated sequencing detection of so-called “viral miRNAs”. Furthermore, none of the other predicted “EBOV miRNAs” were detected in our sequencing data sets at any significant level, with most being entirely absent. However, as mentioned above, many of the predicted viral small RNAs were in fact detected in sera from infected humans, nonhuman primates, and mice (333). This study also reports detection of the GP vncRNA in liver tissue from an infected rhesus macaque, although no attempt was made to detect the published predicted RNAs. It is possible that our sequencing failed to detect the other vsRNAs due to ligation bias known to be associated with certain small RNA sequencing methodologies, including the one used here. Together, these findings indicate that both the vncRNAs reported here, and some number of the predicted vsRNAs are indeed produced during natural infection. Given that we were unable to demonstrate a role for the vncRNAs we detected, and the fact that some of the previously predicted “Ebola virus miRNAs” were indeed very similar to those I report, I believe that they are unlikely to function as viral miRNAs. Aside from being products of stochastic nuclease activity (likely for some of the predicted vsRNAs detected *in vivo* in other reports that we failed to detect), there are a number of potential alternative explanations for their existence are discussed later. The biochemistry of most small RNA sequencing techniques, including the one used here, require 5'-monophosphate and a 3'-hydroxyl moieties for adapter ligation. After size selection for 19-32 nt RNAs, this methodology should, in principle, be highly selective for miRNAs, siRNAs, and piwi-RNAs. In addition to this, it is generally assumed that high abundance miRNAs detected via these sequencing methodologies must be associated with RISC because the fact that they lack typical mRNA features like a 5'-cap or polyA tail would presumably leave them particularly vulnerable to digestion by any of a number of cellular nucleases (286). Interestingly, however, it has been reported that this may not be true. In fact, it seems that the fraction of the total miRNA pool in a given cell that is actively associated with RISC is actually relatively small (345, 346). Furthermore, there is evidence

that the association of mature miRNAs with RISC is not automatic, and is actually dependent upon a number of factors, to include the abundance of mRNA targets (347). Although these findings would initially seem to provide plausible alternative explanations for the lack of EBOV vncRNA association with Argonaute, the lack of such association in combination with the fact that the EBOV vncRNAs have no association with any other component of the miRNA pathway forces the conclusion that there is little evidence to suggest that these molecules are genuine miRNAs.

The conclusion that the EBOV vncRNAs are not genuine miRNAs is bolstered by the finding that they have no measurable biological effects. The vncRNAs had no significant antiviral siRNA activity in bat or human cells, and were not measurably capable of suppressing translation of a target transcript. This finding eliminates an initial hypothesis of this study, namely that bats use RNAi to control infection.

The function of miRNAs is strongly tied to stoichiometry; that is, they must reach a certain abundance in the cell to have any measurable effect (348). In the two cell lines that support strong replication of EBOV (HepG2 and EpoNi/22.1), vncRNA abundance was found to range from several hundred to over one thousand copies per cell (if a cellular RNA content of 10pg/cell is assumed) (**Fig. 9A**). This straddles the stoichiometric line above which a miRNA would be expected to exert some biological effect (348). Despite this, the vncRNAs failed to suppress the translation of luciferase in EBOV-infected human and bat cells transfected with a construct expressing a luciferase reporter mRNA with EBOV GP vncRNA MBEs in the 3'-UTR. However, siRNAs homologous to the EBOV GP vncRNA did successfully suppress translation of luciferase from the same construct when transfected into either human or bat cells. This demonstrates that the sequence itself is capable of functioning in a suppressive role, but the vncRNAs either do not load into RISC at all (supported by the AGO-RIP data), do not reach a sufficient abundance in cells, or both. Despite the clear lack of function in human and *Epomops* cells, the enhanced ratio of both EBOV vncRNAs relative to vRNA/mRNA in the *Rousettus aegyptiacus* cell line

RO6EJ suggests that their production may serve as some sort of restriction factor in these bats. A significant loss of viral mRNAs to the production of vncRNAs would likely impair the ability of the virus to replicate potentially explaining the reduced susceptibility of *Rousettus aegyptiacus* cells to EBOV, and the fact that *Rousettus aegyptiacus* bats are apparently largely refractory to EBOV infection (211). This hypothesis is supported by the fact that no such disparity in the vncRNA:vRNA/mRNA ratio was observed in MARV infected RO6EJ cells. *Rousettus aegyptiacus* bats are the known reservoir of MARV, and support replication of the virus to a degree that allows for transmission (209).

Table 2: Comparison of EBOV vncRNAs to genuine miRNAs

		Genuine miRNAs	EBOV vncRNAs
Biogenesis	Drosha processing	✓	X
	Dicer processing	✓	X
Function	RISC loading	✓	X
	Transcript silencing	✓	X
	Translational repression	✓	X
Biochemistry	5'-phosphate	✓	✓
	3'-OH	✓	✓
	22 nt length	✓	Variable

As there is no evidence supporting the existence of a significant siRNA response to EBOV or MARV infection in bats or humans, and no evidence that the vncRNAs function as viral miRNAs, (see **Table 2**) we are left with the question of the actual function, if any, of the EBOV vncRNAs. Firstly, we must discuss potential mechanisms of biogenesis that would imply that they are either random degradation products, or simply byproducts of another process, such as transcription. The fact that aligned reads were found to derive almost exclusively from the 5' gene start sequences/secondary structures of mRNAs would seem to make any sort of truly random degradation process fairly unlikely. More specific host nuclease digestion of the 5' stem-loop/hairpin structures is not entirely out of the question, given the range of size classes produced during EBOV infection. The implications of this possibility with regards to the status of the vncRNAs as spandrels are

discussed later. However, in data not presented here, I found that MARV-infected cells produced similar vncRNAs that were almost entirely restricted to the 22 nt size class. It is possible that this is a result of differences in the structure of the 5' stem-loops, but this could also be interpreted as evidence of a more directed process.

Other than digestion, production as a byproduct of transcription must also be considered. Given that most sRNA sequencing methods (including the one used for this study) use a requirement for a 5'-monophosphate and a 3'-hydroxyl to select for sRNAs, it could be imagined that the vncRNAs described here could be abortive transcripts. As nonsegmented negative strand (nsNSV) viruses, filovirus transcription is coupled with the addition of a 5' methylguanosine cap by the viral polymerase (32). In the case of VSV, the prototypical nsNSV, the cap is added after the addition of nucleotide 31 (349). The addition of the 5'-cap is thought to be a quality control step, as most abortive transcripts will not receive caps, and therefore will be rapidly degraded by cellular 5'-3' ribonucleases (97). It is likely that EBOV follow the same or a similar pattern, given the degree of sequence homology between VSV L polymerase and EBOV L polymerase. This would appear to lend credence to this hypothesis. However, in the case of VSV, premature termination of transcription does not occur immediately. Instead, VSV abortive transcripts are found to range in length from 40 to 500 nt, which is considerably larger than the RNAs sequenced in this study (97, 350, 351). In respiratory syncytial virus, abortive transcripts are also 40 nt or greater (352, 353). In fact, the sequencing methodology used here would not detect EBOV abortive transcripts at all if EBOV follows this pattern (again, likely due to homology), as they would be size selected out of the libraries during the library generation process. In addition the remarkable consistency in relative abundance between biological replicates, cell lines, and even wild-type and mutant viruses in the data presented here would seem to make such a stochastic process an unlikely explanation. In particular, the fact that the most abundant vncRNAs are 22-25 nt in length, at least 15 nt shorter than the

shortest common abortive transcripts lends more doubt to the hypothesis that the vncRNAs are abortive transcripts.

Another alternative hypothesis is that the EBOV vncRNAs function in a manner similar to that of the short leader RNAs (leRNAs) produced by other nsNSVs (281, 354, 355). IAV is known to produce small RNAs from the 5' ends of its cRNAs. Instead of acting as post-transcriptional regulators, these small RNAs appear to have a role in the switch from transcription of viral mRNA to the production of viral genomes by acting to enhance IAV polymerase activity (281, 283). Importantly, these RNAs are not known to associate with RISC or have any RNAi activity. Moreover they have a 5' triphosphate (281, 283), unlike the EBOV vncRNAs, which have a 5' monophosphate. Although it is possible that EBOV uses a similar mechanism, this seems unlikely, as EBOV and MARV utilize a different system for regulating the transcription/replication balance that is dependent upon the phosphorylation state of VP30 (91).

One intriguing hypothesis assumes that the lack of association of the EBOV vncRNAs with RISC is at least in part a result of their low absolute abundance, and that if present in sufficiently large quantities, they would associate with RISC to some degree. In this instance, the vncRNAs are a mechanism for viral evasion of host defenses that has lapsed into vestigiality by virtue of having been supplanted by far more potent protein-based mechanisms. EBOV VP35, VP40, and VP30 proteins, along with MARV VP35, are reported to act as VSRs (38, 39, 305, 307). Therefore, a reasonable hypothesis would be that an ancestral virus, which sacrifices a portion of its mRNA pool to generate miRNAs to target host transcripts, would eventually evolve mechanisms to control and suppress this process once an infection foothold is attained. Over evolutionary time, the cost of sacrificing mRNAs to produce miRNAs may have led to this strategy being supplanted in favor of the more rapid accumulation of viral proteins, two of which (VP35 and VP24 in EBOV, and VP35 and VP40 in MARV) are known to be potent suppressors of innate immunity themselves (356). The fact that IFN inhibition and apparent VSR activity due to

dsRNA binding often overlap and are difficult to parse apart further complicates interpretation of what is a direct and what is an indirect function of a protein (293). At least in the case of VP35, which exhibits both VSR and IFN-inhibition, was the VSR activity of VP35 selected for independently of its IFN inhibition properties, or is it just coincidental? If they did evolve independently, which occurred first? These questions appear difficult to resolve. Indeed, we do observe an overall decline in the abundance of vncRNAs specifically as infection progresses from 12 to 24 hpi, indicating that nucleolytic digestion of the substrate molecule declines over the course of infection. A recent study may shed light on these questions and lend support to this hypothesis. Edwards et al. explored the functionality of integrated filovirus VP35 homologs in the microbat genus *Myotis* (229). They found that compared to VP35 from extant filoviruses, *Myotis* VP35 only modestly antagonized type I IFN signaling, and that this inhibition likely occurs upstream of RIG-I activation by PACT. This stands in contrast with MARV and EBOV VP35, which directly interact with PACT to prevent RIG-I activation (33, 229). Importantly, the authors also concluded that the reduced IFN inhibition activity of *Myotis* VP35 relative to EBOV and MARV VP35 is not a result of post-integration evolution, but is instead a phenotype maintained from the time of integration, greater than 18 million years ago. Moreover, *Myotis* VP35 apparently lacks the ability to bind dsRNA, and does not act as a VSR. In addition to this, a recent report by Shi et al. suggests that filoviruses are even more ancient than previously assumed, perhaps by hundreds of millions of years (9). The viral genomes identified in this study, which were sequenced from ray-finned fishes, are phylogenetically basal to *Myotis* VP35 as well as extant filoviruses. A cursory analysis of the sequences obtained in this study suggests that these viruses do not exhibit 5'-UTR secondary structures similar to those predicted for the mammal-associated ebolavirus, marburgvirus and cuevavirus mRNAs, thereby implying that the viruses identified in this study do not produce similar vmiRNAs as those identified in this report (Appendix D. Conversely, analysis of the 5'-UTRs of mRNAs from the newly described Bombali virus (BOMV), an

ebolavirus, yielded predicted structures similar to those of other mammalian-hosted filoviruses (examples in Appendix D. While our interpretation of these data does not preclude the possibility that an RNAi-based mechanism of immune antagonism was a derived trait appearing after filoviruses adapted to bats, it does challenge assumptions that said mechanism pre-dated adaptation to an IFN-competent host. Investigations into the potential VSR/IFN-inhibition activity of the analogous VP35 and VP30 proteins of these viruses would therefore help to resolve this question.

Finally, we are bound to consider the possibility that EBOV vncRNAs are merely byproducts of the secondary structures produced by the GS sequences of EBOV mRNAs that themselves confer no selective advantage, but are nonetheless required for another function (likely regulation of transcription). They would, therefore, represent spandrels in the classic sense (357). Only the secondary structure present in the 5'-UTR of the NP gene of EBOV has been conclusively described as being required for VP30-dependent regulation of transcription (96), the fact that stable stem-loop structures are predicted in the 5'-UTR (formed by the GS sequence and downstream nucleotides) of every known mammalian filovirus is highly suggestive that there is some selective pressure to maintain these structures (32). In this model of vncRNA production, the loss of mRNAs to nucleolytic degradation (producing the vncRNAs) is a cost of maintaining these transcriptionally vital structures. As such, any biological activity they may have would be purely incidental, and not the primary source of the selective pressure that is maintaining them. Instead the requirement for the structures they are derived from would be the locus of selection. If the vncRNAs therefore provide no selective advantage, or have a positive effect that is overwhelmed by the loss of viral mRNAs, we would expect the virus to develop mechanisms to limit the loss of mRNAs to the production of vncRNAs. There is evidence that this occurs. While the absolute abundance of GS-vncRNAs increased from 12 to 24 hpi, the *relative* abundance (compared to total viral RNA) actually decreased, suggesting the efficiency of production drops over time. This is likely associated with the

progressive hijacking of cellular functions by the virus. It is also possible that the secondary structures themselves have undergone some degree of selection to limit their vulnerability to nucleolytic digestion without compromising their primary role(s).

Although this study, and indeed the entirety of the work presented here is focused on EBOV, it would be neglectful to ignore the untested possibility that the MARV vncRNAs are indeed associated with the miRNA pathway, and may serve some function. The marburgviruses differ from the ebolaviruses in a number of crucial aspects of their biology, including the identity of their secondary IFN antagonists, and the lack of an editing site in MARV GP (4, 356). The marburgviruses are also phylogenetically basal to the ebolaviruses (4). The MARV vncRNAs were also more consistently 22nt in size, to the point that the 22nt form was almost the exclusive form detected in some cases. We must therefore consider the possibility that MARV has retained some RNAi-based mechanism that EBOV has lost. This would support the vestigiality hypothesis discussed previously. For these reasons, the potential function of the MARV-derived vncRNAs should be assessed independently of the findings related to the EBOV vncRNAs presented here.

This report represents the first rigorous biological assessment of the potential for EBOV to produce miRNA-like RNAs. Despite *in silico* predictions and assumptions derived from these predictions, little to no evidence was found to suggest that EBOV produces genuine, functional viral miRNAs. Although molecules that resemble miRNAs are produced, these are not associated with any component of the miRNA pathway, and lack any RNAi activity. A valuable contribution of this work, aside from the identification of multiple novel vncRNAs associated with EBOV and MARV infection, is the validation of the existence of two previously predicted EBOV miRNA-like RNAs. However, contrary to the reports in which they were initially described, there is no evidence to support the assertion that they have any identifiable biological function. Therefore, the conclusion must be that EBOV vncRNAs either have functions unrelated to RNAi/miRNA-mediated gene silencing, are vestigial, or exist as non-functional spandrels. This leaves a significant

opportunity for further investigation of the biology of these RNAs, including those derived from MARV.

Species-specific evolution of Ebola virus during replication in human and bat cells

INTRODUCTION

Viral evolution in context

Generally, evolution is a process that must be observed over long stretches of time. Although there are counterexamples of relatively rapid evolution in complex multicellular organisms, most evolutionary processes require multiple generations over centuries, or, more often, millennia, to produce readily observable changes (358). Viruses provide a fascinating counterpoint, evolving rapidly to conform themselves to ever-changing fitness landscapes, and, in some cases, to adapt to entirely novel replicative environments. Although generation times are obviously far shorter for viruses than for cellular life, they face the same basic principles of evolution as every other organism (231).

Evolution is the result of natural selection and genetic drift acting upon natural variation resulting from stochastic processes including random mutation and semi-stochastic processes such as genetic recombination, biased mutation, and genetic hijacking (358). Genetic drift is an almost entirely stochastic mechanism by which random events unrelated to the phenotype of an organism affect the frequency of a given allele in a population (358, 359). By contrast, natural selection is a non-stochastic process by which alleles with enhanced fitness relative to their cohorts increase in frequency within a population at the expense of less fit alleles (360). These processes require time in the form of successive generations of organisms, and allelic variation to provide the raw material upon which they can act (358). Viruses provide a special case for each of these, as they have far shorter generation times than cellular life, and generate diversity far more rapidly, both through increased rates of mutation and recombination, and through other mechanisms unique to viruses, such as complementation (231). Recombination and

reassortment have been compared to sexual reproduction, and along with other mechanisms, allow viruses to minimize the effects of Mueller's ratchet (361), although it is likely that viruses escape Mueller's ratchet primarily through their very large population sizes, given the very low rates of recombination observed for many RNA viruses (231).

Viral evolution as a multilevel phenomenon

Viral evolution can be argued to occur on two fundamentally interlinked but marginally distinct levels. Micro-scale viral evolution occurs within a host, and is restricted to the effects of drift and selection upon the diversity generated within the host from the founder population (231). By contrast, macro-scale evolution occurs within the population of the virus as a whole. Evolution in the former case inevitably affects the latter, as particularly advantageous new alleles that arise within a single host will, absent drift effects that eliminate them, quickly rise to prominence within the larger population (231). In this case, "generation time" may be thought of as having two levels as well, one at the micro level, consisting of the absolute time to the generation of a new viral particle from the time of initial infection of a cell, and another at the macro level, consisting of the time from the infection of a given host to transmission to a subsequent host. In fact, genetic drift and natural selection act in this manner as well. For example, at the micro level, the stochastic processes involved in the infection of individual organs or even cells within a host inevitably lead to founder effects (a drift mechanism in which only a small and likely non-representative portion of a given parent population establishes a new population) within the populations established (231, 362). At the epidemic level of macroevolution, founder effects are apparent when a single host travels to a new location and initiates a new chain of transmission independent of the diversity of the parent epidemic. On this level of evolution, the rate at which a virus evolves is influenced by a number of factors, including population bottlenecks, changes in host species and demographics (such as rapid expansion

during an outbreak), etc. This yields an evolutionary rate, typically expressed in the format “X substitutions/site/year” (231).

The origins of diversity in viral evolution

Diversity in viral populations originates primarily from random mutation (231), though viruses also undergo forms of genetic recombination that increase diversity (363). However, the ability of specific viruses to undergo recombination varies (363, 364). Negative strand viruses in particular exhibit very low rates of recombination due to the nature of their replication (365, 366). However, fairly convincing phylogenetic evidence exists for at least some degree of recombination in zoonotic EBOV populations (235). Unique to viruses is the process of reassortment by which segmented viruses can exchange homologous segments during a cellular coinfection (367). With regards to random mutation, viruses are unique, as they generally have high rates of mutation per nucleotide relative to cellular organisms (368). In particular, RNA viruses exhibit particularly error-prone replication due to their lack of a proofreading polymerase (excepting the coronaviruses) (368, 369). This subject will be discussed in more detail later, as it relates to the influence of population genetics on viral fitness. In addition to replication error, viruses are subject to mutations resulting from host editing enzymes, including ADARs and APOBECs (247, 248). Recent work has demonstrated that these editing enzymes may serve as significant drivers of viral evolution (370).

Mechanisms of viral evolution: Drift and selection

GENETIC DRIFT

Genetic drift is particularly influential in viral evolution, as viruses regularly undergo events at the macro and micro levels of evolution that result in random changes in the frequency of alleles within the population (231). In particular, the initial infection of a given host represents a particularly notable case of a founder effect, as the total viable

population transferred to a new host will often be several orders of magnitude smaller than the total intrahost population (371). Moreover, the population transferred is almost certain to not be representative of the diversity of the parent population; only the most frequent alleles in the population have any significant guarantee of being transferred to a new host, while less fit alleles will inevitably be included by random chance (371). Within the macro scale population, drift effects are similarly influential as described above.

NATURAL SELECTION

Broadly, natural selection is divided into two distinct mechanisms. Positive selection follows the classic Darwinian model of selection, with an allele increasing in frequency in the population owing to the increased reproductive success of individuals with that allele due to increased fitness (358, 360). By contrast, negative or purifying selection works to remove an allele from the population as a result of decreased reproductive success of individuals carrying that allele due to a fitness impairment conferred by it (358). The relative roles of positive and negative selection can be determined by taking the ratio of the number of nonsynonymous mutations within a protein coding gene to the number of synonymous mutations, known as the d_N/d_S ratio, or ω (372). An ω value significantly greater than 1 implies that the majority of mutations are beneficial, indicating that positive selection is acting on the gene, whereas a ω value significantly lower than 1 implies that most mutations are deleterious, therefore indicating that negative selection is acting on the gene (372). Values of ω approximating 1 indicate a neutral state that could be the result of a relative lack of selection, or a balance of positive and negative selective pressures acting on the gene (373). RNA viruses nearly always have average ω values less than 1, indicating that deleterious mutations are far more common than beneficial mutations; a finding that would be expected given the density of RNA virus genomes and the high degree of pleiotropy observed (231).

COMPARING THE ROLES OF DRIFT AND SELECTION IN VIRAL EVOLUTION

For much of the 20th century, the role of selection in evolution relative to genetic drift was a matter of debate as the “Modern Synthesis” of the early part of the century was challenged by new findings (374). Some argued that most evolution is solely or largely the result of natural selection acting upon neutral mutations (375). Subsequently, a consensus has emerged that views these mechanisms as complimentary and situationally dependent (358). In relatively stable fitness environments, evolution tends to be dominated by the stochastic mechanisms of genetic drift, and by purifying selection (374). By contrast, positive selection plays a greater role when an organism is not sufficiently fit for its environment, and therefore may face more directional selective pressures (358). In the context of viral evolution at the macro level, purifying selection tends to be far more common in virus/vector and virus/reservoir relationships, as the virus must maintain a given level of adequate fitness that facilitates these relationships (231). Arboviruses in particular face a unique situation in that they must successfully replicate in an arthropod vector and a vertebrate reservoir/host (249). In this instance, both hosts provide strong purifying selection that favors the maintenance of a given phenotype that permits the complex natural history of these viruses (249). Our ability to assess the role of selection in viral evolution more broadly is complicated by the fact that differentiation between selection and drift requires knowledge of both the effective population size (N_e)³ and the selection coefficient (s)⁴ (231). Determining N_e for a given viral population is complicated by a number of previously discussed elements of viral life cycles, and calculating s is similarly complicated, particularly by the fact that synonymous mutations are frequently not fitness neutral in RNA virus genomes (231). However, when reasonable assumptions are made, calculated times required for the fixation of a mutation by drift alone are often far longer than those observed for real mutations, suggesting that selection is indeed acting

³ The number of individuals in a given population that contribute alleles to the next generation.

⁴ A factor derived from the effect a given mutation has on fitness, positive or negative.

on viral populations (231). Although it can occur within a seemingly neutral fitness environment, selection often occurs when viruses are introduced to new replicative environments/fitness landscapes. This may occur when a virus spills over into a new host, or in response to immunological pressure (such as the emergence of escape mutants) (231).

Selection above the level of the individual: Hypotheses and basic mechanisms

GROUP SELECTION AND KIN SELECTION

Although the conventional Darwinian view of natural selection focuses on the individual as the unit upon which selection acts, some authors have historically suggested that selection can act upon traits that benefit a group of individuals instead of a single individual, a hypothesis referred to as group selection (376). Such a mechanism would allow for traits that are deleterious to the individual but advantageous to the group to be selected for (376). This notion of group selection has been largely dismissed for a number of reasons, including the fact that it is inefficient as a means of selection for the traits it proposes to explain (such as altruism) (377, 378). However, kin selection, a refined form of group selection that restricts the relevant group to related individuals, does appear to be relevant in instances where individuals are very closely related, such as in a colony of bees, or an intrahost population of viral genomes (231, 377-380). In this manner, we may think of the genetic diversity produced during a viral infection as a kin group within which individuals with deleterious mutations may make a contribution to the fitness of the kin group that allows the “altruistic” trait in question to survive at a certain level within the viral population.

INTERGENOMIC INTERACTIONS

A number of mechanisms at play in viral infections lead to selection acting upon more than one viral genome. Perhaps most prominent among these is complementation (231). Complementation occurs during coinfection (often, but not exclusively, cellular

coinfection) when either the wild-type genome or another (distinct) mutant genome facilitates the replication of a defective genome (381). Quite clearly, this can facilitate the maintenance of deleterious alleles in the population, at least at the micro scale (382). Although this can have a direct fitness cost, it has been shown that in some instances, the increase in diversity and robustness that results can offset this, and even provide indirect fitness benefits (381). Complementation in this context is a form of intergenomic interaction in which a deleterious allele is protected from purifying selection via these interactions (231). However, intergenomic interactions may also occur between genomes of varying fitness, resulting in emergent phenotypic effects that would not be apparent otherwise (381). This is in large part a consequence of the high mutation rates of RNA viruses, and will be discussed in detail in the next section.

High mutation rates and their consequences

THE ERROR-PRONE NATURE OF RNA VIRUS REPLICATION

RNA viruses represent a special case within the broader context of viral evolution. RNA viruses evolve at a rate unmatched by any other life form, primarily as a result of their low-fidelity genome replication (231, 368). Although RNA viruses are often said to have error-prone polymerases, this is not strictly correct. With the exception of the coronaviruses, viral RNA dependent RNA polymerases (RdRps) lack a proofreading capability, and as a result misincorporated bases are not corrected (231, 369). It is primarily as result of this, and not any aspect of the biochemical/catalytic function of the core RdRp domain of the polymerase, that RNA virus replication is error-prone (383). The absolute error rates of RNA virus RdRps are difficult to calculate, as most techniques for doing so are only capable of detecting genomes that are at least somewhat viable (384). As such, non-viable genomes are not typically detected, thereby reducing the calculated error rate as a result of their not being counted (384, 385). Novel methods of calculating mutation rate, including some utilizing polymerases in cell-free systems, typically yield higher

absolute error rates as a result (385). It is at this point that we must draw a distinction between polymerase error rate, which is a function of the biochemistry of the polymerase protein itself, and mutation rate, which is an estimate of the rate at which mutations are introduced into a given viral population (384). Polymerase error rates are largely inflexible, and will only change if the biochemistry of RNA synthesis is altered (as occurs when viruses are passaged in the presence of ribavirin) (386). Mutation rates, by contrast, are a product of the inherent error rate of the polymerase along with any number of external factors, which may include various host factors such as RNA editing enzymes (247, 384, 387). For all practical purposes, however, it can generally be said that depending upon the virus, RNA virus mutation rates range from 10^{-3} - 10^{-6} substitutions/nucleotide/cycle of replication (384). This tends to result in an average of one (rarely two) mutations per genome per replicative cycle (231). This high rate of mutation requires us to address the question of mutational robustness within the context of RNA virus evolution. Mutational robustness allows organisms to maintain specific phenotypes/functions in spite of the introduction of random mutations that might otherwise be deleterious (388). Although one might expect RNA virus to have significant degree of mutational robustness as a result of their high mutation rates, it is these high mutation rates that leave RNA viruses with genomes that are far less robust than others. This is because the small genome size imposed by high mutation rates prevents RNA viruses from using many of the typical means of genome hardening used by other organisms, such as redundant genes, pseudogenes, large stretches of “nonfunctional” sequence, etc. (231). However, RNA viruses do use alternative mechanisms to establish a degree of what could be thought of as mutational robustness (although not under the strict definition of the term), and under some models of RNA virus evolution, they may possess a significant degree of robustness (231, 388-390). Epistasis serves as one of these additional mechanisms for preserving traits within viral populations subject to high mutation rates, though unlike in organisms with lower mutation rates, epistasis tends to be antagonistic, with epistatic mutations having the effect of opposing

the fitness effects of one another (231). This then may be seen as a mechanism for reducing the impact of deleterious mutations on the fitness of an individual genome. Given the prominent role of RNA secondary structure in RNA virus biology, it is logical that mutations that correct or compensate for deleterious changes in secondary structure would arise (231). In general, however, the very large population sizes of RNA viruses (at both the micro and macro levels) provide what may be the strongest mechanism for preserving fitness in the face of high mutation rates (231).

INTRAHOST GENETIC DIVERSITY AS AN EVOLUTIONARY STRATEGY

RNA viruses therefore produce considerable genomic diversity within a host as the result of low-fidelity replication of a single “master genome” (the consensus sequence) giving rise to a large, complex, and interconnected mutant swarm of variant genomes of varying degrees of fitness relative to the master genome (231, 391). Although it is quite likely that the selective origin of this phenomenon lies in a need to balance polymerase speed with fidelity (231, 368), two schools of thought exist as to its biological role. The predominant model posits that large and diverse populations provide a fitness advantage to viruses by providing them with the flexibility to rapidly adapt to changes in the fitness landscape and cooperatively overcome barriers to successful replication (231, 391). In this model, the intrahost genetic diversity of RNA viruses might therefore be thought of as a “mutant swarm” that reacts to selection not just on the level of the individual genome, but as a population (391, 392). This idea is closely tied to the concept of “survival of the flattest”, which asserts that a population occupying a lower but broader portion of the fitness landscape will ultimately outcompete a population that occupies a higher but narrower peak (393). This makes diversity a selective end in and of itself.

An alternative hypothesis largely dismisses any fitness advantage provided by diversity as incidental to selection for an optimal balance between polymerase speed and polymerase fidelity (383). In this model, the diversity of viral populations is effectively a

spandrel⁵. It should be noted that contrary to some misconceptions, spandrels can have some moderately positive effects on fitness, but these are not their primary locus of selection (357). It is quite likely that some degree of balancing between speed and fidelity does indeed occur, and it is very possible that the selective *origin* of high mutation rates is in a speed/fidelity tradeoff, convincing evidence exists to support the hypothesis that diversity has become a locus of selection in and of itself, as described previously. Additionally, considerable evidence exists to support the diversity-driven model, as will be described below.

The impact of intrahost genetic diversity is well-described for multiple RNA viruses, such as chikungunya virus (394), hepatitis C virus (395), West Nile virus (396, 397), and several enteroviruses, poliovirus in particular (398-400). A lack of diversity/genetic homogeneity, whether the result of natural phenomena or laboratory manipulation, has been found to significantly impair the fitness of RNA viruses. In fact, polio virus with a high-fidelity polymerase loses pathogenicity *in vivo*, including its neuroinvasive phenotype (399). A similar effect was observed with chikungunya virus, which lost the ability to infect mosquitoes in addition to a loss of pathogenicity (394). However, excessively high mutation rates are equally detrimental (401). At some threshold (commonly referred to as the “error threshold”), the mutation rate becomes so high that the population can no longer maintain the presence of enough individual genomes of adequate fitness, compromising the ability of the virus to replicate (402). Once this point is crossed, the population is said to experience “error catastrophe”, and will eventually become extinct (403). Certain antiviral drugs, such as ribavirin, artificially increase the error rate of the RdRp, and are thought to induce this state (404).

⁵ A trait that is the byproduct of selection for another trait, not adaptive selection. Spandrels provide no selective advantage, but are not removed by selection because they are not particularly deleterious.

QUASISPECIES THEORY

The properties and phenotypes of a mutant swarm can be the result of individual contributions from specific subsets of the population, or they can originate from interactions between genomes, such as complementation and epistasis, as previously discussed. In particular, complementation can permit a genome that provides some fitness advantage to the population as a whole to persist despite low individual fitness (382). This idea that selection-susceptible phenotypes can emerge from positive and negative interactions between genomes, and not merely from contributions from individual genomes is central but not exclusive to quasispecies theory (382, 401). However, quasispecies theory goes further, asserting that the entire mutant swarm acts as a largely unified unit of selection, and that the fitness of the mutant swarm is not merely the sum of the individual fitness values of its constituents (382, 401). At this point, it would be wise to establish a clear definition of quasispecies theory and assess whether or not it is applicable to RNA virus evolution. Quasispecies theory should not be confused with the broader concept of the fitness role of intrahost genetic variation/diversity (231). Both concepts require the production and maintenance of large and diverse populations, and posit that this diversity increases fitness of the population. Furthermore, it is possible for viral genomes to interact with each other in ways that produce emergent phenotypes that have impacts upon the intrahost population as a whole in the absence of quasispecies dynamics (231). The fundamental divide is that quasispecies theory is only operative when mutation rates are so high that the frequency of a given variant is determined not only by its own fitness, but also by the frequency by which it is produced by random mutation. As such, closely related sequences can interconvert via random mutation, with each possible variant related to the master genome, but also to other closely related variants through the probability of interconversion via random mutation. Therefore the fitness of a given genome is inherently entangled with that of its neighbors in sequence space to such an extent that the entire

swarm behaves in many ways as a single selective unit (231, 405). As such, quasispecies theory does require that interactions between the variant genomes that comprise intrahost diversity produce emergent phenotypes that have an effect on fitness (392, 401). Quasispecies theory provides a convenient framework for explaining a variety of phenomena observed in RNA virus evolution. In the aforementioned examples of polio and chikungunya virus, the loss of fitness observed can be ascribed to a loss of emergent properties/phenotypes that originate in the diversity of the mutant swarm/intrahost genetic diversity as a result of the mutation rate dropping to a level too low to sustain quasispecies dynamics (399). While complementation is usually a positive or neutral interaction, negative interactions also occur (382). Of these, interference is one of the most common. Acting in a role similar to dominant negative phenotypes in more complex systems, interference occurs when a member of a population actively reduces the efficiency of a given process (231). Consider Protein X, a protein that only functions in a pentameric form. Assume that a cell is co-infected with two genomes, one of which (Genome A) carries a functional copy of the Protein X gene, and another (Genome B) which carries a defective copy. If the defective version of Protein X is incapable of interacting with the functional version, the functional copy will complement the defective one, and the fitness of that infection will not be greatly compromised. However, if the defective version can interact with the functional copy and form nonfunctional Protein X pentamers, a large number of Protein X pentamers will be nonfunctional, and the overall fitness of the infection will be compromised (382). In this case, the defective Protein X is said to interfere with the functional one. This type of interaction can occur on a higher level, particularly in the case of viruses that exhibit superinfection exclusion. As an example, defective interfering particles that are able to enter cells and initiate superinfection exclusion prevent viable genomes from replicating in that cell (382). A balance between these positive and negative interactions is thought to be one determinant of population fitness in quasispecies theory (382, 401). Therefore, within this construction, interference is closely tied to error

catastrophe, as an excessive number of interfering genomes will accumulate relative to functional genomes, dramatically reducing fitness and eventually forcing the population toward extinction (382, 403).

Although quasispecies theory provides a useful conceptual framework for understanding RNA virus evolution, a number of critiques have been leveled largely centered around two objections: 1) that the use of the term “quasispecies” within virology is often inappropriate and not related to the original mathematical model proposed by Eigen and others, and 2) that it has not been definitively established that RNA viruses actually have mutation rates high enough to achieve quasispecies population dynamics in real infections (231). Domingo and others have argued that the operation of quasispecies dynamics within real viral populations are not precisely the same as those described in theoretical mathematical models of quasispecies behavior, and that the requirements imposed by these models are not inevitably applicable to viral populations in actual infections (405, 406).

Host effects on EBOV evolution

The evolution of many viruses is closely tied to their ecology, and EBOV is no exception. The effects of reservoirs and vectors on the evolution of viruses are well-described, particularly for arboviruses (231, 396, 397). In the case of EBOV, the influence of bats on the evolution of the virus is effectively unknown, and given that bats are known to utilize a number of unique antiviral strategies (330, 407), this represents a significant gap in knowledge. In fact, the entirety of our direct knowledge of the evolution of the virus comes from the study of human outbreaks, which almost certainly involve dramatically different selective pressures and models of population growth (231). However, the study of human outbreaks does provide an opportunity to examine the evolution of the virus in a radically different host, and identify adaptive changes that occur.

While the effects of host-specific conditions on the observed mutation rate of EBOV are unknown, and may or may not differ between reservoir and non-reservoir hosts, the factors that dictate evolutionary rate during circulation (i.e. positive/negative selection, genetic drift) likely vary (230). Experimental data demonstrate that the animal passage history of EBOV influences its infectivity and virulence during subsequent infection of a new host species (408), and a similar effect is presumed to occur in natural settings (409).

Host RNA editing enzymes and their effects on viral evolution

Recent work has highlighted the role of host RNA editing enzymes in viral biology, both as antiviral and proviral factors. Some RNA editing enzymes appear to be components of host antiviral defenses, including several members of the APOBEC family of cytidine deaminases, which generate C to U mutations via the deamination of cytidine. Others appear to primarily have roles in cell biology functions including posttranscriptional regulation, particularly in development. This group includes the adenosine deaminase acting on double stranded RNA (ADAR) family. ADARs recognize double stranded RNA and convert adenosine to inosine via deamination, with the inosine being read as a guanosine. This results in an A-to-G mutation. Despite their prominent and well-described roles in cell biology, ADARs have been described as being both proviral and antiviral for a wide variety of RNA viruses. Interestingly one ADAR, ADAR1, is interferon-inducible. It is thought that in addition to potentially disabling genomes by introducing random mutations, hypermutation induced by host RNA editing enzymes may push viral populations closer to the error threshold. On the other hand, if viral populations do not cross the error threshold, these enzymes may in fact increase fitness by increasing the diversity of the population. In particular, RNA editing enzymes are a major contributing factor to the discrepancy between the inherent error rate of the polymerase and the observed mutation rate of a given viral population.

Prior work

The 2013-2016 West African EBOV epidemic generated an unprecedented abundance of sequencing data. Several fixed putative adaptive mutations were identified. Furthermore, at least two, and possibly three of these were under positive selection (233, 410, 411). Despite exhibiting increased fitness in cell culture, no obvious difference in pathogenicity from the parental virus was found in mouse and Rhesus macaque models of EBOV infection (412). However, mice do not recapitulate human or NHP disease, and the size of the Rhesus macaque groups used was insufficient to detect a possible shift in pathogenicity. Furthermore, no significant attempt was made to determine any effect of the mutants on transmission, a significant contributor to the fitness of a virus during an outbreak. In the present study, we sought to characterize EBOV adaptation to cells of bat and human origin. In order to assess changes in mutation rates and the structure of EBOV populations during serial passage through either human (293T) or bat (EpoNi/22.1, *Epomops buettikoferi*) renal cell lines, we utilized Circular Sequencing (CirSeq) (413). CirSeq is an Illumina platform based ultra-deep sequencing approach that uses specialized library preparation and computational protocols to eliminate the vast majority of sequencing errors, reducing the error rate of sequencing to as low as 10^{-12} per base. This permits variant calls at a far lower threshold. We identified several clusters of high-frequency mutations in each cell line. From these, we identified a number of host-specific mutations which appeared to have undergone positive selection. In addition, a particularly prominent cluster of mutations in the region spanning the glycan cap (GC) and mucin-like domain (MLD) of the glycoprotein (GP) of EBOV passaged in EpoNi/22.1 cells was identified. Finally, we selected several mutants from each cell line for further investigation using both infectious EBOV prepared via reverse genetics, and the EBOV minigenome system. Along with characterization of replication kinetics in each cell line, co-infection experiments were performed to assess the fitness of the selected mutant viruses relative to

wild-type EBOV. Our results offer insight into the effects of host factors on the evolution of EBOV, and highlight the capacity of the virus to rapidly develop potentially adaptive mutations in diverse hosts.

MATERIALS AND METHODS

Experimental passages

Recombinant Ebola virus (strain Mayinga) was rescued as described previously (153). The eGFP transgene present in the EBOV FLC provided by Drs. John Towner and Stuart Nichol (CDC) was excised by restriction digest with BSiWI. The EBOV NP, VP35, L, VP30, and T7 polymerase support plasmids were provided by Dr. Yoshihiro Kawaoka (University of Wisconsin). The input virus stock (passage 2) was blind passaged three times (“adaptation passages”) in either 293T or EpoNi/22.1 cells, followed by two sequential rounds of terminal dilution in the respective cell lines, from which three clonal virus populations were selected. Isolated viruses were amplified by two passages in their respective cell lines to generate viruses with sufficient titers for experimental passages. Titration was performed by inoculating confluent monolayers of Vero E6 cells with serially diluted virus allowing the virus to adsorb for 1 hour at 37°C, 5% CO₂. Following adsorption, a 0.6% methylcellulose, 2% FBS MEM overlay was added and the cells were incubated for 5 days at 37°C, 5% CO₂. Plaques were visualized by plaque immunostaining using an anti-GP polyclonal primary antibody (IBT Bioservices). For experimental passages, confluent monolayers of 293T or EpoNi/22.1 were inoculated at MOI 0.1 PFU/cell, except for the first passage in 293T, which was performed at MOI 0.01 PFU per cell due to low titers following the amplification passages. Cells were incubated at 37°C/5% CO₂ for 5 days, after which the supernatants were collected, clarified by centrifugation at 2,000xg, and frozen at -80°C prior to titration and purification. This process was repeated for an additional six experimental passages, at MOI 0.1 PFU/cell. Viruses were purified for RNA extraction by sucrose gradient ultracentrifugation.

Supernatants were layered over 25% sucrose (w/v, diluted in 1X STE buffer), and centrifuged at 175,000xg for 2 hours at 4°C. Pelleted virus was resuspended in 0.5 mL of STE buffer and sonicated in a water bath (amplitude 95 Hz) for 30 seconds. Sonicated samples were layered over a 20%-60% sucrose gradient, topped with 1X STE buffer, and centrifuged at 207,000xg for 90 minutes at 4°C. The virus band at the sucrose cushion was collected with a pipette, diluted with 1X STE, and then centrifuged at 207,000xg for 1 hour at 4°C. Purified virus pellets were resuspended in 100 µL of 1X STE buffer prior to inactivation in 1 mL of Trizol reagent for removal from the BSL-4 and subsequent RNA extraction following the manufacturer's recommended protocol. For 293T cells, host cell rRNA contamination necessitated removal using the GeneRead rRNA Depletion Kit (Qiagen).

Sequencing and processing

Libraries for Circular Sequencing (CirSeq) were generated as described previously (414). 300 cycle, single end reads were generated on an Illumina HiSeq 2500 or HiSeq 4000. Resulting fastq files were analyzed as in Acevedo et al. 2014 (413). Count files from multiple rounds of sequencing and CirSeq processing were combined to obtain final datasets for analysis. Average coverage per base ranged from 94,461 to 509,722 across all EpoNi/22.1 sequenced libraries. For the 293T libraries, values ranged from 80,797 to 296,930.

Determination of population mutation frequency

A maximum likelihood estimation performed on the CirSeq datasets using the script "MaximumLikelihoodEstimation_Q20_Zach.R" was used to determine individual mutation frequencies for each nucleotide variant type (A-to-C, G-to-A, etc.). Only genomic positions with coverage greater than 100,000 were factored in to the calculation. Mann-Whitney U (stats::wilcox.test() in R) and Kolmogorov-Smirnov (stats::ks.test() in R) tests

were used to assess significance between mutation frequencies of a given variant type. Significance testing for each variant type was performed between the 21 data points (7 passages x 3 replicates) from EpoNi/22.1- and 293T-derived viral populations.

Identification of ‘ADAR’ motif

The highest frequency A-to-G (genomic strand)/T-to-C (coding strand) mutations centered around GP’s mucin-like domain and glycan cap were identified for each replicate (the specific region analyzed was from coding strand nucleotide 6,723 to 7,540 of the EBOV clone used). Positions containing variant frequencies at or above the indicated quantile in all three replicates were used for motif analysis via forICELOGO_v5.R. For example, a given mutation needed to be at or above the 0.8 quantile in passage 7 of EpoNi A, B, and C to be included. Each sequence consisted of the position of interest and its surrounding 10 nucleotides (5 upstream and 5 downstream; 11 nucleotides total). Sequence logos were created using ggSeqLogo. Variants were only considered if their coverage was greater than $3 \times 1/(\text{variant frequency})$.

Average number of mutations per read

The Python package ‘pySam’ was used to parse SAM files output by the CirSeq pipeline (data.sam) (script: mismatchesPerRead_combo_AtoGorTtoC_usingPySam.py). These SAM files represent the consensus read sequences resulting from comparing the head-to-tail repeats generated during the CirSeq workflow. The number of T-to-C (or A-to-G) mutations per consensus read was determined, and the average number of each type of mutation per read was determined over the course of passaging for each replicate. Only reads 80 nt or longer and base calls with a quality score ≥ 20 were used.

Fitness estimation of variants

Fitness values were calculated using FitnessEstimator (manuscript in preparation), using a window size of 6 passages and a bottleneck of 106. Significant fitness variants for a given cell line were variants exhibiting beneficial fitness (**wrel.ciLower** [minimum fitness value in 95% confidence interval] > 1) or deleterious (**wrel.ciUpper** [maximum fitness value in 95% confidence interval] < 1) in at least two of three clones. Additionally, it was generally required that at least 5 of the 7 passages had high enough coverage at the position of interest to support the calculated frequency (binomial value in FitnessEstimator). The average of these fitness values were used to compare variant fitness between cell lines.

Visualization of PDB files and determination of $\Delta\Delta G$ value

PDB files were visualized using PyMol. A previously published structure of VP30 (5T3T) was used for visualization and stability testing. A structural prediction of the EBOV L protein was constructed using MODELER, using VSV L (5a22) as a template. $\Delta\Delta G$ values were estimated using FoldX. Each PDB file was first repaired (FoldX command=RepairPDB), then a model was built containing the mutation of interest (command=BuildModel).

Calculation of average Shannon entropy

Entropy was calculated (using shannonEntropy_avgByPassage.R) for each nucleotide position in the EBOV genome (at each passage for each replicate). Shannon entropy for an individual nucleotide was calculated as

Entropy single nucleotide = $-\sum(f * \log_4(f))$ where 'f' is the frequency (i.e. probability) of each possible nucleotide at that position. The average of this value across the genome was calculated for each replicate at each passage, then plotted. C to U (genomic strand) variants were excluded from the calculations. Effect size (Cohen's d) in the region

of the glycan cap and mucin-like domain (defined as nucleotide positions 6723 to 7540), was determined using the `cohen.d` function from the 'effsize' R package. Distribution of Shannon entropy per base in all 293T clones was compared to the distribution of Shannon entropies per base for all EpoNi/22.1 clones at each passage. Only positions with coverage greater than 100,000 were evaluated in any calculation of Shannon entropy.

ADAR RT-qPCR

EpoNi/22.1 and 293T cells were infected with recombinant EBOV (passage 2) at MOI 3 PFU/cell. Monolayers were collected in Trizol 24 hours later. Following RNA extraction, cDNAs were prepared using the iScript cDNA synthesis kit (BioRad) using 20 ng of total RNA. 1 ng of cDNA was used in subsequent qPCR reactions, performed with the iTaq universal SYBR Green mastermix kit (BioRad). Primers and standards for absolute quantitation were obtained from Integrated DNA Technologies (IDT). Bat ADAR1 and 18s rRNA primers were designed using the *Pteropus vampyrus* genome due to the lack of a published Epomops genome. Table 3 provides primers and standards used. qPCR was performed on a QuantStudio 6 thermal cycler. ADAR1 copy number was normalized to 18s rRNA copy number. Significance was tested using a 1-way ANOVA with a Tukey's post-hoc test for multiple comparisons.

Table 3: ADAR RT-qPCR primers

	Forward Primer	Reverse Primer
<i>P. vampyrus</i> ADAR1	ACTTTGAAAACGGCCAGTGG	TAGAAGGACGGCATCTCCATG
Human ADAR1	ATCAGCGGGCTGTTAGAATATG	AAACTCTCGGCCATTGATGAC
<i>P. vampyrus</i> 18s	CACGGCGACTACCATCGAA	CGGCGACGACCCATTC
Human 18s	GTAACCCGTTGAACCCATT	CCATCCAATCGGTAGTAGCG
Standards		
<i>P. vampyrus</i> ADAR1	ACTTTGAAAACGGCCAGTGGGGCCACCGACGACATCCCGGACGACCTG AACAGCATCCGCGCGGGCCCCAGGCGAGTTCCGGGCCATCATGGAGAT GCCGTCCTTCTA	
Human ADAR1	ATCAGCGGGCTGTTAGAATATGCCAGTTCGCTAGTCAAACCTGTGAG TTCAACATGATAGAGCAGAGTGGACCACCCATGAACCTCGATTAAA TTCCAGGTTGTCATCAATGGCCGAGAGTTT	
<i>P. vampyrus</i> 18s	CGGCGACGACCCATTGGAACGTCTGCCCTATCAACTTTCGATGGTAGT CGCCGTG	
Human 18s	GTAACCCGTTGAACCCATTGATGGGGATCGGGGATTGCAATTAT TCCCATGAACGAGGAATCCAGTAAGTGCGGGTGATAAGCTTGCCT TGATTAAGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATT GGATGG	

Rescue of EBOV mutants

Mutated full length EBOV clones were prepared using the wild-type full length clone (FLC) backbone described above. Briefly, for all clones but GP F252S, GP L256P, and the GP triple mutant, site-directed PCR mutagenesis (New England Biolabs) was used to introduce point mutations as identified in our sequencing datasets. For GP mutants, double-stranded DNA fragments (gBlocks, Integrated DNA Technologies) were first subcloned into the Sall and BbsI sites of a pUC19 construct containing the portion of the EBOV FLC between Sall and SacI. The Sall/SacI fragment of this new construct was digested out and inserted via restriction cloning between the Sall and SacI sites of the FLC plasmid. The bicistronic MG was prepared in the Bukreyev lab from a monocistronic MG provided by Dr. Elke Mühlberger (Boston University) (29). This MG consists of the 3' genomic leader, plus the NP 5' UTR controlling transcription of a firefly luciferase ORF, followed by the VP40-GP gene junction region, including the GP 5' UTR, which controls

transcription of a Renilla luciferase ORF. The Renilla luciferase ORF is followed by the L gene 3' UTR and the 5' genomic trailer. Rescue and MG support plasmids (excepting the codon-optimized L plasmid) were as described above. The codon optimized L polymerase plasmid was synthesized by Genescript.

Viruses were rescued following a modified version of the protocol described by Tsuda et al. (415). Briefly, 90% confluent 6-well plates of Huh-7 cells in standard maintenance media were transfected with 1 μ g pCEZ-NP, 0.5 μ g pCEZ-VP35, 0.3 μ g pCEZ-VP30, 2 μ g pCEZ-L-co, 1 μ g PLASMID-T7, and 1 μ g of the appropriate FLC plasmid. Transfection complexes were prepared using *transIT-LT1* (Mirus), with a ratio of 2 μ L of transfection reagent per microgram of plasmid DNA. The next day, media was replaced with fresh DMEM high glucose with 2% FBS and 0.1% gentamicin. Five days post-transfection, supernatants were pooled and adsorbed onto T75 flasks of Huh-7 overnight, with fresh media added the next day. Five days post infection, viruses were collected. To produce stocks of sufficient titer for experiments, viruses were passaged one time on Vero E6 cells infected at MOI 0.1 PFU/cell. Stocks were titrated by plaque immunostaining in 96 well plates.

Phenotypic characterization of mutants

Prior to characterization, stocks were sequenced to ensure population homogeneity at the mutated nucleotides. Multistep replication kinetics assays were performed for all mutants on both EpoNi/22.1 and 293T. Samples were collected at 24, 48, 72, and 96 hours post infection. For each timepoint, three wells of a 24 well plate were infected at MOI 0.01 PFU/cell, and the virus was allowed to adsorb for 30 minutes, after which wells were washed twice with PBS before fresh media was added. Time course samples were titrated by plaque immunostaining in 96 well plates. Significance was tested using 2-way ANOVA, with a Dunnett's post-hoc correction.

Competition and complementation assays were performed by mixing selected mutant viruses 1:1 by plaque immunostaining titer. For competition assays, 293T or EpoNi/22.1 cells in 12 well plates were infected for passage 1 at MOI 0.1 PFU/cell in triplicate, otherwise following the infection protocol as described for kinetics assays. Passage 1 infections for complementation assays were performed at MOI 3 PFU/cell. When significant CPE was observed (3 dpi for complementation assays, 4 dpi for competition assays), supernatants were collected. For all subsequent passages, cells for competition assays were infected with a 1:100 dilution of the supernatant from the previous passage. For complementation assays, EpoNi/22.1 cells were infected with $\frac{1}{4}$ of the supernatant from the previous passage, while 293T cells were infected with $\frac{1}{2}$ of the supernatant from the previous passage, to account for the larger number of 293T cells per well and ensure an MOI greater than 1 PFU/cell. Although this approach does not allow for a precise MOI, and may introduce complications (such as defective interfering particle accumulation), my experience has been that EBOV titers do not appreciably increase with passaging at high MOI. As such, it is likely that the effective MOI was similar across passages. Samples of inocula (in triplicate), and supernatants were removed from BSL-4 in Trizol (ThermoFisher) for column-based RNA extraction (Zymo Direct-Zol RNA micro-prep). RT-PCR amplicons were generated (Qiagen One Step RT-PCR kit) using the primers provided in **Table 4**, and were Sanger sequenced with technical duplicates following enzymatic PCR cleanup (Genewiz). Analysis of sequence data was performed via poly-SNP (416), using the area under the curve method to determine the relative proportion of each virus within the sequenced population. Significance between passages within cell lines was tested with 2-way ANOVA, with a Tukey post-hoc correction. Significance between cell lines was tested via two-tailed Wilcoxon nonparametric test.

Table 4: Competition assay PCR and sequencing (indicated by *) primers

	Forward Primer	Reverse Primer
NP N5566S	CAGGCTTATTGATTGTCAAA	TGTCACTGTCCTGGTTCCTG*
GP L256P	AAGGTGTCGTTGCATTTCTG*	CTCGTGTGGTGTCTCTCTGC
VP30 E205G	AGTACCGTCAATCAAGGAGC	ATCAGACCATGAGCATGTCC*
L C1211R	CATCAACTCCTGTTATGAGT	GATCGTTGTACCTGTGAACA*
L S1992G	AGGTGCTGGTGCCTTACTAT	CGAATCTCTGCTCTAAGATG*

Minigenome transfections were performed in triplicate in 6 well plates as previously described (45). A codon-optimized L polymerase plasmid was used. Control transfections omitting the L polymerase plasmid were performed with both the wild-type and mutant bicistronic MGs. Dual-luciferase assays were performed to assess the efficiency of translation of the Renilla luciferase ORF relative to the firefly luciferase open reading frame by taking the FFL:RL signal ratio. Comparison of the efficiency of translation from the wild-type MG to translation from the mutant MG was determined by dividing the mutant ratio by the wild-type ratio and taking the reciprocal. This yields a value representative of the loss of efficiency resulting from the mutation. Data presented is representative of three independent experiments. Significance was tested using one-way ANOVA, with a Dunnett's post-hoc correction.

RESULTS

Experimental evolution through serial passaging of EBOV in human and bat cells

The cell line EpoNi/22.1, derived from renal epithelia of an adult *Epomops buettikoferi*, was selected. This species is a close relative of *Epomops franqueti* (417). 293T cells, derived from human embryonic kidney, were used for comparison. These cell lines were selected due to their similar tissue origin and the fact that they replicate the virus to similar titers. The latter is an important consideration for population genetics, as vastly divergent population sizes complicate analysis.

EBOV was rescued from the full-length clone plasmid in 293T cells. Passage 2 (p2) virus was blind passaged three times in either EpoNi/22.1 or 293T cell lines for initial “adaptation”. This step reduces the risk of interference from any extremely high-fitness mutations associated with early passage in a given cell line that may either obscure or artificially inflate the fitness of lower frequency mutations. Next, viruses were put through two rounds of terminal dilution. Three clonal isolates were selected from each cell line. Resulting titers were low, and two rounds of amplification in their respective cell lines were required. For experimental passages, monolayers of 293T or EpoNi/22.1 were inoculated at MOI 0.1 PFU/cell. The first passage in 293T was performed at MOI 0.01 PFU/cell for all replicates as the titer was low following the amplification passages. This process was repeated for a total of seven experimental passages. Supernatants from each passage were collected, and virus was purified via sucrose gradient for subsequent RNA extraction and sequencing. **Figure 17** presents an overview of the experiment.

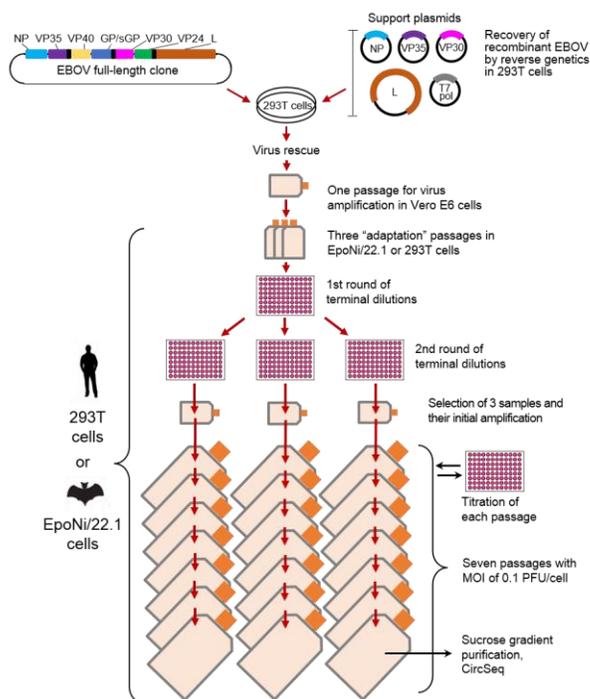


Figure 17: Schematic representation of experimental design

In 293T, titers for Clones A and C remained relatively stable throughout the passage series; however, starting in passage 6 a precipitous decline in titer was observed for Clone B (**Fig. 18**). No substantial difference was observed in the three clones passaged in EpoNi/22.1.

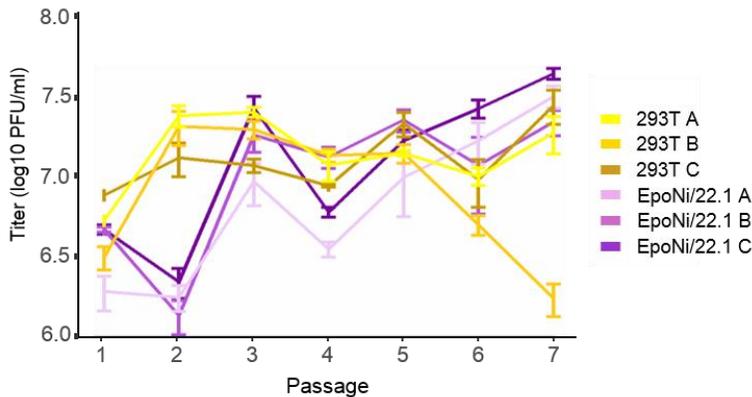


Figure 18: Titer of clones across passage series

The titer of clone 293T B drops dramatically during passages 5 and 6. Limit of detection 10 PFU/mL.

EBOV takes distinct evolutionary paths in human-derived and bat-derived cell lines

Viral genomic RNA isolated after each passage was used to prepare libraries for Circular Sequencing (CirSeq) (413, 414). First, the data was used to calculate individual mutation frequencies for each possible nucleotide variant (A-to-G, C-to-A, U-to-C, etc.) (**Fig. 19A**). Averaged over all clones and passages, we found that overall mutation frequencies were similar between the EpoNi/22.1 and 293T-derived viruses. However, frequencies of A-to-G, G-to-A, and U-to-A mutations were all significantly different between the cell lines (**Fig. 19A**). In particular, there was a clear increase in the frequency of G-to-A transition

mutations (with respect to the genomic strand) in the EpoNi/22.1 derived viral populations relative to 293T (**Fig. 19A**).

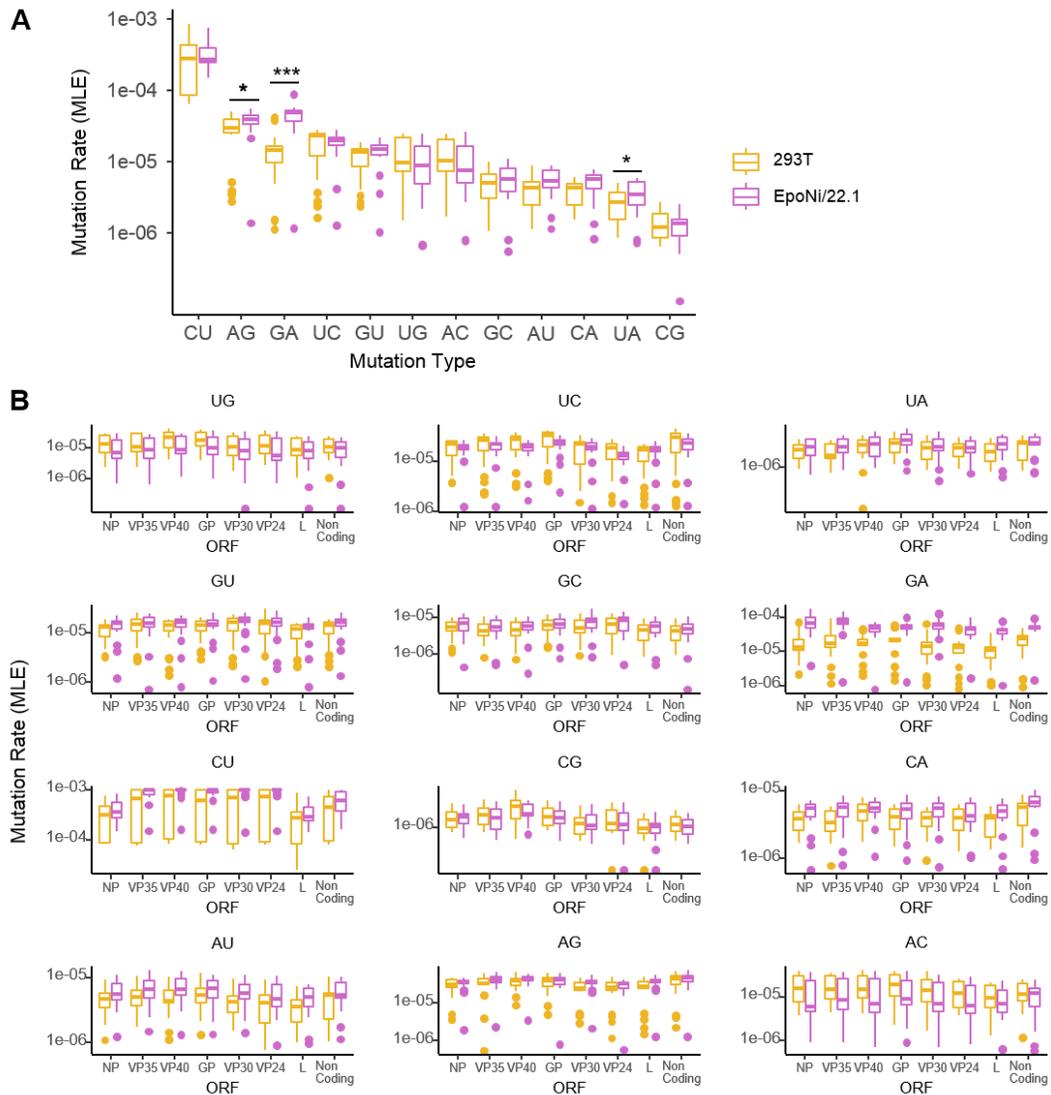


Figure 19: Population-level characterization of passaged EBOV populations

(A) Mutation frequencies for each type of nucleotide substitution. Each boxplot represents 21 data points (three clones over seven passages). (B) Identical to panel A, but mutation frequencies were calculated on a per ORF basis, rather than across the entire genome. G-to-A mutations are significantly elevated in EpoNi/22.1 cells across all ORFs. “Non-coding” represents all non-coding regions as a single entity. * = p-value < 0.05 (by both Mann-Whitney U and Kolmogorov-Smirnov); *** = p-value < 10^{-6} .

To determine if mutation rates within individual viral genes differed between cell lines, we recalculated individual mutation frequencies, treating each ORF as an independent region. Frequencies were consistent, indicating no gross differences in the spontaneous RdRP mutation rate due to genomic position (**Fig. 19B**). This also held true for the increased G-to-A frequency observed in EpoNi/22.1-derived viral genomes, with all examined regions exhibiting a similar pattern relative to 293T derived virus (**Fig. 19B**).

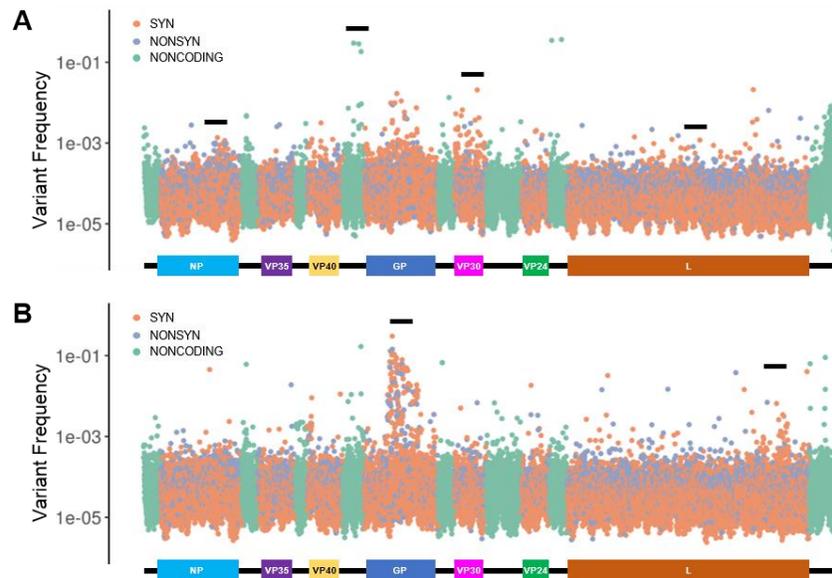


Figure 20: Average variant frequency among replicates

Average frequency for every variant in the EBOV genome across (A) 293T-derived and (B) EpoNi/22.1-derived virus populations at passage 7. C to U(-/genomic)/G-to-A(+/-coding) mutations are not shown due to their high frequency to improve visual clarity. These mutations are extremely common, and would obscure the general pattern observed were they included. As some are not immediately apparent in these representative figures, black horizontal bars denote locations of variant clusters discussed in text.

Clear differences in variant frequencies at the final passage highlight the distinct evolutionary paths of the 293T-passaged and EpoNi/22.1-passaged populations (**Fig. 20**). Comparing Shannon Entropy over time, we found a high degree of homogeneity in passage 1, with increasing heterogeneity over the course of passaging (**Fig. 21**). Overall, the average genomic Shannon Entropy estimated in the EpoNi/22.1-passaged replicates was

moderately higher than in their 293T-passaged counterparts (**Fig. 21**). As a technical note, we did observe some differences in mutation frequency of certain nucleotide substitutions depending upon the sequencer used (Illumina HiSeq 2500 vs. 4000). The HiSeq 2500 tended to exhibit lower rates of mutation for UG, AC, and CG. However, the differences were not particularly remarkable, and had no meaningful impact on our findings.

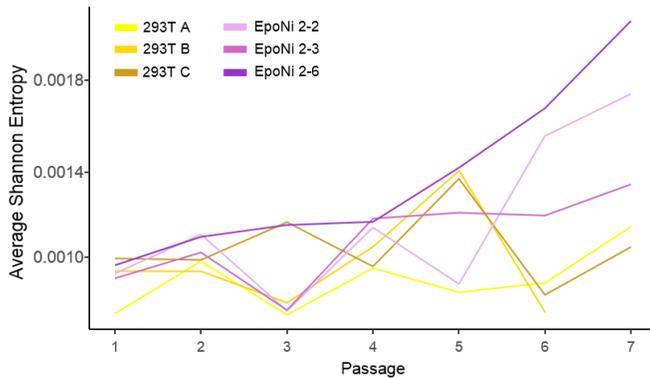


Figure 21: Average Shannon entropy of passaged virus

The average Shannon Entropy increases faster in the EpoNi-derived viral genomes than in 293T-derived genomes. The increase in the average genome-wide Shannon entropy in the EpoNi/22.1-passaged lines was largely due to the region of GP's glycan cap and mucin-like domain, where a small, but measurable effect size could be detected (Cohen's $d = 0.31$ at passage 6).

EBOV populations passaged in bat-derived cells exhibit a ‘spike’ of high frequency mutations consistent with ADAR activity in the glycan cap/mucin-like domain region of the GP protein

An intriguing pattern of mutations arose during passaging of EBOV in EpoNi/22.1 cells. Two of the three EpoNi/22.1-passaged viral populations acquired localized peaks of high frequency variants over the course of passaging (**Fig. 22A**). These mutations arose primarily after passage 4, and consisted almost entirely of adenosine (A) to guanine (G) substitutions (genomic sense). These peaks of mutations were localized within the region spanning the glycan cap (GC) and mucin-like domain (MLD) regions of GP (**Fig. 22B**).

An increased frequency of A-to-G substitutions in these regions was also detected in the other passage series (including the 293T-derived populations) at a very low frequency.

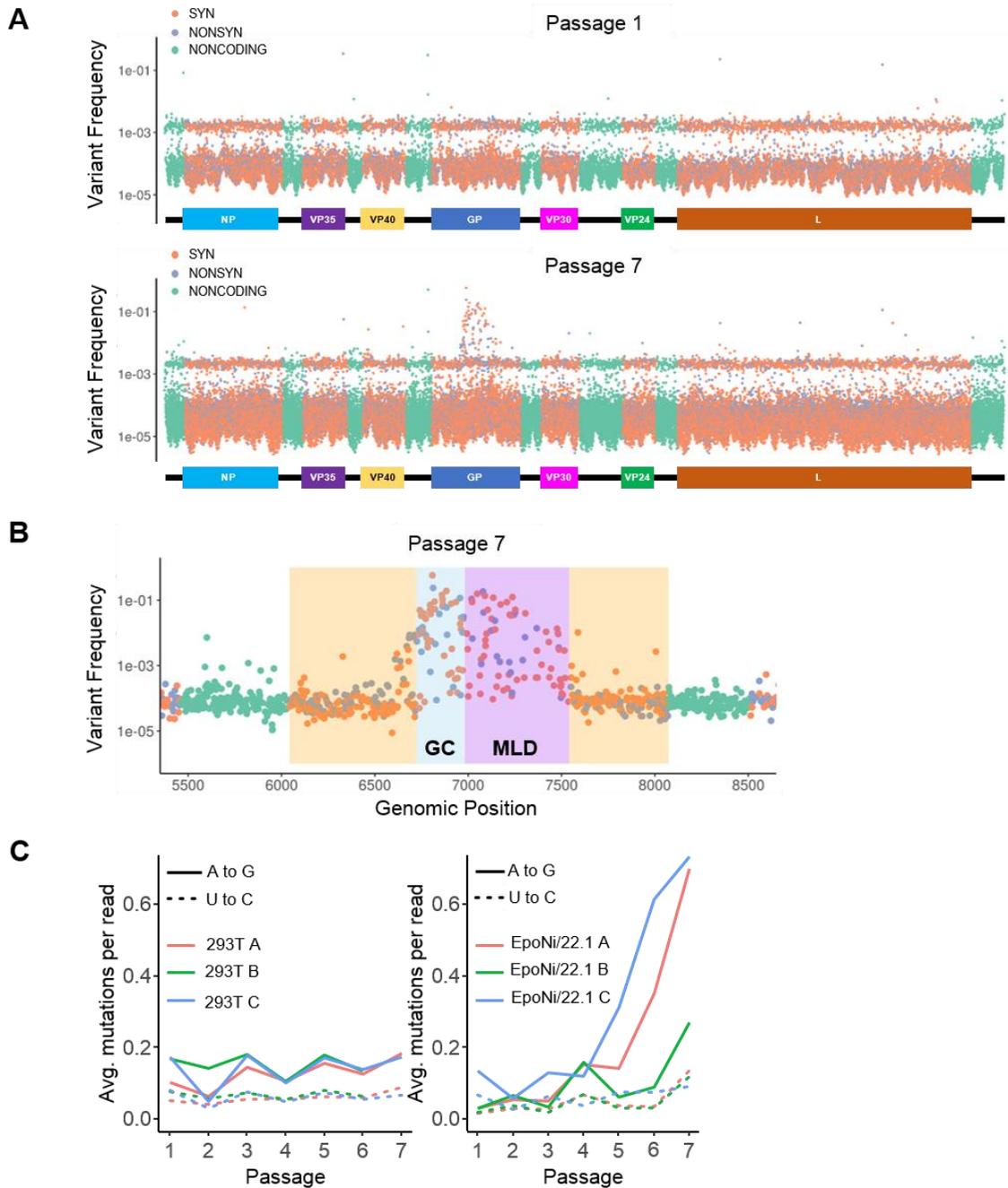


Figure 22: Bat cell-passaged virus accumulates high-frequency A-to-G mutations

(A) Variant frequencies plotted against nucleotide position, color coded by type of amino acid change at passage 1 and 7 of the EpoNi/22.1 “C” population. **(B)** Detailed view of the GP coding region, showing only A-to-G (genomic sense) mutations. The majority of

mutations in the high frequency ‘spike’ are A-to-G, and localize primarily within in the glycan cap (light blue) and mucin-like domain (purple) of GP (yellow). GC, glycan cap, MLD, mucin-like domain. (C) The average number of A-to-G (solid lines) and U-to-C (dashed lines) mismatches per read in EpoNi/22.1-derived (left panel) and 293T-derived (right panel) virus genomes. Only reads longer than 80 bp were examined.

As a result of the relatively small size of the region in which they were located, we were able to determine if multiple mutations appeared on a single genome. In two of the three replicates passaged in EpoNi/22.1, the average number of A-to-G substitutions (genomic strand) per read (i.e. genome) increased dramatically over the course of passaging (**Fig. 22C** - solid lines), while U-to-C substitutions did not (**Fig. 22C** - dashed lines). EBOV therefore accumulates these A-to-G mutations on the same genomes without apparent detriment. Such a high level of mutational robustness is supported by previous studies in other viruses (390). Synonymous and nonsynonymous mutations were accumulated in a relatively unbiased manner, though nonsense mutations did not appear at any significant level (data not shown). It is likely that these were removed through purifying selection.

A compelling explanation for this phenomenon is editing activity associated with the ADAR family of RNA editing enzymes. (418). ADARs edit dsRNA by creating adenosine to inosine mutations, ultimately resulting in A-to-G substitutions. ADARs have been implicated in editing the genomes of a number of viruses (247, 370, 387, 419, 420), including EBOV (98, 132, 234, 421, 422). We investigated whether any ADAR motifs were enriched in the highest frequency variants in this region. Examination of the 10 nucleotides surrounding the most frequent A-to-G (genomic strand) variants revealed a motif matching that expected of ADAR editing (5'-(U/A/C)AG/U-3') (**Fig. 23A,B**) (423). Lending further support to the hypothesis of ADAR editing in viral populations derived from EpoNi/22.1, we found that EpoNi/22.1 cells express approximately 12-fold more ADAR1 mRNA than 293T cells. EBOV infection did not significantly increase ADAR1 expression in either cell line (**Fig. 23C**).

In summary, we have found that during passaging in bat cells, a region encompassing parts of the glycan cap and mucin like domain of EBOV GP undergoes hypermutation in the form of a drastic increase in the rate of A-to-G mutations. These mutations are consistent with the described editing activity of ADAR, an isoform of which was found to be expressed in significantly greater quantities in the bat cells used relative to the human cells used.

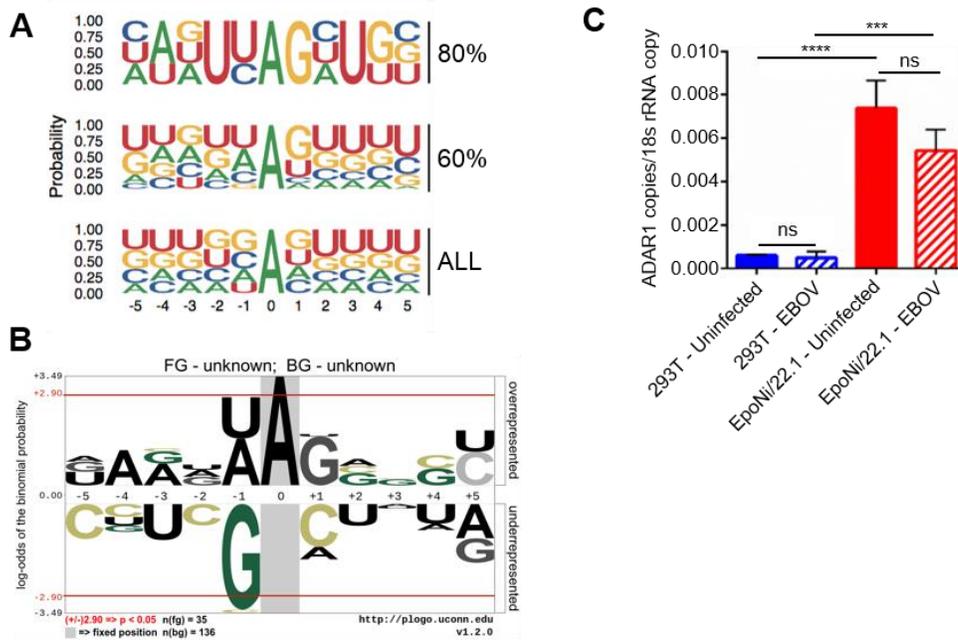


Figure 23: High-frequency A-to-G mutations in EpoNi/22.1 are likely the result of ADAR activity

(A) Sequence logo of the nucleotides surrounding the most frequent A-to-G (genomic sense) mutations. Each tier represents the sequence logo of A-to-G mutations that were in the top 80% (by frequency, the upper panel) or top 60% (the middle panel) in all 3 populations from EpoNi/22.1-passaged virus at passage 7. “All” (the lower panel) represents the sequence logo of surrounding nucleotides for all adenosines in the examined region. (B) pLogo plot of high frequency A-to-G mutations in EpoNi/22.1 passages (genomic sense). The 0.5 quantile of the most frequent A-to-G substitutions (determined with the same methodology as in A) was compared to the entire region from nucleotide 6,723 to 7,540. Only the depletion of ‘G’ in the +1 position relative to suspected ADAR target sites was identified as statistically significant ($p < 0.5$). (C) RT-qPCR (absolute quantitation) for ADAR1 normalized to 18s rRNA copy number. Cells were infected with EBOV at MOI of 3 PFU/cell, or mock infected, then collected in Trizol 24 hours later. **** $p < 0.0001$, ns - not significant ($p > 0.05$), one-way ANOVA with Tukey’s post-hoc test.

Human and bat cell passage produced viruses have distinct population structures

In addition to the ‘spike’ of mutations in GP unique to EpoNi/22.1-passaged viruses, we identified individual mutations that rose in frequency over the course of passaging. To identify these, we searched for variants that rose in frequency in at least two of the three EBOV clones passaged in cells derived from a particular host. Several variants in 293T-derived populations identified were in regions associated with transcriptional regulation. These included mutations in NP, VP30, and the gene end/transcription termination signal of the VP40 gene (**Figs. 24, 25**). In NP, variants were found within the protein phosphatase 2 (PP2) interaction domain (93), while those in VP30 were near the region of the protein responsible for interaction with NP (**Fig. 26**). Protein modeling found that mutations identified in VP30 were predicted to decrease the stability of the protein (**Fig. 27A**). The mutations identified in the VP40 gene-end/transcription termination signal (ATTAAGAAAAA) (95) are as follows, with mutated nucleotides underlined: TTTAAGAAAAA, ATTATGAAAAA and ATCAAGAAAAA. Mutations were never found on the same read. Also identified in 293T-passages was a variant cluster within the capping domain of the L ORF (**Fig. 26B**). Other than the spike of mutations in GP, the only significant variant cluster identified in EpoNi/22.1 was within the methyltransferase domain of the L ORF. **Fig. 27B** illustrates the predicted impacts of the mutations identified in both cell lines on the stability of L polymerase.

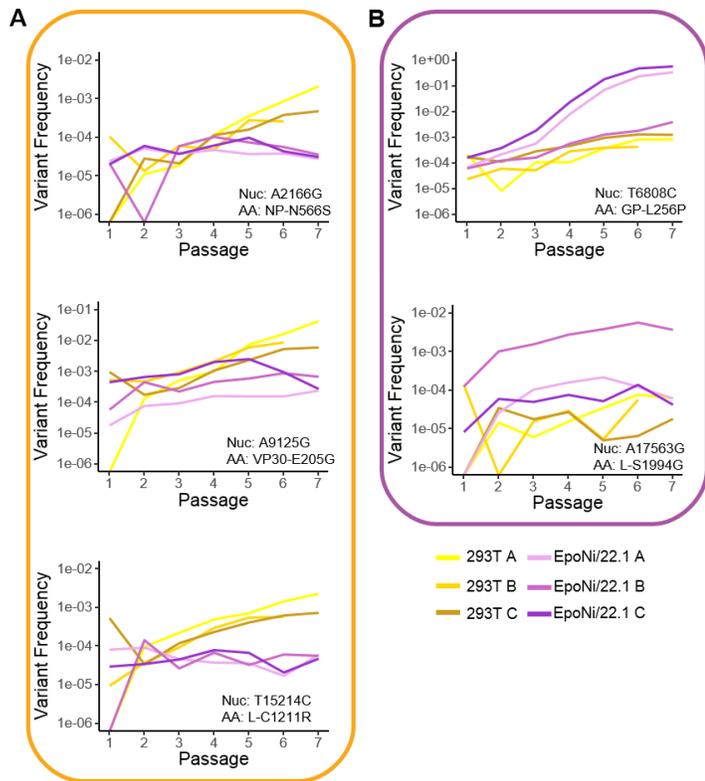


Figure 24: Frequency trajectory of selected variants identified within variant clusters identified in human and bat cells

(A) Mutations identified which accumulate faster in 293T-derived populations as compared to EpoNi/22.1-derived virus over the course of passaging. **(B)** Mutations which accumulate faster in EpoNi/22.1 compared with 293T-derived virus genomes.

Notably, most of the identified mutations did not closely approach fixation in either 293T or EpoNi/22.1, and could only be identified using CirSeq, without resorting to a large number of passages. This demonstrates the utility of CirSeq in experimental evolution studies. Variant frequency trajectories for representative mutants that exhibited higher fitness in one cell line are shown in **Fig. 24**. Overall, we found that passaging had host-specific effects on population structure. Therefore, we sought to determine the effects of these differences by characterizing the infection phenotypes of representative mutants.

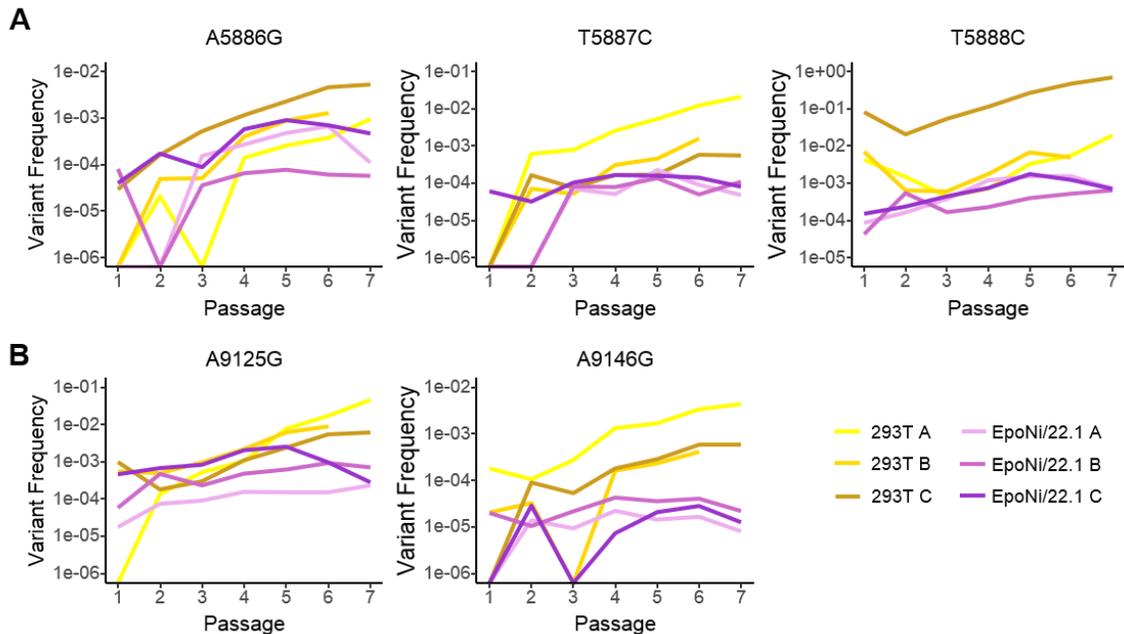


Figure 25: Trajectories of select high-fitness mutations identified in human cell passaged virus

Plots of frequency versus passage for mutations in the VP40 gene end (A) and VP30 (B). 293T replicates are in shades of yellow; EpoNi/22.1 replicates are in shades of purple.

Human and bat cell-derived mutants displayed cell-specific fitness patterns

A total of six mutants, five of which are identified in **Fig. 24**, were selected for characterization. Four mutants were identified in 293T-passaged virus, while two were from EpoNi/22.1-passaged virus. Mutants selected displayed a consistent upward trend in at least two clones during passaging, and were generally the most fit within their variant cluster. Mutations in ORFs were rescued through generation of recombinant viruses using the EBOV full length clone, while the single untranslated region (UTR) mutant, VP40 t5884c (not included in Fig. 24) was tested in a VP40/GP bicistronic minigenome (MG) that was developed for this purpose. Replication kinetics assays were performed under multistep conditions in both 293T and EpoNi/22.1 cells. In this context, multistep kinetics assays provide a more reliable estimate of the fitness of a given mutant, and minimize the potential for complementation by revertant genomes. Given the apparent functional

relatedness of the mutants, we attempted to construct a double mutant containing the NP and VP30 mutant genes identified in **Fig. 24**. Multiple attempts to rescue this virus failed, suggesting that it is nonviable (data not shown).

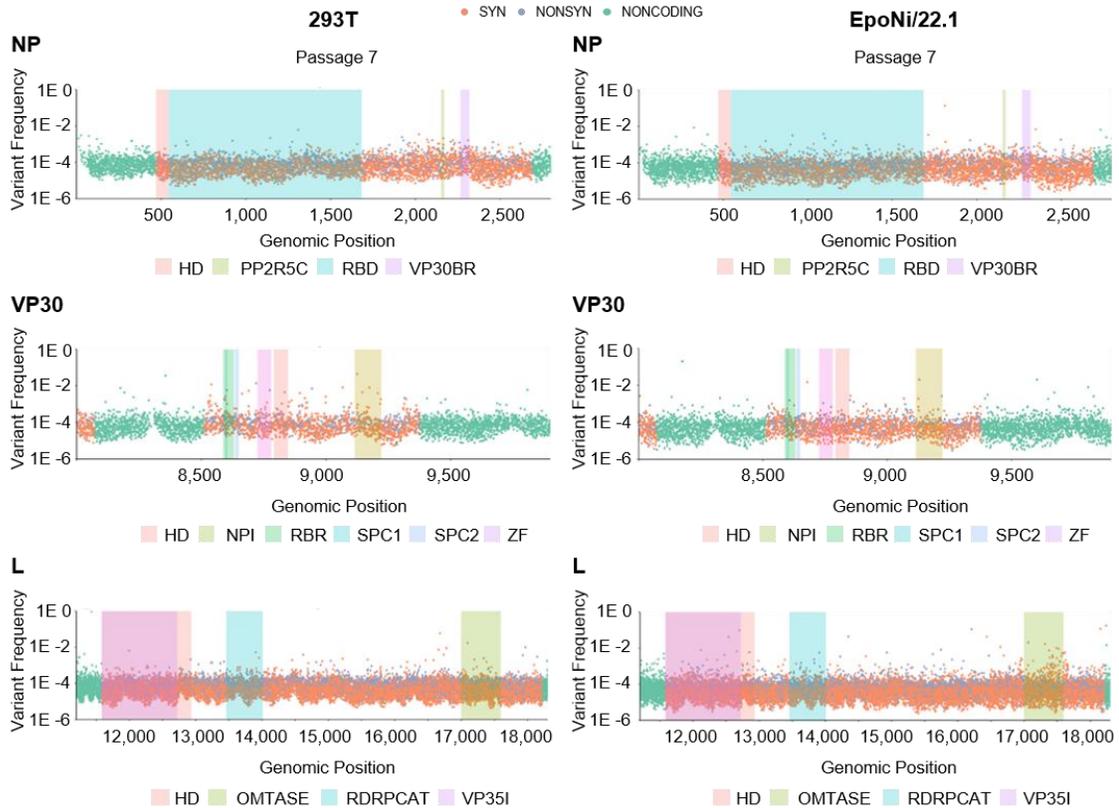


Figure 26: Clusters of variants detected in specific protein domains

Maximum variant frequency at passage 7 among all 293T or EpoNi/22.1 clones is plotted for NP, VP30, and L. Specific domains within each protein are indicated by shaded boxes and indicate the following: NP (Homoligomerization domain; PPP2R5C binding motif; RNA binding domain; VP30 binding region), VP30 (homoligomerization domain; NP interaction; RNA binding region; Serine phosphorylation cluster 1; Serine phosphorylation cluster 2; zinc finger) and L (Homoligomerization domain; SAM-dependent 2'-O-Mtase; RdRp catalytic domain; VP35 interaction).

All mutant viruses had an apparent replicative advantage over the parental wild-type virus in 293T (**Fig. 28A**). In EpoNi/22.1, only GP L256P (identified in EpoNi/22.1) had a statistically significant advantage over wild-type. (**Fig. 28A**). We also identified cell-specific differences in the infection phenotype of both L polymerase mutants. L C1211R, identified in 293T, exhibited a marked deficiency in EpoNi/22.1. Although this did not

reach statistical significance, the loss of replicative efficiency led to an approximately one log reduction in titer across the experiment (**Fig. 28A**).

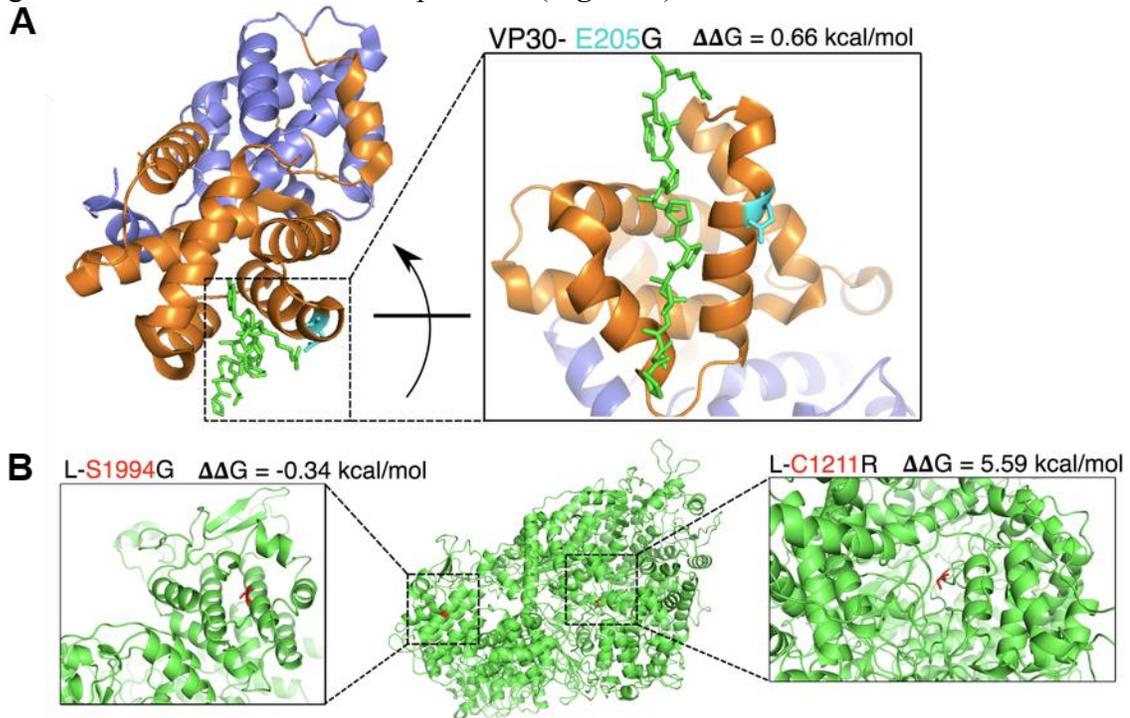


Figure 27: Predicted impact of selected mutations on protein stability

(A) VP30-E205G (highlighted blue residue) on PDB structure 5T3T, visualized in PyMol. The stick structure is a portion of NP, which interacts with the alpha-helix to which E205 belongs. (B) The positions of L-S1994G and L-C1211R (highlighted in red) on the predicted structure of EBOV L protein (based on the VSV L protein). The predicted impact of each mutation, as determined by FoldX, is shown in kcal/mol.

The virus also had a small plaque phenotype on Vero E6 (**Fig. 28B**). By comparison, this mutation was neutral to mildly beneficial in 293T. By contrast, single step kinetics assays performed in EpoNi/22.1 found that L S1994G (identified in EpoNi/22.1) had a significant advantage over wild-type under these conditions (data not shown). L S1994G had a large plaque phenotype on Vero E6 (**Fig. 28B**). Finally, a dual luciferase MG assay was used to demonstrate the potential role of noncoding mutations. We found that the mutation identified in the VP40 gene end/transcription termination signal, t5884c, impaired translation of the second (Renilla luciferase) ORF downstream of the disrupted gene-end signal (**Fig. 28C**).

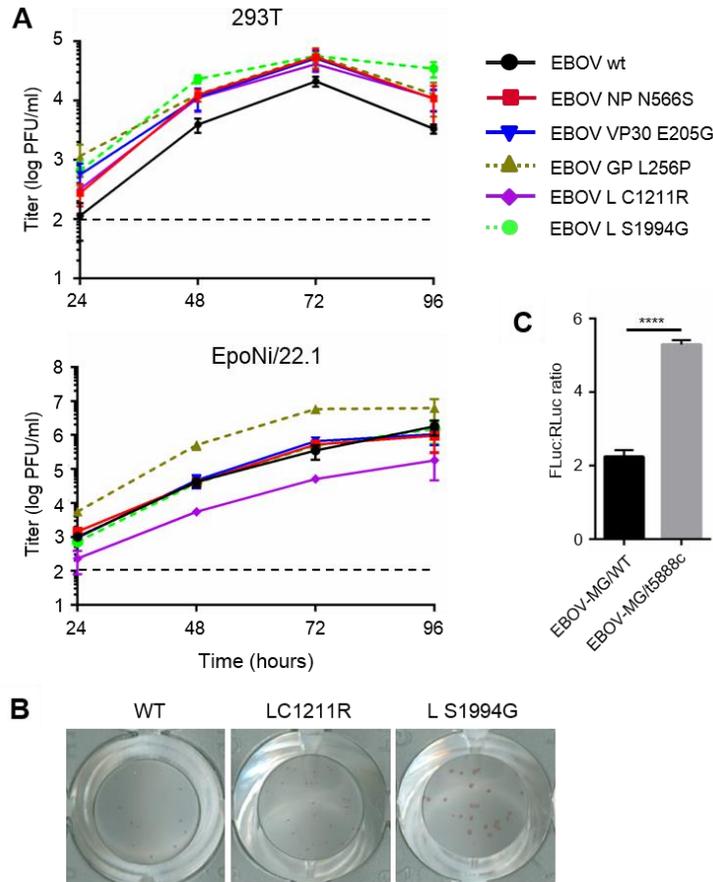


Figure 28: Phenotypic effects of selected mutants

(A) Multi-step growth kinetics of EBOV wt and mutant viruses in 293T and EpoNi/22.1 cells. Nearly confluent monolayers of 293T or EpoNi/22.1 cells were infected with a MOI 0.01 PFU/cell in triplicates, and supernatants collected at the indicated timepoints. Supernatants were titrated on Vero E6 cells. Replication kinetics of 293T-derived and EpoNi/22.1-derived mutants are shown with solid lines and dashed lines, respectively. The limit of detection (2 log PFU/mL) indicated by the dashed line. Relevant statistical comparisons (2-way ANOVA with Dunnett's correction) are discussed in text. (B) Plaque immunostaining was performed as described in Materials and Methods in 96 well plates, developed using AEC chromogen, and imaged. (C) VP40 gene-end region mutant reduces translation of the second ORF in a bicistronic minigenome. Ratio of firefly luciferase (first ORF) to Renilla luciferase (second ORF) reflects the efficiency of translation of the second ORF. **** $p < 0.0001$ (unpaired t-test with Welch's correction). No L and no minigenome controls were performed but are not shown. Firefly luciferase activity for controls was significantly lower than in full plasmid transfections ($p < 0.05$ by 1-way ANOVA with Tukey's post-hoc test).

To better understand the fitness relationships between the mutants and wild-type EBOV, competition assays were performed in both cell lines as shown in **Fig. 29A**. All mutants were observed to displace the wild-type virus under low MOI “competition” conditions in both cell lines (final passage proportion significantly greater than input, maximum $p < 0.01$). The exception was L C1211R in EpoNi/22.1 cells (**Fig. 29B**), repeating the results of the kinetics assays (passage three proportion was significantly less than input, $p < 0.001$). However, the kinetics of replacement were variable between viruses and cell lines. Although a non-parametric comparison between cell lines failed to detect a statistically significant difference between the two cell lines, the fitness of the 293T-origin mutant L C1211R was very cell line dependent, in contrast to L S1994G, which was detected in EpoNi/22.1 (**Fig. 29B**). While the EpoNi/22.1-origin GP L256P was more fit than wild-type in both cell lines, its kinetics of displacement appeared to be more rapid in EpoNi/22.1 cells, consistent with our replication kinetics results (**Fig. 29B**). However, this difference did not reach statistical significance. The fitness of mutations in polymerase accessory proteins showed little cell line dependency, and statistical significance was not reached in any instance. The only notable trend under high MOI “complementation” conditions, was the slow displacement of NP N566S with VP30 E205G (**Fig. 29B**) (final passage proportion was significantly greater than input, $p < 0.05$). Although possible in principle due to the large number of passages performed, it is unlikely that any significant degree of reversion occurred, primarily because (with a single exception) the mutations in question were more fit than wild type virus, making directional selection for reversion extremely unlikely. Ideally, watermarked viruses with noncoding mutations in proximity to the mutations would be prepared to entirely eliminate this possibility, but due to technical limitations associated with the EBOV reverse genetics system, this was not possible.

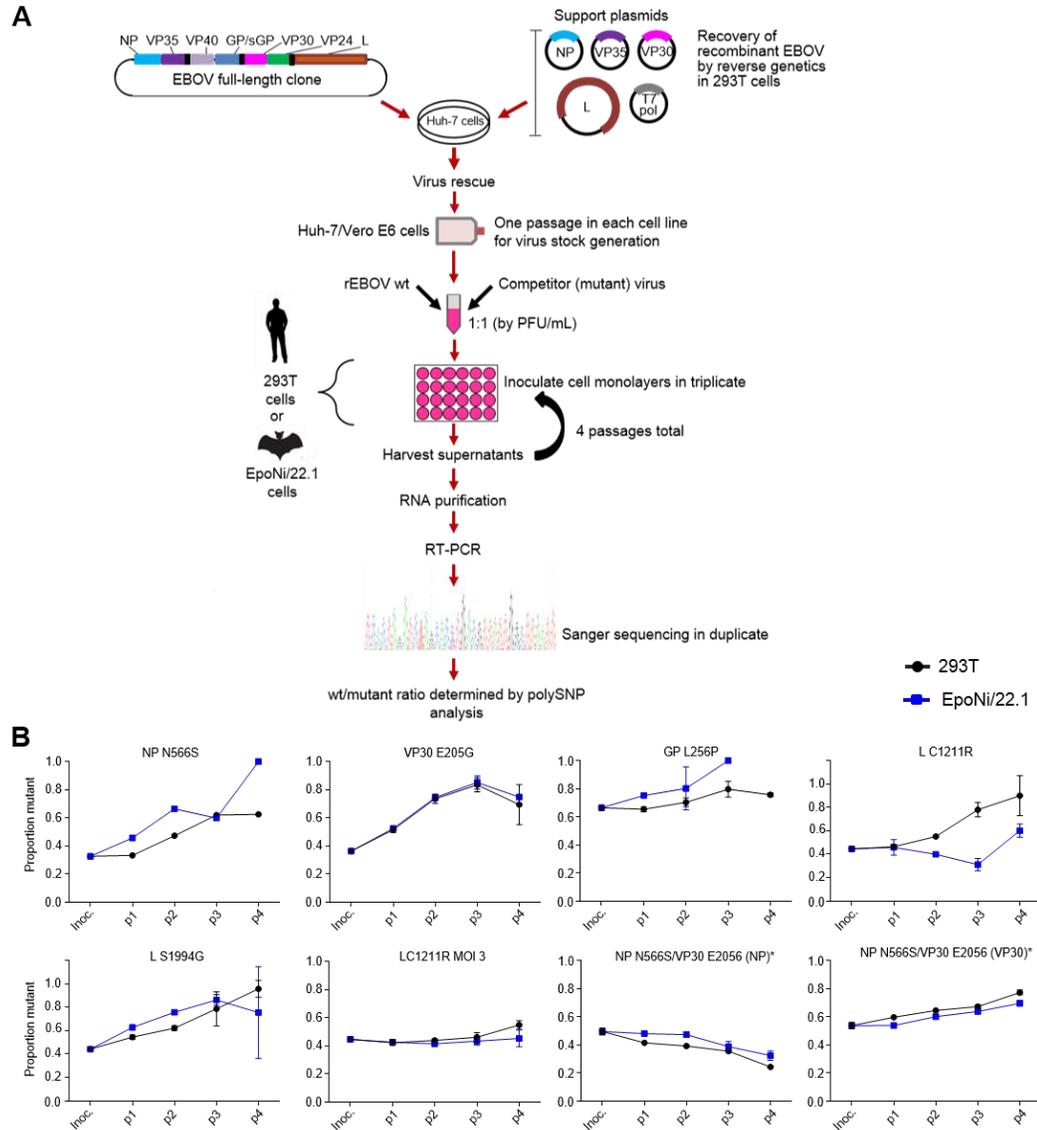


Figure 29: Competition and complementation assays

(A) Schematic representation of experimental design. (B) Data shown is representative of a single passaging experiment for each cell line, with three biological replicates for most competition/complementation assays. Two replicates were analyzed for the following: 293T EBOV wt/NP N566S, EBOV wt/L C1211R p4 competition assays, EpoNi/22.1 EBOV wt/NP N566S and EBOV GP L256P p3 competition assays, and EBOV wt/L C1211R p4 competition assays. One replicate was analyzed for the EpoNi/22.1 EBOV wt/VP30 E205G p4 competition assay. Additionally, due to the inability to generate an amplicon suitable for Sanger sequencing, EpoNi/22.1 EBOV wt/NP N566S and wt/GP L256P p4 competition assays were completely omitted from the analysis. Sequencing of RT-PCR amplicons was performed in duplicate, with the average of the two proportions being used. Error bars denote \pm SD. Asterisks denote complementation assays. Relevant statistical comparisons described in text.

DISCUSSION

The evolution of EBOV in EpoNi/22.1 cells during passaging was remarkably different from that observed in 293T cells. Although divergent evolutionary patterns are not unexpected, the degree and nature of the differences were notable. While the observed mutation rates were similar in both cell lines, the finding that the rate of G-to-A substitutions was significantly greater in EpoNi/22.1 is particularly important (**Fig. 19**). A potential explanation for this finding is RNA editing of the positive sense complementary RNA (cRNA) by a host factor. C to U mutations in the cRNA, such as those catalyzed by the APOBEC family, would produce G-to-A mutations in the resulting genomic RNA (248). Thus it would appear that rates of specific mutation types across the EBOV genome can be dramatically influenced by its cellular host.

More evidence for RNA-editing enzyme activity in EpoNi/22.1 cells was found in the GP gene. We observed a spike of high frequency A-to-G mutations in a region spanning the glycan cap and mucin-like domain of GP in EpoNi/22.1 passaged EBOV (**Fig. 22**). These regions are known to be favored targets of the humoral immune response during infection (424), and high frequency mutations here would be expected in the presence of such strong selective pressure. However, there were no antibodies present during our passaging, and both coding and non-coding mutations were identified. Additionally, the truly massive number of mutations present, and the rate at which these mutations accumulated after passage 4 in EpoNi/22.1 cells suggests the activity of a host RNA editing factor. Further investigation found that this pattern was likely the result of ADAR family RNA editing enzyme activity. ADAR editing of (-)ssRNA virus genomes is well-documented, and has been shown to have both proviral and antiviral effects (247, 387, 425). Although 293T-passaged viruses had A-to-G mutations similar to those identified in EpoNi/22.1 cells, the frequency was far lower, suggesting that ADAR activity is elevated in EpoNi/22.1 cells relative to 293T cells. Supporting this hypothesis, we found that

EpoNi/22.1 cells produce significantly more ADAR1 mRNA than 293T cells. As expression of ADAR is relatively consistent between all EBOV-relevant tissue types (including renal tissue), it is likely that this is true of other relevant cell types, barring any 293T-specific changes (426). Here a limitation of this study must be acknowledged. Obviously, it would be preferable to obtain direct evidence of ADAR activity by knocking out ADAR in EpoNi/22.1 via CRISPR/Cas9 or a similar technique. Unfortunately, there are no sequence data available for this or any closely related species, which would make designing guide RNAs and validating the resulting cell lines, including confirming the absence of serious off-target deletions, extremely difficult, and well beyond the scope of this study.

EBOV's MLD and GC are quite flexible, and in cell culture, appear to be at least partially expendable for GP-pseudotyped vesicular stomatitis virus (427). However, whether this applies to genuine virus is not well-established. We would thus expect editing of these regions to be well-tolerated. However, instead of simply being fitness-neutral, we found that some of the observed mutations may have been subject to positive selection in EpoNi/22.1 cells, as in some cases they closely approached fixation, (**Fig. 24B**). One mutant reconstituted using the reverse genetics system rapidly displaced wild-type virus in competition assays (**Fig. 29B**). The latter implies that clustering of these mutations on a single genome was not required for increased fitness. Thus, the rapid rise in A-to-G mutations in the EpoNi/22.1-derived populations was likely the result of enzymatic activity and selection.

Evidence of ADAR editing of EBOV genomes has been found in sequences obtained from human cases (132, 234, 421). Specifically, ADAR-like mutations in GP have been reported, although the activity was less specific than what we have observed (421). Our findings raise the possibility that there has been selective pressure to make GP a favorable target for ADAR. The nucleotide compositions of the mucin-like domain and glycan cap show a distinct enrichment for 'G' nucleotides, and depletion of 'A' nucleotides

(Fig. 30A) (428). This increased frequency of ‘G’ contributes to a uniformly high concentration of the 5'-AG-3' dinucleotide, part of ADAR's preferred 5'(C/A/U)-AG-3' target motif (423, 429) across the entire GC/MLD region, as compared to the other dinucleotides (shown as 3'-GA-5' in Fig. 30B). This specific region is one of the few in the entire genome where 5'-AG-3' is the most prevalent 5'-AX-3' dinucleotide. However, the frequency of 5'-GA-3' (a motif not preferred by ADAR) is also increased in this region. The notion of ADAR-driven evolution is not novel, and has previously been proposed in the context of Zika virus and rhabdovirus sigma, a member of the order *Mononegavirales* (370). However, our findings are novel in that we are reporting evidence for ADAR-driven evolution of portions of the envelope glycoprotein that are heavily targeted by the humoral immune response. Taking these facts into consideration, increased susceptibility of this region to ADAR editing may be a strategy to provide an intrinsic means of rapidly generating antibody escape mutants. This would be important in the context of actual *in vivo* infection, as bats do not appear to maintain a chronic circulating viremia when infected with MARV (210, 218). Several reports have found that potentially transmissible viremia is of relatively short duration, persisting no more than two weeks after inoculation (210). Bats also develop neutralizing antibodies in response to MARV within two to three weeks of inoculation (210). Although these factors clearly do not inhibit transmission in the wild, and bat to bat transmission of MARV has been experimentally demonstrated, the forest-dwelling bats thought to act as reservoirs for EBOV are not as gregarious as cave-dwelling *Rousettus aegyptiacus*. In particular, despite the fact that most of these forest-dwelling species do form social groups, these groups tend to be smaller than colonies of *Rousettus aegyptiacus* (214, 215). Given that repeated attempts to capture an actively infected bat have consistently failed (207), it must also be assumed that the incidence of EBOV infection in bat populations must be quite low, which is similar to what is observed in *Rousettus aegyptiacus* with MARV (209). It is therefore reasonable to conclude that when considering these issues, opportunities for the virus to transmit to an entirely naïve host

may be limited. Any mechanism that would increase the duration of a transmissible viral load would therefore be very advantageous. Furthermore, it is possible that a mechanism for rapidly generating escape mutants such as the one proposed here may allow the virus to temporarily instigate a productive infection in an individual with low anti-EBOV antibody titers before an anamnestic response could be mounted to clear the infection, particularly given the fact that bat antibodies are thought to have relatively low affinity due to the suspected inability of bats to undergo affinity maturation via somatic hypermutation (213). As such, bats would therefore be expected to mount a high avidity, but low affinity humoral response to infection. As a result, we would expect the diversity of the potential viral escape mutant pool to be of particular importance in bat infection.

Similar immune escape mechanisms are found in viruses associated with chronic infections, such as HIV, and viruses such as influenza that circulate in host populations with low turnover in which reinfection is a necessary for maintenance. Since EBOV does not chronically infect bats, HIV is not a relevant comparison. Influenza may seem to be a more relevant comparison to a first approximation. However, using the MARV/*Rousettus* system as a basis for comparison, we find that there is less analogy than there may appear to be. Influenza is associated with seasonal bursts of infection in which large numbers of individuals become infected, and eventually become immune to the circulating virus. As a result, influenza viruses undergo selection that alters their antigens (primarily HA and NA) to escape this immunity so as to re-infect these hosts (231). This process is known as antigenic drift, and represents a macroevolutionary trend readily identifiable from the ladder-like phylogeny of seasonal influenza viruses. By contrast, very small numbers of bats are infected with MARV at any given point (209), and the naïve population is refreshed twice per year by the biannual birthing cycle of *Rouesttus* bats, a trait shared with the bats implicated as potential EBOV reservoirs (214, 215, 220). This pattern would not favor the macroevolutionary antigenic drift mechanism of immune escape used by influenza virus,

but as discussed previously, may favor a microevolutionary form of immune escape. Given that escape mutations often revert in the absence of selective pressure (430), it is possible that such mutations would only rarely become fixed within the population at large, although GP is the most variable protein coding region of the EBOV genome (4).

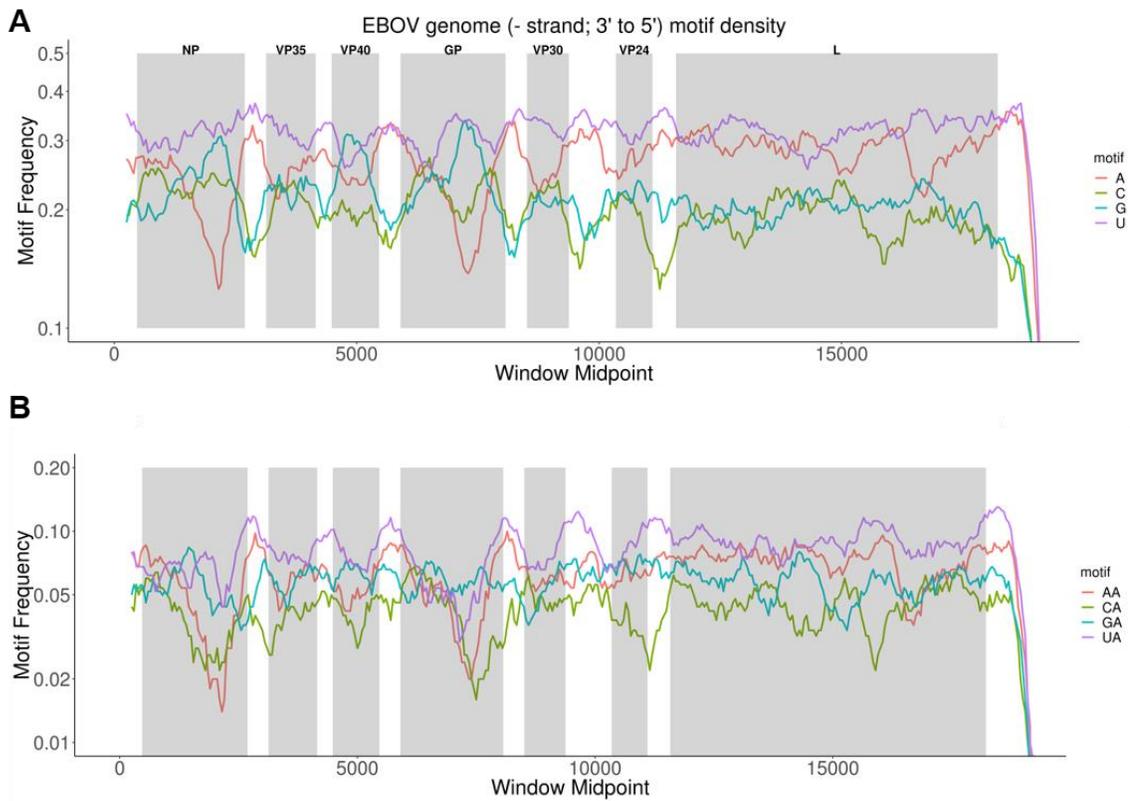


Figure 30: Nucleotide composition of the EBOV genome

Frequencies of individual nucleotides (**A**) or the dinucleotides 3'-XA-5' (**B**) are shown. The genome is displayed 3' to 5' with respect to the negative sense genomic strand. Motif frequencies were calculated with a window size of 500 nucleotides and a step size of 50 nucleotides.

It is important to emphasize that the extreme frequency of the mutants (but not the rate of mutation itself) observed in our data is the result of the artificial conditions imposed by our experimental design. Owing to the minimal purifying selection imposed by cell culture conditions, the ADAR mutants were able to accumulate without being removed by purifying selection or bottlenecks, permitting their detection. Under normal *in vivo* conditions, they would not normally accumulate. However, the increased rate of mutation

would increase the effective diversity of the GP gene within the mutant swarm, thereby providing a larger pool of potential escape mutants than would otherwise exist.

In addition to investigating the direct effects of host factors on viral evolution, we also describe changes in population structure that occurred as the result of the virus responding to the replicative environments imposed by the cell lines used. Broadly speaking, purifying selection of EBOV genomes appeared to be a predominant factor in EpoNi/22.1 cells. This is demonstrated by our observation of an increased rate of specific mutations and moderately higher average Shannon entropy in EpoNi/22.1 cells (**Figs. 19, 21**). Thus diversity was higher, but we identified fewer mutations that exhibited positive fitness compared to 293T cells (**Figs. 24, 20**). Moving beyond this global view, interpretation of our data must be conservative. Given the disparities in complexity, *in vitro* evolution cannot always be directly compared to *in vivo* evolution. However, patterns of mutations can be reasonably examined for the purposes of understanding aspects of the more general nature of viral evolution and adaptation in a given species. Therefore, our goal was to identify regions of the genome that appeared to be responding to the selective pressures imposed by each cell line. In doing this, our focus was on clusters of mutations rather than individual point mutations. This approach has been used previously in tandem with CirSeq in the context of poliovirus (413).

We identified a number of variant clusters associated with passaging in each cell line. In the ORFs, we identified one cluster in NP proximal to the VP30 binding domain, one in VP30 proximal to the NP binding domain, and one within the capping region of the L gene (**Figs. 20, 25**). An additional set of mutations was identified in the gene-end signal of VP40. A single representative mutant was selected from each identified cluster. The nature of the NP and VP30 mutants is particularly notable. NP N566S falls within a region reported to be important for interaction with host protein phosphatase 2 (PP2) (93), which is recruited to viral intracytoplasmic inclusion bodies by NP (93). PP2 participates in the

regulation of EBOV transcription via dephosphorylation of VP30, a requirement for transcription initiation (42, 44). VP30 E205G, meanwhile, is proximal to the NP interaction domain of VP30 (85), and would likely disrupt an α -helix, significantly disturbing the conformation of the binding domain. Given that NP/VP30 interaction is required for the dynamic phosphorylation of VP30 (93, 94), both of these mutants are predicted to affect EBOV transcription. A double mutant incorporating both VP30 E205G and its equivalent in NP (NP N566S) was nonviable and failed to rescue, implying that the mutants are not complementary, but may instead represent an example of convergent evolution. The relative lack of information regarding the structure and function of EBOV L polymerase makes discussion of the potential effects of the capping domain mutant L C1211R difficult, though the mutant and its associated variant cluster may affect the efficiency of mRNA production or on mRNA stability. By contrast, the likely implications of the VP40 gene-end signal mutations are more predictable. Disruption of this highly conserved sequence is almost certain to lead to the production of bicistronic mRNAs, as has been previously described (32, 95, 96). The second open reading frame in a bicistronic EBOV mRNA is translated at a drastically reduced frequency, therefore reducing the production of the resulting protein (96). Our findings with a representative mutant in the EBOV MG system are consistent with this. In EpoNi/22.1 cells, two clusters were identified, the putative ADAR cluster in GP, and another in the methyltransferase domain of L polymerase that likely has similar effects to the L cluster in 293T (**Figs. 25, 26**).

In competition and complementation assays performed with rescued mutant viruses, we found that most had a fitness advantage over wild-type virus in both cell lines, suggesting that they were genuinely under positive selection during our passage series (**Fig. 29B**). However, there were notable cell line-dependent differences in fitness for specific mutants. These phenotypes are likely the result of differences in the cellular microenvironments of the cell lines, and may be worthy of future exploration. The performance of L C1211R was somewhat unexpected. Though selection appeared to be

strong during passaging, it had a very marginal fitness advantage in 293T, the cell line in which it was identified, and was less fit than wild-type in EpoNi/22.1. The displacement of NP N566S by VP30 E205G during complementation assays in both cell lines was also unexpected because the fitness of both of these viruses relative to wild-type in kinetics assays appeared to be similar. It is likely that the difference is relatively small, and that as a result was only observable when the viruses were in direct competition. This hypothesis is bolstered by the low absolute frequency of the mutants in the sequencing data. Our failure to rescue the NP N566S/VP30 E205G double mutant suggests that these mutations are mutually exclusive, and that cellular co-infection would not be productive. With complementation impossible, if one virus had even a narrow competitive advantage over the other, it would eventually become the dominant genome.

Taken as a whole, our results validate the utility of CirSeq as a predictive tool for the identification of variants and variant clusters associated with increased fitness/adaptive evolution. Moreover, we were able to detect these mutants within a relatively short passage series. In both cell lines, variant clusters in the polymerase are associated with potentially adaptive evolution, but 293T cells produced more adaptive variant clusters. These findings are consistent with reports from the 2014-2016 West African epidemic which found that prolonged passaging in humans induced mutations in the NP, VP30, GP, and L genes (233). Two caveats must be noted. Firstly, the NP mutation identified in the West African epidemic did not increase fitness *in vitro*, and did not appear to be under true positive selection during the epidemic (233). Secondly GP mutations were not identified in human cells during our passage series. Despite this, the identification of potentially adaptive mutations in the VP30 and L genes in this report, and in circulation during a human outbreak points to a role for these two proteins in adaptation of the virus to humans. By extension, and considering the roles of these proteins in the life cycle of the virus, this suggests that elements of their roles in replication and transcription are not identical, and/or experience divergent selective pressures in bats. VP30 in particular is interesting in this

respect, for two reasons. Firstly, of the two proteins, it has the most direct interaction with the host in its function as a transcriptional activator, as it must be dynamically phosphorylated by host phosphatases (44, 93). The kinetics of interactions with host phosphatases are an obvious point at which selective pressures imposed by a new host could lead to adaptive mutation. Related to this, the NP mutant identified in this report falls within the domain that interacts with host phosphatases (93), suggesting a similar conclusion. Secondly, VP30 also a role as a VSR that was discussed previously (39), and the importance and mechanism of this function may differ in bat and human cells.

Although it is a powerful tool for understanding the dynamics of RNA virus evolution within cell culture systems, CirSeq does have a number of serious limitations that are relevant to my work. Due to the requirement for a large quantity of very pure viral RNA, CirSeq is almost exclusively restricted to use within cell culture systems. In addition, effective genome read lengths must be short, to accommodate the requirement for multiple copies of a given genome fragment within a single sequencing read. This makes assessment of linkage between mutants separated by any significant distance difficult or impossible. However, by using CirSeq in tandem with sequencing chemistries that permit longer read lengths, this limitation can be ameliorated.

We have identified a number of key differences in the evolution of EBOV in a human cell line relative to a cell line derived from a close relative of a potential reservoir host. By comparison to the dramatic differences in replicative and fitness environments faced by arboviruses in their host/vector lifecycles, the cell lines used in this study are not extraordinarily divergent as both are of mammalian origin. In this light, our identification of a number of meaningful differences in the short-term evolution of the virus in these cell lines is in and of itself remarkable. We have presented evidence suggesting that RNA editing enzymes play a greater role in the replication and evolution of EBOV in bat cells. As a result of our findings, we propose that ADAR, a host RNA editing enzyme, may have a role in the evolution of the virus in at least one of the cell lines used. We further propose

that this role is intimately related with the biology of the likely bat reservoir of EBOV. Furthermore, we identified regions within the viral genome associated with potentially adaptive evolution resulting from passaging in these cell lines, and characterized selected mutations from these regions. Some of the mutations are in genes previously identified as being associated with EBOV adaptation to humans (233). In addition to identifying potential adaptive strategies, this also works in reverse to identify viral proteins that may have roles in bat infection that diverge from those that they are known to have in infection of humans. Curiously, many of the mutants identified in variant clusters associated with passaging in these cell lines displayed similar, but not identical fitness in each cell line, suggesting that relatively minor differences in selective pressures could be responsible for the evolutionary divergence we observed. Overall, our findings would suggest that evolution of EBOV in EpoNi/22.1 cells, and potentially by extension in bats, is driven to a significant degree by host factors acting on the genome. By contrast, EBOV evolution in 293T cells appears to be adaptive, with emphasis on regulation of transcription and transcript stability, as evidenced by variant clusters found within regions of the genome associated with these functions. This pattern fits expectations for a virus that uses bats as a natural reservoir, as evolution in the reservoir host would likely be drift-driven, while evolution in an incidental host would be more likely to favor positive selection for adaptation (231, 233).

CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

THE UNIQUE BIOLOGY OF BATS

Bats have a very unique place in the field of reservoir biology. They are either the known or suspected reservoir host of a staggering number of viruses across a large number of taxa. In other cases, they have played some role in the evolutionary history of a virus, such as SARS and MERS (407). This astounding ability to harbor viruses, often without obvious clinical illness, has been a topic of a great deal of interest for some time. Here, two studies aimed at understanding parts of this fascinating biology have been presented. Given how little is known about the biology of bats in this context, the biology of human EBOV infection was used as a better-characterized comparison in both cases.

STUDY 1: VIRAL SMALL RNAs

The first study focused on the small RNA biology of filovirus infections in bats, and found that although there was no evidence for an antiviral RNAi response to infection in bats or humans, filovirus infection in both species results in the production of small noncoding RNAs from viral mRNAs. The vncRNAs produced by EBOV were not found to have any association with the host microRNA machinery, and lacked any RNAi function. However, in *Rousettus aegyptiacus* cells some evidence was found that suggests that the digestion of viral mRNAs to produce vncRNAs may have some antiviral effect. Overall, it seems that in the case of EBOV, these vncRNAs are either the vestigial remnants of some now lost host immune suppression mechanism, or are spandrels resulting from some mechanism likely associated with transcription.

STUDY 2: EXPERIMENTAL EVOLUTION

The second study was designed to use experimental evolution to reveal aspects of the biology of EBOV that might otherwise be difficult to identify. Using a unique ultra-deep sequencing approach, this study characterized the changes in population structure that occurred over a prolonged passaged series in bat and human cells. Two major conclusions can be drawn directly from the data obtained. Firstly, evolution in bat cells was dominated by drift and selective pressure resulting from host factors, such as RNA editing enzymes. More specifically, the data establish a strong circumstantial case implicating the host RNA editing enzyme ADAR in editing of the EBOV genome. Given the location of the focused hypermutation attributed to ADAR in a flexible region of the spike glycoprotein gene, I believe that this represents a mechanism for the rapid production of antibody escape mutants that may facilitate the transmission of the virus in low-density populations with a low incidence of infection by extending the infectious period. Secondly, evolution in human cells was adaptive. Two genes found to accumulate potentially adaptive mutations, VP30 and L were also implicated in human adaptation during the West African epidemic. As such, it is likely that these two genes represent major foci of adaptive evolution during human passaging of EBOV. Working backward, I believe that it is therefore also reasonable to extrapolate that the functions of these genes, may diverge in bats and humans. In the case of VP30, it seems likely given my findings that its dynamic phosphorylation may be affected by the identity of the host, as this is dependent upon host phosphatases.

THOUGHTS ON METHODOLOGY

Much of the published literature on the topic of bat/virus interactions focuses on a single gene or pathway (216, 431, 432). This approach has the advantage of providing specific and well-defined results, but comes with a number of pitfalls. In particular, two issues are common. The first is over interpretation of the data obtained, often seen as an attempt to apply findings in one species of bat to bats as a whole. This is an issue of some

importance, as multiple reports have found that many aspects of seemingly fundamental immune mechanisms are, in fact, not broadly shared between bats (431, 433-436). The second issue is more fundamental, and relates to the design of these experiments. By focusing on a single gene or pathway, it is very easy to miss or ignore the systems implications of findings, or to obtain data of limited value due to the presence of significant and sometimes unknown/unidentified confounding factors. Narrowly focusing on a specific aspect of bat biology and/or using biased approaches such as qPCR can also cause one to miss things that would have otherwise been identified by less focused approaches. Systems-based approaches, particularly minimally biased ones that permit *de novo* discovery such as next generation sequencing can minimize or avoid these problems. In the case of the studies presented here, next-generation sequencing enabled the identification of a novel class of viral small RNAs produced by EBOV, and the detailed characterization of the evolution of the virus a likely reservoir host, and an incidental host.

Future directions

STUDY 1: VIRAL SMALL RNAs

Identification of the biogenic origin of EBOV vncRNAs

Although knockdown of CPSF3L was found to have an effect on the production of the vncRNAs, this finding is not sufficient to definitively link it to their biogenesis. One approach to further exploring the association of CPSF3L with the vncRNAs would be to perform an RNA immunoprecipitation for CPSF3L, and use miRNA-specific RT-qPCR to probe for the EBOV vncRNAs in the resulting RNA pool. Identifying and knocking down functional partners of CPSF3L both with and without simultaneous knockdown of CPSF3L would provide a means of identifying any potential cofactors that may be required.

Characterization of the MARV vncRNAs

The most obvious future direction in this study is the characterization of the MARV vncRNAs, which were largely ignored in the process of attempting to validate the miRNA status of the EBOV vncRNAs. It is possible that the MARV vncRNAs do in fact function as miRNAs, unlike their EBOV counterparts. As such, following an experimental program similar to that used to characterize the EBOV vncRNAs, including investigation of Dicer and Argonaute association, would be well worth the time and effort. Given that one of the hypotheses of the evolutionary origin of the EBOV vncRNAs posits that they are a lost mechanism of host immunosuppression, finding a similar mechanism active in a basal clade (MARV is phylogenetically basal to EBOV) would go a long way toward bolstering this idea.

STUDY 2: EXPERIMENTAL EVOLUTION

Experimental evolution in other bat cell lines

As has already been discussed extensively, *Epomops* bats are not the only suspected reservoirs of EBOV, and the ecology of EBOV is not defined to include only a single bat species. It is possible that another bat species, such as *Hypsignathus monstrosus* is a reservoir for the virus. It is also possible that another bat species with more frequent contact with humans, such as *Mops condylurus* acts as a sort of bridging reservoir between epomophorine forest bats, where a cycle of transmission between these bats occurs somewhat analogous to the sylvatic cycle of yellow fever, and humans, where human to human transmission establishes an urban cycle of transmission (**Illustration 6**). Cell lines from both of these species exist that are tissue-matched to those used in the study presented here. An identical study conducted using these cell lines would both allow for a comparison of evolution between bat species, potentially expanding the scope of the finding of ADAR hypermutation, and/or identifying mutations that may occur in a potential bridging

reservoir like *Mops condlyurus*. Additionally, an identical experiment with MARV in *Rousettus aegyptiacus* cells would be a valuable contribution to the field of bat virology.

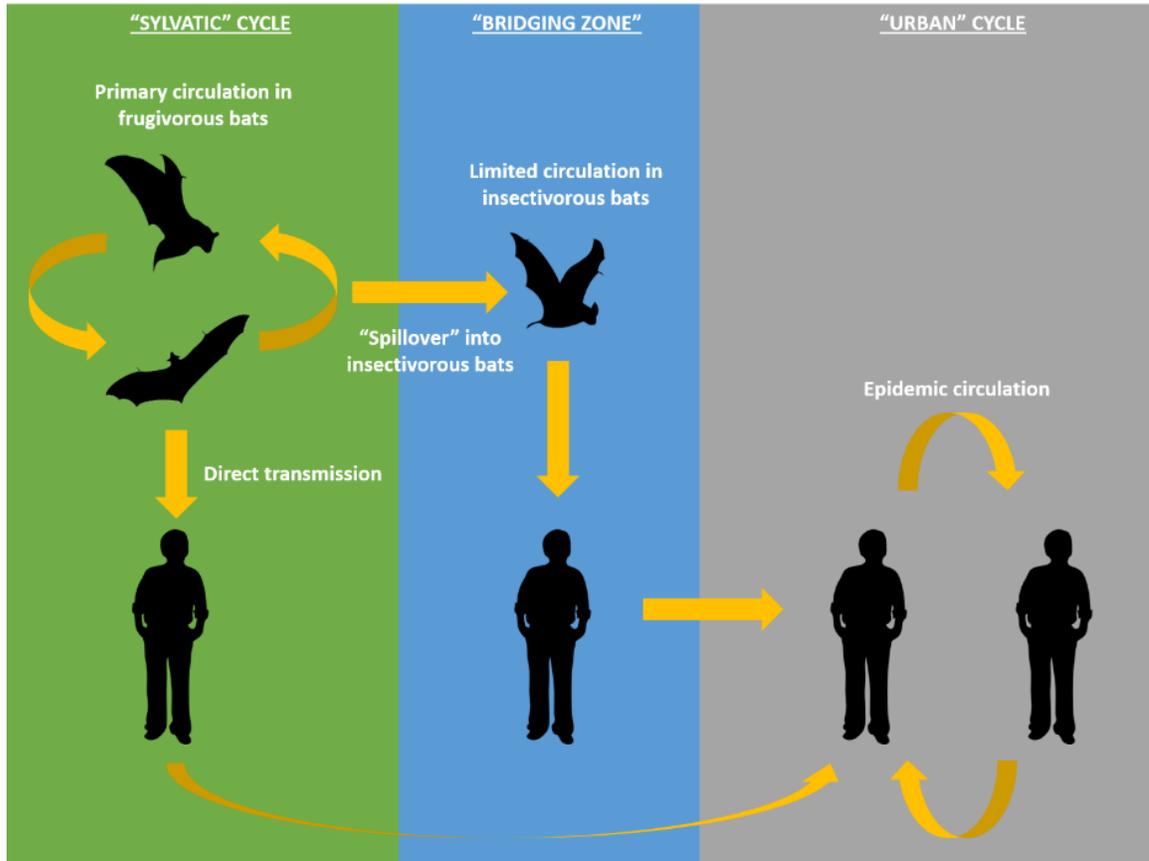


Illustration 6: Hypothetical multispecies ecology of EBOV in bats

Further validation of ADAR editing activity

Perhaps the most obvious future direction of this entire work is the further validation of ADAR's apparent editing of the GP mucin-like domain and glycan cap regions. Although it was beyond the scope of the original study, the most obvious route would be to sequence the transcriptome of an *Epomops* bat and use this to design guide RNAs for use in generation of a CRISPR/Cas9 knockout cell line. Although there are some concerns about the behavior of such a cell line, as ADAR has a role in reducing cellular immune responses to endogenous dsRNA (418), passaging the virus in such a cell line in

a manner identical to that performed in the study would provide a definitive answer as to the association of ADAR with the mutations observed. If they are not present, it can be concluded that they are almost certainly the result of ADAR editing. If they appear in spite of ADAR knockout, some other mechanism would have to be identified to explain the phenomenon.

Validation of escape mutant generation via ADAR hypermutation

One of the major points of speculation resulting from the ADAR data involved the possibility of ADAR hypermutation providing a rapid means of generating antibody escape mutants. One way to explore this possibility would be to produce virus in 293T cells, EpoNi/22.1 cells, and the aforementioned ADAR knockout EpoNi/22.1 cell line. Viruses would be produced both in the presence and in the absence of neutralizing human and ideally bat (likely from inoculation of *Rousettus aegyptiacus* with EBOV) immune sera. These viruses could then be used in a neutralization assay with human and bat immune sera. If neutralization of virus produced in EpoNi/22.1 requires a higher antibody titer than that required to neutralize virus produced in 293T or the knockout cell line, it could be reasonably concluded that ADAR hypermutation contributes to increased resistance to the humoral immune response. A simpler approach would be to test the neutralization resistance of virus produced in a variety of human cell lines to that of viruses produced in bat cell lines. This would carry with it the advantage of demonstrating broader applicability of the phenomenon.

Appendix A – Detailed sequencing data for vncRNA chapter

Quality data

Cell line	Virus	Timept.	Rep.	Index	Reads PF (Lane 1)	Reads PF (Lane 2)	Total Reads PF	% ≥Q30 (Lane 1)	% ≥Q30 (Lane 2)	Mean % ≥Q30 Lanes 1 & 2	Mean Quality Lane 1	Mean Quality Lane 2
HepG2	EBOV wt-eGFP	12hpi	1	1	7,379,261	7,563,230	14,942,491	95.95	95.75	95.85	38.15	38.08
HepG2	EBOV wt-eGFP	12hpi	2	2	7,413,331	7,634,890	15,048,221	96.1	95.94	96.02	38.18	38.11
HepG2	EBOV wt-eGFP	12hpi	3	3	7,370,205	7,507,346	14,877,551	95.92	95.72	95.82	38.14	38.07
HepG2	EBOV wt-eGFP	24hpi	1	4	7,617,757	7,820,670	15,438,427	95.87	95.67	95.77	38.13	38.06
HepG2	EBOV wt-eGFP	24hpi	2	5	8,993,781	9,258,236	18,252,017	96.21	96.07	96.14	38.24	38.19
HepG2	EBOV wt-eGFP	24hpi	3	6	9,014,626	9,267,487	18,282,113	96.2	96.05	96.125	38.22	38.16
HepG2	EBOV VP35 R312A-eGFP	12hpi	1	7	7,734,371	7,950,619	15,684,990	96.62	96.55	96.585	38.34	38.3
HepG2	EBOV VP35 R312A-eGFP	12hpi	2	8	6,984,260	7,178,179	14,162,439	96.87	96.81	96.84	38.43	38.39
HepG2	EBOV VP35 R312A-eGFP	12hpi	3	9	6,674,832	6,832,057	13,506,889	96.67	96.59	96.63	38.36	38.32
HepG2	EBOV VP35 R312A-eGFP	24hpi	1	10	6,005,100	6,182,405	12,187,505	96.78	96.73	96.755	38.41	38.38
HepG2	EBOV VP35 R312A-eGFP	24hpi	2	11	7,140,274	7,292,363	14,432,637	96.94	96.87	96.905	38.45	38.41
HepG2	EBOV VP35 R312A-eGFP	24hpi	3	12	7,392,357	7,580,979	14,973,336	96.83	96.76	96.795	38.42	38.38
HepG2	Mock Infected	12hpi	1	7	8,434,184	8,673,215	17,107,399	95.77	95.6	95.685	38.09	38.02
HepG2	Mock Infected	12hpi	2	8	7,141,132	7,364,544	14,505,676	96.11	95.96	96.035	38.2	38.14
HepG2	Mock Infected	12hpi	3	9	7,636,244	7,831,128	15,467,372	95.98	95.8	95.89	38.16	38.1
HepG2	Mock Infected	24hpi	1	10	4,933,595	5,089,516	10,023,111	96.09	95.93	96.01	38.21	38.14
HepG2	Mock Infected	24hpi	2	11	9,321,044	9,481,658	18,802,702	96.14	95.96	96.05	38.2	38.13
HepG2	Mock Infected	24hpi	3	12	8,113,795	8,349,501	16,463,296	96.09	95.94	96.015	38.2	38.14
EpoNi/22.1	EBOV wt-eGFP	12hpi	1	13	6,919,337	7,099,144	14018481	96.08	95.89	95.985	38.2	38.13
EpoNi/22.1	EBOV wt-eGFP	12hpi	2	14	8,114,970	8,341,881	16456851	96.11	95.94	96.025	38.2	38.14
EpoNi/22.1	EBOV wt-eGFP	12hpi	3	15	7,935,109	8,140,277	16075386	95.76	95.58	95.67	38.16	38.09
EpoNi/22.1	EBOV wt-eGFP	24hpi	1	16	9,094,849	9,355,624	18450473	95.89	95.73	95.81	38.13	38.06
EpoNi/22.1	EBOV wt-eGFP	24hpi	2	17	7,521,135	7,670,909	15192044	96.06	95.9	95.98	38.2	38.13
EpoNi/22.1	EBOV wt-eGFP	24hpi	3	18	6,795,573	6,931,934	13727507	95.99	95.82	95.905	38.17	38.1
EpoNi/22.1	EBOV VP35 R312A-eGFP	12hpi	1	19	6,963,520	7,112,501	14076021	96.37	96.3	96.335	38.29	38.25
EpoNi/22.1	EBOV VP35 R312A-eGFP	12hpi	2	20	5,608,771	5,713,733	11322504	96.56	96.48	96.52	38.34	38.3
EpoNi/22.1	EBOV VP35 R312A-eGFP	12hpi	3	21	6,879,139	7,032,806	13911945	96.77	96.71	96.74	38.4	38.37
EpoNi/22.1	EBOV VP35 R312A-eGFP	24hpi	1	22	6,401,439	6,535,306	12936745	96.33	96.27	96.3	38.27	38.23
EpoNi/22.1	EBOV VP35 R312A-eGFP	24hpi	2	23	6,510,379	6,631,972	13142351	96.74	96.68	96.71	38.4	38.37
EpoNi/22.1	EBOV VP35 R312A-eGFP	24hpi	3	24	6,040,227	6,148,198	12188425	96.4	96.32	96.36	38.29	38.24
EpoNi/22.1	Mock Infected	12hpi	1	19	6,584,325	6,734,732	13319057	95.9	95.7	95.8	38.15	38.08
EpoNi/22.1	Mock Infected	12hpi	2	20	5,763,730	5,877,128	11640858	95.98	95.78	95.88	38.16	38.08
EpoNi/22.1	Mock Infected	12hpi	3	21	4,738,758	4,834,260	9573018	96.05	95.89	95.97	38.18	38.12
EpoNi/22.1	Mock Infected	24hpi	1	22	7,671,408	7,833,974	15505382	95.91	95.73	95.82	38.14	38.07
EpoNi/22.1	Mock Infected	24hpi	2	23	7,478,394	7,617,908	15096302	96.25	96.1	96.175	38.26	38.2
EpoNi/22.1	Mock Infected	24hpi	3	24	6,874,931	7,012,612	13887543	95.95	95.77	95.86	38.15	38.08
R06	EBOV wt-eGFP	12hpi	1	13	3,271,391	3,392,444	6,663,835	97.03	96.88	96.955	38.44	38.38
R06	EBOV wt-eGFP	12hpi	2	14	7,818,718	8,104,437	15,923,155	97.14	97.02	97.08	38.47	38.42
R06	EBOV wt-eGFP	12hpi	3	15	3,429,697	3,556,278	6,985,975	96.89	96.72	96.805	38.44	38.38
R06	EBOV wt-eGFP	24hpi	1	16	11,873,949	12,177,676	24,051,625	97.03	96.89	96.96	38.44	38.38
R06	EBOV wt-eGFP	24hpi	2	12	3,177,149	3,254,031	6,431,180	97.09	96.94	97.015	38.46	38.4
R06	EBOV wt-eGFP	24hpi	3	18	7,193,349	7,327,943	14,521,292	97.06	96.89	96.975	38.45	38.38
R06	EBOV VP35 R312A-eGFP	12hpi	1	19	8,571,240	8,743,537	17,314,777	96.94	96.78	96.86	38.42	38.36
R06	EBOV VP35 R312A-eGFP	12hpi	2	20	3,845,871	3,921,154	7,767,025	97.07	96.9	96.985	38.45	38.39
R06	EBOV VP35 R312A-eGFP	12hpi	3	21	5,753,138	5,887,032	11,640,170	97.15	97.02	97.085	38.48	38.43
R06	EBOV VP35 R312A-eGFP	24hpi	1	22	5,362,408	5,477,959	10,840,367	96.88	96.73	96.805	38.39	38.33
R06	EBOV VP35 R312A-eGFP	24hpi	2	23	3,985,410	4,058,889	8,044,299	97.21	97.08	97.145	38.5	38.45

R06	EBOV VP35 R312A-eGFP	24hpi	3	24	10,133,667	10,332,841	20,466,508	96.95	96.8	96.875	38.41	38.35
R06	Mock Infected	12hpi	1	37	13,230,949	13,544,880	26,775,829	97.07	96.93	97	38.44	38.38
R06	Mock Infected	12hpi	2	38	4,647,495	4,836,270	9,483,765	97.21	97.12	97.165	38.51	38.47
R06	Mock Infected	12hpi	3	39	3,095,957	3,172,894	6,268,851	96.91	96.75	96.83	38.4	38.34
R06	Mock Infected	24hpi	1	40	4,605,128	4,736,448	9,341,576	97.01	96.85	96.93	38.43	38.37
R06	Mock Infected	24hpi	2	41	4,448,067	4,577,722	9,025,789	97.22	97.11	97.165	38.5	38.46
R06	Mock Infected	24hpi	3	42	4,286,698	4,396,418	8,683,116	97.06	96.92	96.99	38.46	38.41
HepG2	rMARV wt-eGFP	12hpi	1	43	6,879,511	7,036,931	13,916,442	97.08	96.92	97	38.45	38.39
HepG2	rMARV wt-eGFP	12hpi	2	44	11,471,605	11,768,032	23,239,637	96.83	96.66	96.745	38.38	38.31
HepG2	rMARV wt-eGFP	12hpi	3	45	11,862,606	12,154,183	24,016,789	96.84	96.65	96.745	38.38	38.31
HepG2	rMARV wt-eGFP	24hpi	1	46	8,193,841	8,399,568	16,593,409	96.98	96.81	96.895	38.42	38.35
HepG2	rMARV wt-eGFP	24hpi	2	47	9,380,600	9,590,850	18,971,450	97.21	97.06	97.135	38.49	38.43
HepG2	rMARV wt-eGFP	24hpi	3	48	8,068,793	8,235,961	16,304,754	97.2	97.04	97.12	38.49	38.42
HepG2	rMARV VP35 R301A-eGFP	12hpi	1	1	7,898,495	7,975,798	15,874,293	95.64	95.31	95.475	38.1	38.02
HepG2	rMARV VP35 R301A-eGFP	12hpi	2	2	5,072,258	5,137,319	10,209,577	95.76	95.44	95.6	38.12	38.04
HepG2	rMARV VP35 R301A-eGFP	12hpi	3	3	8,617,496	8,628,495	17,245,991	95.77	95.44	95.605	38.15	38.06
HepG2	rMARV VP35 R301A-eGFP	24hpi	1	4	7,446,964	7,548,318	14,995,282	95.68	95.35	95.515	38.13	38.05
HepG2	rMARV VP35 R301A-eGFP	24hpi	2	5	5,900,952	5,972,393	11,873,345	95.86	95.54	95.7	38.19	38.11
HepG2	rMARV VP35 R301A-eGFP	24hpi	3	6	5,958,021	6,022,612	11,980,633	95.87	95.56	95.715	38.17	38.09
HepG2	Mock Infected	12hpi	1	7	9,979,465	10,136,017	20,115,482	95.67	95.35	95.51	38.11	38.03
HepG2	Mock Infected	12hpi	2	8	5,762,858	5,852,271	11,615,129	95.83	95.52	95.675	38.17	38.09
HepG2	Mock Infected	12hpi	3	9	9,948,310	10,058,825	20,007,135	95.72	95.4	95.56	38.13	38.05
HepG2	Mock Infected	24hpi	1	10	6,984,613	7,098,743	14,083,356	95.86	95.55	95.705	38.19	38.11
HepG2	Mock Infected	24hpi	2	11	5,165,903	5,200,865	10,366,768	96.12	95.82	95.97	38.24	38.17
HepG2	Mock Infected	24hpi	3	12	5,770,921	5,856,945	11,627,866	95.93	95.62	95.775	38.2	38.12
EpoNi/22.1	rMARV wt-eGFP	12hpi	1	7	7,995,755	8,200,371	16,196,126	96.67	96.63	96.65	38.35	38.32
EpoNi/22.1	rMARV wt-eGFP	12hpi	2	8	7,417,696	7,606,724	15,024,420	96.84	96.82	96.83	38.42	38.39
EpoNi/22.1	rMARV wt-eGFP	12hpi	3	9	6,160,075	6,296,083	12,456,158	96.67	96.64	96.655	38.37	38.34
EpoNi/22.1	rMARV wt-eGFP	24hpi	1	10	7,114,078	7,298,434	14,412,512	96.74	96.73	96.735	38.4	38.37
EpoNi/22.1	rMARV wt-eGFP	24hpi	2	11	6,833,901	6,953,460	13,787,361	96.91	96.87	96.89	38.44	38.4
EpoNi/22.1	rMARV wt-eGFP	24hpi	3	12	6,829,984	6,995,783	13,825,767	96.78	96.76	96.77	38.4	38.38
EpoNi/22.1	rMARV VP35 R301A-eGFP	12hpi	1	13	6,771,135	6,932,140	13,703,275	96.66	96.62	96.64	38.36	38.33
EpoNi/22.1	rMARV VP35 R301A-eGFP	12hpi	2	14	5,460,674	5,592,177	11,052,851	96.7	96.68	96.69	38.37	38.34
EpoNi/22.1	rMARV VP35 R301A-eGFP	12hpi	3	15	6,659,233	6,824,083	13,483,316	96.55	96.52	96.535	38.36	38.34
EpoNi/22.1	rMARV VP35 R301A-eGFP	24hpi	1	16	7,456,776	7,636,095	15,092,871	96.66	96.63	96.645	38.35	38.32
EpoNi/22.1	rMARV VP35 R301A-eGFP	24hpi	2	17	6,641,418	6,780,737	13,422,155	96.73	96.71	96.72	38.39	38.36
EpoNi/22.1	rMARV VP35 R301A-eGFP	24hpi	3	18	6,730,511	6,867,786	13,598,297	96.68	96.64	96.66	38.36	38.33
EpoNi/22.1	Mock Infected	12hpi	1	19	5,371,613	5,492,690	10,864,303	96.7	96.65	96.675	38.38	38.34
EpoNi/22.1	Mock Infected	12hpi	2	20	6,910,944	7,042,757	13,953,701	96.74	96.7	96.72	38.38	38.35
EpoNi/22.1	Mock Infected	12hpi	3	21	6,138,866	6,271,788	12,410,654	96.82	96.79	96.805	38.4	38.38
EpoNi/22.1	Mock Infected	24hpi	1	22	6,469,763	6,625,091	13,094,854	96.57	96.53	96.55	38.32	38.29
EpoNi/22.1	Mock Infected	24hpi	2	23	5,503,917	5,603,952	11,107,869	96.82	96.8	96.81	38.4	38.37
EpoNi/22.1	Mock Infected	24hpi	3	24	6,999,464	7,137,176	14,136,640	96.65	96.61	96.63	38.35	38.32
RO6	rMARV wt-eGFP	12hpi	1	13	5,195,476	5,268,931	10,464,407	95.83	95.51	95.67	38.17	38.09
RO6	rMARV wt-eGFP	12hpi	2	14	5,683,921	5,767,424	11,451,345	96.04	95.75	95.895	38.23	38.16
RO6	rMARV wt-eGFP	12hpi	3	15	5,738,606	5,822,014	11,560,620	95.74	95.43	95.585	38.19	38.11
RO6	rMARV wt-eGFP	24hpi	1	16	8,477,413	8,599,880	17,077,293	95.81	95.51	95.66	38.15	38.08
RO6	rMARV wt-eGFP	24hpi	2	17	5,033,189	5,088,725	10,121,914	96.01	95.71	95.86	38.23	38.15
RO6	rMARV wt-eGFP	24hpi	3	18	7,089,878	7,158,356	14,248,234	95.96	95.64	95.8	38.21	38.13
RO6	rMARV VP35 R301A-eGFP	12hpi	1	19	6,992,844	7,086,787	14,079,631	95.86	95.56	95.71	38.19	38.11
RO6	rMARV VP35 R301A-eGFP	12hpi	2	20	12,582,846	12,685,092	25,267,938	95.9	95.59	95.745	38.19	38.11
RO6	rMARV VP35 R301A-eGFP	12hpi	3	21	8,081,183	8,185,654	16,266,837	96.02	95.74	95.88	38.23	38.16
RO6	rMARV VP35 R301A-eGFP	24hpi	1	22	7,769,052	7,851,358	15,620,410	95.65	95.33	95.49	38.11	38.03
RO6	rMARV VP35 R301A-eGFP	24hpi	2	23	10,867,938	10,959,686	21,827,624	96.09	95.8	95.945	38.25	38.18
RO6	rMARV VP35 R301A-eGFP	24hpi	3	24	6,476,412	6,522,259	12,998,671	95.75	95.43	95.59	38.13	38.05
RO6	Mock Infected	12hpi	1	1	8,245,933	8,441,397	16,687,330	96.43	96.41	96.42	38.3	38.27
RO6	Mock Infected	12hpi	2	2	6,086,270	6,233,509	12,319,779	96.54	96.52	96.53	38.3	38.27
RO6	Mock Infected	12hpi	3	3	7,307,400	7,415,916	14,723,316	96.6	96.54	96.57	38.34	38.3
RO6	Mock Infected	24hpi	1	4	7,054,101	7,230,245	14,284,346	96.62	96.59	96.605	38.36	38.33
RO6	Mock Infected	24hpi	2	5	7,305,619	7,466,761	14,772,380	96.69	96.68	96.685	38.38	38.35
RO6	Mock Infected	24hpi	3	6	6,232,923	6,377,488	12,610,411	96.7	96.68	96.69	38.36	38.33

Aligned reads data

Cell line	Virus	Timept.	Rep.	Index	Trimmed Reads (19-32 nts only)	Virus aligned reads (19-32 nts)	% Total 19-32 nt Trimmed Rds.	Total Aligned reads
HepG2	EBOV wt-eGFP	12hpi	1	1	9133722	1095	0.011988541	1169
HepG2	EBOV wt-eGFP	12hpi	2	2	10462950	1362	0.013017361	1466
HepG2	EBOV wt-eGFP	12hpi	3	3	8994747	1273	0.014152705	1351
HepG2	EBOV wt-eGFP	24hpi	1	4	9178588	20339	0.221591818	20837
HepG2	EBOV wt-eGFP	24hpi	2	5	11578790	28710	0.24795337	29228
HepG2	EBOV wt-eGFP	24hpi	3	6	11145035	26811	0.24056452	27331
HepG2	EBOV VP35 R312A-eGFP	12hpi	1	7	9994957	944	0.009444763	993
HepG2	EBOV VP35 R312A-eGFP	12hpi	2	8	9099491	751	0.008253209	802
HepG2	EBOV VP35 R312A-eGFP	12hpi	3	9	7826047	795	0.010158385	831
HepG2	EBOV VP35 R312A-eGFP	24hpi	1	10	6073391	12917	0.212681844	13188
HepG2	EBOV VP35 R312A-eGFP	24hpi	2	11	9156825	23243	0.253832524	23801
HepG2	EBOV VP35 R312A-eGFP	24hpi	3	12	9996438	22455	0.224630013	22821
HepG2	Mock Infected	12hpi	1	7	8346212	3	3.59445E-05	3
HepG2	Mock Infected	12hpi	2	8	7711688	2	2.59347E-05	2
HepG2	Mock Infected	12hpi	3	9	8889841	4	4.49952E-05	4
HepG2	Mock Infected	24hpi	1	10	6255280	5	7.99325E-05	6
HepG2	Mock Infected	24hpi	2	11	10723599	10	9.32523E-05	12
HepG2	Mock Infected	24hpi	3	12	10209175	11	0.000107746	14
EpoNi/22.1	EBOV wt-eGFP	12hpi	1	13	11212015	3903	0.03481087	4483
EpoNi/22.1	EBOV wt-eGFP	12hpi	2	14	13325086	4641	0.034829044	5408
EpoNi/22.1	EBOV wt-eGFP	12hpi	3	15	12931862	4748	0.036715517	5496
EpoNi/22.1	EBOV wt-eGFP	24hpi	1	16	14054704	126131	0.897429074	139333
EpoNi/22.1	EBOV wt-eGFP	24hpi	2	17	11336869	83808	0.739251728	91086
EpoNi/22.1	EBOV wt-eGFP	24hpi	3	18	10439129	85109	0.815288325	92795
EpoNi/22.1	EBOV VP35 R312A-eGFP	12hpi	1	19	9646843	1420	0.014719842	1619
EpoNi/22.1	EBOV VP35 R312A-eGFP	12hpi	2	20	8702840	1121	0.012880853	1314
EpoNi/22.1	EBOV VP35 R312A-eGFP	12hpi	3	21	11630078	1195	0.010275082	1345
EpoNi/22.1	EBOV VP35 R312A-eGFP	24hpi	1	22	9377733	32042	0.341681726	35169
EpoNi/22.1	EBOV VP35 R312A-eGFP	24hpi	2	23	9718952	31776	0.326948832	35002
EpoNi/22.1	EBOV VP35 R312A-eGFP	24hpi	3	24	9048275	29091	0.321508796	31911
EpoNi/22.1	Mock Infected	12hpi	1	19	10178507	14	0.000137545	15
EpoNi/22.1	Mock Infected	12hpi	2	20	9163838	10	0.000109125	12
EpoNi/22.1	Mock Infected	12hpi	3	21	7163614	4	5.58377E-05	4
EpoNi/22.1	Mock Infected	24hpi	1	22	12405994	24	0.000193455	27
EpoNi/22.1	Mock Infected	24hpi	2	23	12185554	13	0.000106684	14
EpoNi/22.1	Mock Infected	24hpi	3	24	10874904	6	5.51729E-05	6
R06	EBOV wt-eGFP	12hpi	1	13	4405623	32	0.000726344	32
R06	EBOV wt-eGFP	12hpi	2	14	10125723	87	0.000859198	94
R06	EBOV wt-eGFP	12hpi	3	15	4202623	44	0.001046965	46
R06	EBOV wt-eGFP	24hpi	1	16	15673499	10912	0.069620702	12010
R06	EBOV wt-eGFP	24hpi	2	12	3665112	2899	0.079097174	2993
R06	EBOV wt-eGFP	24hpi	3	18	10014035	8167	0.081555537	8825
R06	EBOV VP35 R312A-eGFP	12hpi	1	19	10639461	551	0.005178834	603
R06	EBOV VP35 R312A-eGFP	12hpi	2	20	5023551	309	0.006151027	331
R06	EBOV VP35 R312A-eGFP	12hpi	3	21	8025206	439	0.005470265	479
R06	EBOV VP35 R312A-eGFP	24hpi	1	22	7322886	16756	0.22881689	17975
R06	EBOV VP35 R312A-eGFP	24hpi	2	23	5577808	14620	0.262110134	15730
R06	EBOV VP35 R312A-eGFP	24hpi	3	24	12987884	33553	0.258340774	36148
R06	Mock Infected	12hpi	1	37	17123247	8	4.67201E-05	8
R06	Mock Infected	12hpi	2	38	1396806	1	7.15919E-05	1
R06	Mock Infected	12hpi	3	39	3614334	1	2.76676E-05	1
R06	Mock Infected	24hpi	1	40	4734794	0	0	0

R06	Mock Infected	24hpi	2	41	3857163	2	5.18516E-05	2
R06	Mock Infected	24hpi	3	42	5322986	3	5.63593E-05	4
HepG2	rMARV wt-eGFP	12hpi	1	43	6192499	870	0.014049255	870
HepG2	rMARV wt-eGFP	12hpi	2	44	11580223	1452	0.012538619	1452
HepG2	rMARV wt-eGFP	12hpi	3	45	11251516	1558	0.013847023	1558
HepG2	rMARV wt-eGFP	24hpi	1	46	8783433	1395	0.015882173	1395
HepG2	rMARV wt-eGFP	24hpi	2	47	9845865	1145	0.011629247	1145
HepG2	rMARV wt-eGFP	24hpi	3	48	8681581	958	0.011034856	958
HepG2	rMARV VP35 R301A-eGFP	12hpi	1	1	7260338	698	0.009613877	698
HepG2	rMARV VP35 R301A-eGFP	12hpi	2	2	4918218	440	0.00894633	440
HepG2	rMARV VP35 R301A-eGFP	12hpi	3	3	7371515	829	0.011245992	829
HepG2	rMARV VP35 R301A-eGFP	24hpi	1	4	7416497	942	0.012701414	942
HepG2	rMARV VP35 R301A-eGFP	24hpi	2	5	5539322	558	0.010073435	558
HepG2	rMARV VP35 R301A-eGFP	24hpi	3	6	5917564	569	0.009615443	569
HepG2	Mock Infected	12hpi	1	7	9462297	0	0	0
HepG2	Mock Infected	12hpi	2	8	5382211	0	0	0
HepG2	Mock Infected	12hpi	3	9	9343686	5	5.35121E-05	5
HepG2	Mock Infected	24hpi	1	10	8099650	2	2.46924E-05	2
HepG2	Mock Infected	24hpi	2	11	5646073	1	1.77114E-05	1
HepG2	Mock Infected	24hpi	3	12	6910535	2	2.89413E-05	2
EpoNi/22.1	rMARV wt-eGFP	12hpi	1	7	12372167	2441	0.019729769	2441
EpoNi/22.1	rMARV wt-eGFP	12hpi	2	8	11553937	2135	0.01847855	2135
EpoNi/22.1	rMARV wt-eGFP	12hpi	3	9	9444815	1849	0.019576879	1849
EpoNi/22.1	rMARV wt-eGFP	24hpi	1	10	10791051	6461	0.059873686	6461
EpoNi/22.1	rMARV wt-eGFP	24hpi	2	11	10503318	6276	0.059752547	6276
EpoNi/22.1	rMARV wt-eGFP	24hpi	3	12	10867368	6224	0.057272377	6224
EpoNi/22.1	rMARV VP35 R301A-eGFP	12hpi	1	13	9820179	4629	0.047137634	4629
EpoNi/22.1	rMARV VP35 R301A-eGFP	12hpi	2	14	7551821	3111	0.041195362	3111
EpoNi/22.1	rMARV VP35 R301A-eGFP	12hpi	3	15	8587017	3826	0.044555636	3826
EpoNi/22.1	rMARV VP35 R301A-eGFP	24hpi	1	16	10934542	6503	0.059472084	6503
EpoNi/22.1	rMARV VP35 R301A-eGFP	24hpi	2	17	8943581	5369	0.060031882	5369
EpoNi/22.1	rMARV VP35 R301A-eGFP	24hpi	3	18	9627124	5476	0.056880954	5476
EpoNi/22.1	Mock Infected	12hpi	1	19	7470848	0	0	0
EpoNi/22.1	Mock Infected	12hpi	2	20	9855398	3	3.04402E-05	3
EpoNi/22.1	Mock Infected	12hpi	3	21	9047566	2	2.21054E-05	2
EpoNi/22.1	Mock Infected	24hpi	1	22	8324301	3	3.60391E-05	3
EpoNi/22.1	Mock Infected	24hpi	2	23	7671977	5	6.51722E-05	5
EpoNi/22.1	Mock Infected	24hpi	3	24	10453430	1	9.56624E-06	1
RO6	rMARV wt-eGFP	12hpi	1	13	6780162	147	0.00216809	147
RO6	rMARV wt-eGFP	12hpi	2	14	7895383	152	0.001925176	152
RO6	rMARV wt-eGFP	12hpi	3	15	7216446	232	0.003214879	232
RO6	rMARV wt-eGFP	24hpi	1	16	9749615	1613	0.016544243	1613
RO6	rMARV wt-eGFP	24hpi	2	17	6426667	1089	0.01694502	1089
RO6	rMARV wt-eGFP	24hpi	3	18	9660110	1497	0.015496718	1497
RO6	rMARV VP35 R301A-eGFP	12hpi	1	19	9007800	197	0.002186993	197
RO6	rMARV VP35 R301A-eGFP	12hpi	2	20	15879142	404	0.002544218	404
RO6	rMARV VP35 R301A-eGFP	12hpi	3	21	11198407	209	0.001866337	209
RO6	rMARV VP35 R301A-eGFP	24hpi	1	22	9581441	1158	0.012085865	1158
RO6	rMARV VP35 R301A-eGFP	24hpi	2	23	13517623	2118	0.015668435	2118
RO6	rMARV VP35 R301A-eGFP	24hpi	3	24	7760731	1051	0.013542539	1051
RO6	Mock Infected	12hpi	1	1	10156991	3	2.95363E-05	3
RO6	Mock Infected	12hpi	2	2	7254881	3	4.13515E-05	3
RO6	Mock Infected	12hpi	3	3	9601731	2	2.08296E-05	2
RO6	Mock Infected	24hpi	1	4	8853948	9	0.00010165	9
RO6	Mock Infected	24hpi	2	5	8678884	2	2.30444E-05	2
RO6	Mock Infected	24hpi	3	6	7752454	2	2.57983E-05	2

Appendix B – Significance results for vncRNA chapter figures

Figure	Comparison	Multiple comparisons test	Multiplicity adjusted p value	
1A	HepG2 rEBOV wt-eGFP 12 hpi vs. HepG2 rEBOV wt-eGFP 24 hpi	Tukey	<0.0001	
	HepG2 rEBOV wt-eGFP 12 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 12hpi	Tukey	>0.9999	
	HepG2 rEBOV wt-eGFP 24 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.0001	
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 24hpi	Tukey	0.9887	
	HepG2 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 12 hpi	Tukey	>0.9999	
	HepG2 rEBOV wt-eGFP 24 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 24 hpi	Tukey	<0.0001	
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 12hpi	Tukey	>0.9999	
	HepG2 rEBOV VP35 R312A-eGFP 24 hpi vs. EpoNi/22.1 rEBOV VP35 R312A 24 hpi	Tukey	0.9980	
	HepG2 rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 12 hpi	Tukey	>0.9999	
	HepG2 rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	<0.0001	
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	>0.9999	
	HepG2 rEBOV VP35 R312A-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A 24 hpi	Tukey	0.9891	
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 24 hpi	Tukey	<0.0001	
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi	Tukey	>0.9999	
	EpoNi/22.1 rEBOV wt-eGFP 24 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	>0.9999	
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24hpi	Tukey	>0.9999	
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 12 hpi	Tukey	>0.9999	
	EpoNi/22.1 rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	<0.0001	
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	>0.9999	
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	>0.9999	
	RO6E/J rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	>0.9999	
	RO6E/J rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	>0.9999	
	RO6 E/J rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	>0.9999	
	RO6 E/J rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	>0.9999	
	1B	HepG2:EBOV wt-eGFP 12hpi vs. HepG2:EBOV wt-eGFP 24hpi	Tukey	> 0.9999
		HepG2:EBOV wt-eGFP 12hpi vs. HepG2:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
HepG2:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV wt-eGFP 12hpi		Tukey	> 0.9999	
HepG2:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 12hpi		Tukey	> 0.9999	
HepG2:EBOV wt-eGFP 24hpi vs. HepG2:EBOV VP35 R312A-eGFP 24hpi		Tukey	0.0005	
HepG2:EBOV wt-eGFP 24hpi vs. EpoNi/22.1:EBOV wt-eGFP 24hpi		Tukey	0.9999	
HepG2:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV wt-eGFP 24hpi		Tukey	> 0.9999	
HepG2:EBOV VP35 R312A-eGFP 12hpi vs. HepG2:EBOV VP35 R312A-eGFP 24hpi		Tukey	> 0.9999	
HepG2:EBOV VP35 R312A-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi		Tukey	> 0.9999	
HepG2:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi		Tukey	> 0.9999	
HepG2:EBOV VP35 R312A-eGFP 24hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi		Tukey	> 0.9999	
HepG2:EBOV VP35 R312A-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi		Tukey	0.0001	
EpoNi/22.1:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV wt-eGFP 24hpi		Tukey	> 0.9999	
EpoNi/22.1:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi		Tukey	> 0.9999	
EpoNi/22.1:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 12hpi		Tukey	0.0009	

	EpoNi/22.1:EBOV wt-eGFP 24hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.0001
	EpoNi/22.1:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV wt-eGFP 24hpi	Tukey	0.9985
	EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi	Tukey	> 0.9999
	EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	0.9991
	EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	> 0.9999
	RO6E/J:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 24hpi	Tukey	> 0.9999
	RO6E/J:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	RO6E/J:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	> 0.9999
	RO6E/J:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	> 0.9999
1C	HepG2:EBOV wt-eGFP 12hpi vs. HepG2:EBOV wt-eGFP 24hpi	Tukey	0.0165
	HepG2:EBOV wt-eGFP 12hpi vs. HepG2:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	HepG2:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV wt-eGFP 12hpi	Tukey	0.9971
	HepG2:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 12hpi	Tukey	> 0.9999
	HepG2:EBOV wt-eGFP 24hpi vs. HepG2:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.9954
	HepG2:EBOV wt-eGFP 24hpi vs. EpoNi/22.1:EBOV wt-eGFP 24hpi	Tukey	0.7274
	HepG2:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV wt-eGFP 24hpi	Tukey	0.0125
	HepG2:EBOV VP35 R312A-eGFP 12hpi vs. HepG2:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.1135
	HepG2:EBOV VP35 R312A-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	HepG2:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	HepG2:EBOV VP35 R312A-eGFP 24hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.8954
	HepG2:EBOV VP35 R312A-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.4025
	EpoNi/22.1:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV wt-eGFP 24hpi	Tukey	0.0017
	EpoNi/22.1:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi	Tukey	0.9998
	EpoNi/22.1:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 12hpi	Tukey	0.9784
	EpoNi/22.1:EBOV wt-eGFP 24hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.9595
	EpoNi/22.1:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV wt-eGFP 24hpi	Tukey	0.0001
	EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.0081
	EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.0181
	RO6E/J:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 24hpi	Tukey	> 0.9999
	RO6E/J:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	RO6E/J:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	RO6E/J:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.999
1D	HepG2:EBOV wt-eGFP 12hpi vs. HepG2:EBOV wt-eGFP 24hpi	Tukey	0.9994
	HepG2:EBOV wt-eGFP 12hpi vs. HepG2:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	HepG2:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV wt-eGFP 12hpi	Tukey	0.0262
	HepG2:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 12hpi	Tukey	0.0034
	HepG2:EBOV wt-eGFP 24hpi vs. HepG2:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.7398
	HepG2:EBOV wt-eGFP 24hpi vs. EpoNi/22.1:EBOV wt-eGFP 24hpi	Tukey	0.098
	HepG2:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV wt-eGFP 24hpi	Tukey	> 0.9999
	HepG2:EBOV VP35 R312A-eGFP 12hpi vs. HepG2:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.9925
	HepG2:EBOV VP35 R312A-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi	Tukey	0.0804
	HepG2:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	0.0716

	HepG2:EBOV VP35 R312A-eGFP 24hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.3259
	HepG2:EBOV VP35 R312A-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	> 0.9999
	EpoNi/22.1:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV wt-eGFP 24hpi	Tukey	< 0.0001
	EpoNi/22.1:EBOV wt-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	EpoNi/22.1:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 12hpi	Tukey	< 0.0001
	EpoNi/22.1:EBOV wt-eGFP 24hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.3394
	EpoNi/22.1:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV wt-eGFP 24hpi	Tukey	0.0341
	EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi vs. EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.0025
	EpoNi/22.1:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	< 0.0001
	EpoNi/22.1:EBOV VP35 R312A-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.1068
	RO6E/J:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV wt-eGFP 24hpi	Tukey	0.0074
	RO6E/J:EBOV wt-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 12hpi	Tukey	0.9641
	RO6E/J:EBOV wt-eGFP 24hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.662
	RO6E/J:EBOV VP35 R312A-eGFP 12hpi vs. RO6E/J:EBOV VP35 R312A-eGFP 24hpi	Tukey	0.0014
1E	HepG2 rMARV wt-eGFP 12hpi vs. HepG2 rMARV wt-eGFP 24hpi	Tukey	>0.9999
	HepG2 rMARV wt-eGFP 12hpi vs. HepG2 rMARV VP35 R301A-eGFP 12hpi	Tukey	>0.9999
	HepG2 rMARV wt-eGFP 24hpi vs. HepG2 rMARV VP35 R301A-eGFP 24hpi	Tukey	0.9989
	HepG2 rMARV VP35 R301A-eGFP 12hpi vs. HepG2 rMARV VP35 R301A-eGFP 24hpi	Tukey	>0.9999
	HepG2 rMARV wt-eGFP 12hpi vs. EpoNi/22.1 rMARV wt-eGFP 12hpi	Tukey	>0.9999
	HepG2 rMARV wt-eGFP 24hpi vs. EpoNi/22.1 rMARV wt-eGFP 24hpi	Tukey	<0.0001
	HepG2 rMARV VP35 R301A-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi	Tukey	>0.9999
	HepG2 rMARV VP35 R301A-eGFP 24 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	HepG2 rMARV wt-eGFP 12 hpi vs. RO6E/J EBOV wt 12 hpi	Tukey	>0.9999
	HepG2 rMARV wt-eGFP 24 hpi vs. RO6E/J EBOV wt 24 hpi	Tukey	>0.9999
	HepG2 rMARV VP35 R301A-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	>0.9999
	HepG2 rMARV VP35 R301A-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	>0.9999
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. EpoNi/22.1 rMARV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi	Tukey	>0.9999
	EpoNi/22.1 rMARV wt-eGFP 24 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV wt-eGFP 12 hpi	Tukey	>0.9999
	EpoNi/22.1 rMARV wt-eGFP 24 hpi vs. RO6E/J rMARV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	>0.9999
	EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	RO6E/J rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV wt-eGFP 24 hpi	Tukey	>0.9999
	RO6E/J rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	>0.9999
	RO6E/J rMARV wt-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	>0.9999
	RO6E/J rMARV VP35 R301A-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	>0.9999
1F	HepG2:MARV wt-eGFP 12hpi vs. HepG2:MARV wt-eGFP 24hpi	Tukey	0.4814
	HepG2:MARV wt-eGFP 12hpi vs. HepG2:MARV VP35 R301A-eGFP 12hpi	Tukey	> 0.9999
	HepG2:MARV wt-eGFP 12hpi vs. EpoNi/22.1:MARV wt-eGFP 12hpi	Tukey	> 0.9999
	HepG2:MARV wt-eGFP 12hpi vs. RO6E/J:MARV wt-eGFP 12hpi	Tukey	> 0.9999
	HepG2:MARV wt-eGFP 24hpi vs. HepG2:MARV VP35 R301A-eGFP 24hpi	Tukey	0.9998

	HepG2:MARV wt-eGFP 24hpi vs. EpoNi/22.1:MARV wt-eGFP 24hpi	Tukey	0.3671
	HepG2:MARV wt-eGFP 24hpi vs. RO6E/J:MARV wt-eGFP 24hpi	Tukey	0.7308
	HepG2:MARV VP35 R301A-eGFP 12hpi vs. HepG2:MARV VP35 R301A-eGFP 24hpi	Tukey	0.84
	HepG2:MARV VP35 R301A-eGFP 12hpi vs. EpoNi/22.1:MARV VP35 R301A-eGFP 12hpi	Tukey	0.6966
	HepG2:MARV VP35 R301A-eGFP 12hpi vs. RO6E/J:MARV VP35 R301A-eGFP 12hpi	Tukey	> 0.9999
	HepG2:MARV VP35 R301A-eGFP 24hpi vs. EpoNi/22.1:MARV VP35 R301A-eGFP 24hpi	Tukey	< 0.0001
	HepG2:MARV VP35 R301A-eGFP 24hpi vs. RO6E/J:MARV VP35 R301A-eGFP 24hpi	Tukey	0.9753
	EpoNi/22.1:MARV wt-eGFP 12hpi vs. EpoNi/22.1:MARV wt-eGFP 24hpi	Tukey	0.0078
	EpoNi/22.1:MARV wt-eGFP 12hpi vs. EpoNi/22.1:MARV VP35 R301A-eGFP 12hpi	Tukey	0.9403
	EpoNi/22.1:MARV wt-eGFP 12hpi vs. RO6E/J:MARV wt-eGFP 12hpi	Tukey	0.9978
	EpoNi/22.1:MARV wt-eGFP 24hpi vs. EpoNi/22.1:MARV VP35 R301A-eGFP 24hpi	Tukey	< 0.0001
	EpoNi/22.1:MARV wt-eGFP 24hpi vs. RO6E/J:MARV wt-eGFP 24hpi	Tukey	0.0071
	EpoNi/22.1:MARV VP35 R301A-eGFP 12hpi vs. EpoNi/22.1:MARV VP35 R301A-eGFP 24hpi	Tukey	< 0.0001
	EpoNi/22.1:MARV VP35 R301A-eGFP 12hpi vs. RO6E/J:MARV VP35 R301A-eGFP 12hpi	Tukey	0.4414
	EpoNi/22.1:MARV VP35 R301A-eGFP 24hpi vs. RO6E/J:MARV VP35 R301A-eGFP 24hpi	Tukey	< 0.0001
	RO6E/J:MARV wt-eGFP 12hpi vs. RO6E/J:MARV wt-eGFP 24hpi	Tukey	0.9985
	RO6E/J:MARV wt-eGFP 12hpi vs. RO6E/J:MARV VP35 R301A-eGFP 12hpi	Tukey	> 0.9999
	RO6E/J:MARV wt-eGFP 24hpi vs. RO6E/J:MARV VP35 R301A-eGFP 24hpi	Tukey	> 0.9999
	RO6E/J:MARV VP35 R301A-eGFP 12hpi vs. RO6E/J:MARV VP35 R301A-eGFP 24hpi	Tukey	0.9989
4A	HepG2 EBOV wt 12 hpi vs. HepG2 rEBOV wt-eGFP 24 hpi	Tukey	<0.0001
	HepG2 rEBOV wt-eGFP 12 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 12 hpi	Tukey	>0.9999
	HepG2 rEBOV wt-eGFP 24 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	>0.9999
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	<0.0001
	HepG2 rEBOV wt-eGFP 12hpi vs. EpoNi/22.1 rEBOV wt-eGFP 12 hpi	Tukey	0.9914
	HepG2 rEBOV wt-eGFP 24 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 24 hpi	Tukey	<0.0001
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi	Tukey	<0.9999
	HepG2 rEBOV VP35 R312A-eGFP 24 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.0024
	HepG2 rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 12 hpi	Tukey	>0.9999
	HepG2 rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	<0.0001
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	>0.9999
	HepG2 rEBOV VP35 R312A-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.9974
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi	Tukey	0.99
	EpoNi/22.1 rEBOV wt-eGFP 24 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24hpi	Tukey	<0.0001
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 12 hpi	Tukey	0.8481
	EpoNi/22.1 rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	>0.9999
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.0226
	RO6E/J rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	0.0367
	RO6E/J rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	>0.9999
	RO6 E/J rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	<0.0001
	RO6 E/J rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	<0.0001
4B	HepG2 rMARV wt-eGFP 12hpi vs. HepG2 rMARV wt-eGFP 24hpi	Tukey	>0.9999

	HepG2 rMARV wt-eGFP 12hpi vs. HepG2 rMARV VP35 R301A-eGFP 12hpi	Tukey	0.2751
	HepG2 rMARV wt-eGFP 24hpi vs. HepG2 rMARV VP35 R301A-eGFP 24hpi	Tukey	0.9020
	HepG2 rMARV VP35 R301A-eGFP 12hpi vs. HepG2 rMARV VP35 R301A-eGFP 24hpi	Tukey	>0.9999
	HepG2 rMARV wt-eGFP 12hpi vs. EpoNi/22.1 rMARV wt-eGFP 12hpi	Tukey	0.0075
	HepG2 rMARV wt-eGFP 24hpi vs. EpoNi/22.1 rMARV wt-eGFP 24hpi	Tukey	<0.0001
	HepG2 rMARV VP35 R301A-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi	Tukey	<0.0001
	HepG2 rMARV VP35 R301A-eGFP 24 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	HepG2 rMARV wt-eGFP 12 hpi vs. RO6E/J EBOV wt 12 hpi	Tukey	<0.0001
	HepG2 rMARV wt-eGFP 24 hpi vs. RO6E/J EBOV wt 24 hpi	Tukey	0.2973
	HepG2 rMARV VP35 R301A-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	0.0002
	HepG2 rMARV VP35 R301A-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	0.5152
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. EpoNi/22.1 rMARV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV wt-eGFP 24 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	>0.9999
	EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV wt-eGFP 12 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV wt-eGFP 24 hpi vs. RO6E/J rMARV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	RO6E/J rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV wt-eGFP 24 hpi	Tukey	<0.0001
	RO6E/J rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	>0.9999
	RO6E/J rMARV wt-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	0.7091
	RO6E/J rMARV VP35 R301A-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
6A	HepG2 rEBOV wt-eGFP 12 hpi vs. HepG2 rEBOV wt-eGFP 24 hpi	Tukey	<0.0001
	HepG2 rEBOV wt-eGFP 12 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 12hpi	Tukey	0.9743
	HepG2 rEBOV wt-eGFP 24 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	> 0.9999
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. HepG2 rEBOV VP35 R312A-eGFP 24hpi	Tukey	0.0022
	HepG2 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 12 hpi	Tukey	<0.0001
	HepG2 rEBOV wt-eGFP 24 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 24 hpi	Tukey	0.0021
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 12hpi	Tukey	<0.0001
	HepG2 rEBOV VP35 R312A-eGFP 24 hpi vs. EpoNi/22.1 rEBOV VP35 R312A 24 hpi	Tukey	0.0014
	HepG2 rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 12 hpi	Tukey	0.0364
	HepG2 rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	0.1616
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	0.4387
	HepG2 rEBOV VP35 R312A-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A 24 hpi	Tukey	0.1191
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi	Tukey	0.9993
	EpoNi/22.1 rEBOV wt-eGFP 24 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.9972
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24hpi	Tukey	< 0.0001
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 12 hpi	Tukey	<0.0001
	EpoNi/22.1 rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	0.8204
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	0.0125
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.721
	RO6E/J rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	>0.9999
	RO6E/J rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	0.0026

	RO6 E/J rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.9996
	RO6 E/J rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.0139
6B	HepG2 rMARV wt-eGFP 12hpi vs. HepG2 rMARV wt-eGFP 24hpi	Tukey	<0.0001
	HepG2 rMARV wt-eGFP 12hpi vs. HepG2 rMARV VP35 R301A-eGFP 12hpi	Tukey	0.7612
	HepG2 rMARV wt-eGFP 24hpi vs. HepG2 rMARV VP35 R301A-eGFP 24hpi	Tukey	0.0006
	HepG2 rMARV VP35 R301A-eGFP 12hpi vs. HepG2 rMARV VP35 R301A-eGFP 24hpi	Tukey	<0.0001
	HepG2 rMARV wt-eGFP 12hpi vs. EpoNi/22.1 rMARV wt-eGFP 12hpi	Tukey	>0.9999
	HepG2 rMARV wt-eGFP 24hpi vs. EpoNi/22.1 rMARV wt-eGFP 24hpi	Tukey	0.1336
	HepG2 rMARV VP35 R301A-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi	Tukey	0.4964
	HepG2 rMARV VP35 R301A-eGFP 24 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	HepG2 rMARV wt-eGFP 12 hpi vs. RO6E/J EBOV wt 12 hpi	Tukey	0.0745
	HepG2 rMARV wt-eGFP 24 hpi vs. RO6E/J EBOV wt 24 hpi	Tukey	<0.0001
	HepG2 rMARV VP35 R301A-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	<0.0001
	HepG2 rMARV VP35 R301A-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. EpoNi/22.1 rMARV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi	Tukey	>0.9999
	EpoNi/22.1 rMARV wt-eGFP 24 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	0.0417
	EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi vs. EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi	Tukey	0.0005
	EpoNi/22.1 rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV wt-eGFP 12 hpi	Tukey	0.0234
	EpoNi/22.1 rMARV wt-eGFP 24 hpi vs. RO6E/J rMARV wt-eGFP 24 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV VP35 R301A-eGFP 12 hpi vs. RO6 E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	<0.0001
	EpoNi/22.1 rMARV VP35 R301A-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	0.0434
	RO6E/J rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV wt-eGFP 24 hpi	Tukey	0.0875
	RO6E/J rMARV wt-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 12 hpi	Tukey	0.0003
	RO6E/J rMARV wt-eGFP 24 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	0.9495
	RO6E/J rMARV VP35 R301A-eGFP 12 hpi vs. RO6E/J rMARV VP35 R301A-eGFP 24 hpi	Tukey	<0.0001
9A	HepG2 rEBOV wt-eGFP 12hpi vs. HepG2 rEBOV wt-eGFP 24hpi	Tukey	> 0.9999
	HepG2 rEBOV wt-eGFP 12hpi vs. HepG2 rEBOV VP35 R312A-eGFP 12hpi	Tukey	> 0.9999
	HepG2 rEBOV wt-eGFP 24hpi vs. HepG2 rEBOV VP35 R312A-eGFP 24hpi	Tukey	> 0.9999
	HepG2 rEBOV VP35 R312A-eGFP 12hpi vs. HepG2 rEBOV VP35 R312A-eGFP 24hpi	Tukey	0.9997
	HepG2 rEBOV wt-eGFP 12hpi vs. EpoNi/22.1 rEBOV wt-eGFP 12hpi	Tukey	0.0472
	HepG2 rEBOV wt-eGFP 24hpi vs. EpoNi/22.1 rEBOV wt-eGFP 24hpi	Tukey	0.0221
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi	Tukey	0.0402
	HepG2 rEBOV VP35 R312A-eGFP 24 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.038
	HepG2 rEBOV wt-eGFP 12 hpi vs. RO6E/J EBOV wt 12 hpi	Tukey	> 0.9999
	HepG2 rEBOV wt-eGFP 24 hpi vs. RO6E/J EBOV wt 24 hpi	Tukey	> 0.9999
	HepG2 rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	> 0.9999
	HepG2 rEBOV VP35 R312A-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	> 0.9999
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV wt-eGFP 24 hpi	Tukey	0.9996
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi	Tukey	0.9996
	EpoNi/22.1 rEBOV wt-eGFP 24 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	> 0.9999
	EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi vs. EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.988
	EpoNi/22.1 rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 12 hpi	Tukey	0.0196

EpoNi/22.1 rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	0.0058
EpoNi/22.1 rEBOV VP35 R312A-eGFP 12 hpi vs. RO6 E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	0.0228
EpoNi/22.1 rEBOV VP35 R312A-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.022
RO6E/J rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV wt-eGFP 24 hpi	Tukey	> 0.9999
RO6E/J rEBOV wt-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 12 hpi	Tukey	> 0.9999
RO6E/J rEBOV wt-eGFP 24 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	> 0.9999
RO6E/J rEBOV VP35 R312A-eGFP 12 hpi vs. RO6E/J rEBOV VP35 R312A-eGFP 24 hpi	Tukey	0.9997

Appendix C – Additional reagents for vncRNA chapter

Tag-based strand-specific EBOV qPCR

RNA Target	RT Primer (5'-3')	Forward (5'-3')	Reverse (5'-3')	Probe (5'-3')
vRNA	ATGCCTAGCTGA AGCTAGCGTGAC ATATTACTGCCGC AATGAATT	AATAAATCATAA ATGCCTAGCTGA AGCTAGCGT	AATAAATCATA AGCCAGACCT TTCGTAAAGC T	GCAACATAATA AACTCTGCACT
mRNA	CCAGATCGTTCG AGTCGTTTTTTTT TTTTTCTTAATTA GA	AATAAATCATAA ATCCTCAAATTG CCTGCATGCT	AATAAATCATA ACCAGATCGTT CGAGTCG	GGTTGTTTCA ATCCAAGTAC
Standards	Sequence (5'-3')			
vRNA	AATAAATCATAAGCCAGACCTTTCGTAAAGCTTAATTATAAAGAGTGCAGA GTTTATTATGTTGCGTTAAATTCATTGCGGCAGTAATATGTCACGCTAGCTTCA GCTAGGCATTTATGATTTATT			
mRNA	AATAAATCATAAATCCTCAAATTGCCTGCATGCTTACATCTGAGGATAGCCAG TGTGACTTGGATTGGAAATGTGGAGAAAAAATCGGGACCCATTTCTAGGTTGT TCACAATCCAAGTACAGACATTGCCCTTCTAATTAAGAAAAAAAAAAAAAACG ACTCGAACGATCTGGTTATGATTTATT			

miRNA-specific qPCR

Target	Ordering Reference ID
EBOV GP vncRNA	YCP0032383
EBOV VP40 vncRNA	YCP0032387
EBOV GP vncRNA	YCP0032385
EBOV VP24 vncRNA	YCP0031721
EBOV L vncRNA	YCP0045230
hsa-miR-103a-3p	YP00204063
hsa-let-7a-5p	YP00205727
U6 snRNA	YP00203907
UniSp6 Synthetic RNA Spike-In	YP00203954

Antibodies

Antigen	Host	Vendor	Catalog Number	Dilution
RNase L	Rabbit	Cell Signaling	27281	1:1000 (for WB)
DIS3	Rabbit	Thermo Fisher	PA5-58723	1:500 (for WB)
CPSF3L (INTS11)	Rabbit	Abcam	ab75276	1:2000 (for WB)
pan-actin	Mouse	Thermo Fisher	MA5-11869	1:2500 (for WB)
Mouse IgG (HRP conjugated)	Sheep	Amersham	NA931	1:2500 (for WB)
Rabbit IgG (HRP conjugated)	Donkey	Amersham	NA934	1:2500 (for WB)
AGO1	Mouse	Millipore Sigma	03-249	5 µg (for IP)
AGO2	Mouse	Millipore Sigma	03-110	5 µg (for IP)
AGO3	Rabbit	Cell Signaling	5054	5 µg (for IP)
AGO4	Rabbit	Millipore Sigma	6913	5 µg (for IP)
pan-AGO	Mouse	Millipore Sigma	03-248	5 µg (for IP)
SAM 68	Rabbit	Millipore Sigma	07-415-I	10 µL (for IP)
IgG isotype control antibody	Mouse	Millipore Sigma	CS200621 (part #)	5 µg (for IP)
Rabbit anti-EBOV GP pAb	Rabbit	IBT Bioservices	0301-015	1:8000 (for plaque immunostaining)
Goat Anti-Rabbit IgG (H+L), Mouse/Human adsorbed-HRP	Goat	SouthernBiotech	4050-05	1:1000 (for plaque immunostaining)

siRNAs

siRNA	Manufacturer	Catalog #	Final Concentration used in experiment	Guide Strand Sequence (if available, excludes overhang nts)
Drosha	Thermo Fisher	4390824 (assay ID: s26491)	25 nM	
eIF2C2 (Ago2)	Santa Cruz Biotechnology	sc-44409	50 nM	
RNase L	Santa Cruz Biotechnology	sc-45965	50 nM	
DIS3	Dharmacon	L-015405-01-0005	25 nM	
CPSF3L	Dharmacon	L-013789-01-0005	25 nM	
Scrambled (non-targeting)	Santa Cruz Biotechnology	sc-37007	50 nM	
EBOV VP40 siRNA mimic	Dharmacon	Custom Order	Indicated in Methods/Figure	5'-GAUGAAGAUUAAGAAAAACCUA-3'
EBOV L siRNA mimic	Dharmacon	Custom Order	Indicated in Methods/Figure	5'-GAGGAAGAUUAAGAAAAACUGC-3'
EBOV GP siRNA mimic	Dharmacon	Custom Order	Indicated in Methods/Figure	5'-GAUGAAGAUUAAGCCGACAGUG-3'

pmIRGLO Inserts

Perfect MBE Insert
TAAGCAGGAGCTCGAAGATGGAATGACCTTACACTGTGGCTTAATCTTCATCAACGTATATTCAATGTAATGAA GTCGGAGGATTAACGTGGGAATCGTGCTTCTGTCTAAACAAGTAAGGATATAAAAGTTGTAACCGTTCCTCAAGCGT ACAGGGTGCATTTTGTAAACAATTTGGGAGTCCAAAGACTCGCTGTTTTTAAAATTTATCCTCAAGCGCGAGTATTGA ACCAAGCTTACGTCTAAGAACGTAGCAAGCTGACTCAAACAAAATACATTTTGCCCGCGTTACATATAAAATTAAGTT AAAAGTTATGGAATATAATAACATGTGGATGGCCAGTGGTGGGTTGTTACACCCCTACGGCAATGTTGAAAACACTG TCGGCTTAATCTTCATC TTTAAGCCGTGACACCCGTTATACTCCATAACCGTCTGTAACATAGCTTGTCTGGAT TGGATTGTCAATCTCTCAGAGTATTATGCGTCTGACTAAGCAG
Imperfect MBE Insert
TAAGCAGGAGCTCGAAGATGGAATGACCTTACACTGTGGCATTCTTCATCCAACGTATATTCAATGTAATGAAG TCGGAGGATTAACGTGGGAATCGTGCTTCTGTCTAAACAAGTAAGGATATAAAAGTTGTAACCGTTCCTCAAGCGT AGGGTGCATTTTGTAAACAATTTGGGAGTCCAAAGACTCGCTGTTTTTAAAATTTATCCTCAAGCGCGAGTATTGA CAAGCTTACGTCTAAGAACGTAGCAAGCTGACTCAAACAAAATACATTTTGCCCGCGTTACATATAAAATTAAGTTAA AAGTTATGGAATATAATAACATGTGGATGGCCAGTGGTGGGTTGTTACACCCCTACGGCAATGTTGAAAACACTGTC GGCATTCTTCATC TTTAAGCCGTGACACCCGTTATACTCCATAACCGTCTGTAACATAGCTTGTCTGGATTG GATTGTCAATCTCTCAGAGTATTATGCGTCTGACTAAGCAG
10 nt MBE Insert
TAAGCAGGAGCTCGAAGATGGAATGACCTTAAGACTGGCGTACAATCTTCATCAACGTATATTCAATGTAATGAA GTCGGAGGATTAACGTGGGAATCGTGCTTCTGTCTAAACAAGTAAGGATATAAAAGTTGTAACCGTTCCTCAAGCGT ACAGGGTGCATTTTGTAAACAATTTGGGAGTCCAAAGACTCGCTGTTTTTAAAATTTATCCTCAAGCGCGAGTATTGA ACCAAGCTTACGTCTAAGAACGTAGCAAGCTGACTCAAACAAAATACATTTTGCCCGCGTTACATATAAAATTAAGTT AAAAGTTATGGAATATAATAACATGTGGATGGCCAGTGGTGGGTTGTTACACCCCTACGGCAATGTTGAAAAGACT GGCGTACAATCTTCATC TTTAAGCCGTGACACCCGTTATACTCCATAACCGTCTGTAACATAGCTTGTCTGGAT TGGATTGTCAATCTCTCAGAGTATTATGCGTCTGACTAAGCAG
Key
Underlined Bolded sequences indicate miRNA binding elements (MBEs) <i>SalI</i> Restriction Site <i>Sal I</i> Restriction Site Scrambled+A1:A15 3'-UTRs excluding MBEs are identical between inserts

Appendix D – Novel filovirus 5' mRNA structures

Predicted sequences of Xilǎng virus (XILV) (Genus *Striavirus*)

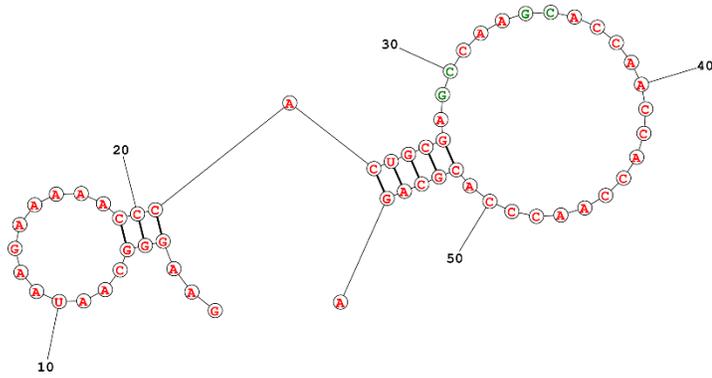
transcription start sites

Signal	Sequence (positive sense)
NP	<i>GAAGGGCAAUAAGACAAC</i>
Internal genes	<i>GAAGGGCAAUAAGAAAAA</i>
L	<i>UUUGGGGAUUAAGAAAAA</i>

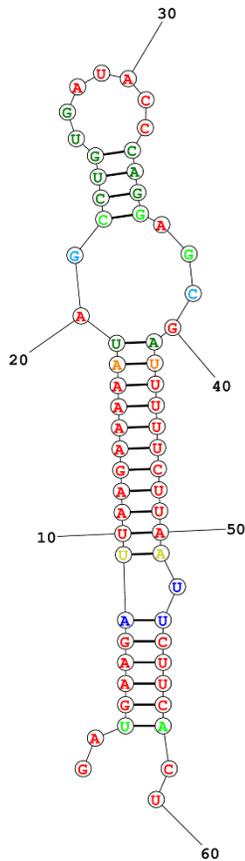
The putative XILV intergenic element can be split into what appear to be “transcription start” (*italics*) and “transcription stop” (underlined) elements, based upon their presence or absence in the single function signals preceding the putative NP ORF and following the putative L ORF. Signals in the gene borders of internal genes contain both components

Example predicted XILV and BOMV 5' mRNA secondary structures

XILV VP40 ANALOGUE



BOMV VP40



Structures prepared using a publically available tool (437).

References

1. **Kuhn JH, Becker S, Ebihara H, Geisbert TW, Johnson KM, Kawaoka Y, Lipkin WI, Negredo AI, Netesov SV, Nichol ST, Palacios G, Peters CJ, Tenorio A, Volchkov VE, Jahrling PB.** 2010. Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. *Arch Virol* **155**:2083-2103.
2. **Bukreyev AA, Chandran K, Dolnik O, Dye JM, Ebihara H, Leroy EM, Muhlberger E, Netesov SV, Patterson JL, Paweska JT, Saphire EO, Smither SJ, Takada A, Towner JS, Volchkov VE, Warren TK, Kuhn JH.** 2014. Discussions and decisions of the 2012-2014 International Committee on Taxonomy of Viruses (ICTV) Filoviridae Study Group, January 2012-June 2013. *Arch Virol* **159**:821-830.
3. **Goldstein T, Anthony SJ, Gbakima A, Bird BH, Bangura J, Tremeau-Bravard A, Belaganahalli MN, Wells HL, Dhanota JK, Liang E, Grodus M, Jangra RK, DeJesus VA, Lasso G, Smith BR, Jambai A, Kamara BO, Kamara S, Bangura W, Monagin C, Shapira S, Johnson CK, Saylor K, Rubin EM, Chandran K, Lipkin WI, Mazet JAK.** 2018. The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses. *Nat Microbiol* **3**:1084-1089.
4. **Emanuel J, Marzi A, Feldmann H.** 2018. Filoviruses: Ecology, Molecular Biology, and Evolution. *Adv Virus Res* **100**:189-221.
5. **Spengler JR, Ervin ED, Towner JS, Rollin PE, Nichol ST.** 2016. Perspectives on West Africa Ebola Virus Disease Outbreak, 2013-2016. *Emerg Infect Dis* **22**:956-963.
6. **W. H. O. Ebola Response Team, Agua-Agum J, Allegranzi B, Ariyaratnam A, Aylward R, Blake IM, Barboza P, Bausch D, Brennan RJ, Clement P, Coffey**

- P, Cori A, Donnelly CA, Dorigatti I, Drury P, Durski K, Dye C, Eckmanns T, Ferguson NM, Fraser C, Garcia E, Garske T, Gasasira A, Gurry C, Hamblion E, Hinsley W, Holden R, Holmes D, Hugonnet S, Jaramillo Gutierrez G, Jombart T, Kelley E, Santhana R, Mahmoud N, Mills HL, Mohamed Y, Musa E, Naidoo D, Nedjati-Gilani G, Newton E, Norton I, Nouvellet P, Perkins D, Perkins M, Riley S, Schumacher D, Shah A, Tang M, Varsaneux O, Van Kerkhove MD.** 2016. After Ebola in West Africa-- Unpredictable Risks, Preventable Epidemics. *N Engl J Med* **375**:587-596.
7. **Feldmann H, Geisbert TW.** 2011. Ebola haemorrhagic fever. *Lancet* **377**:849-862.
 8. **Vinck P, Pham PN, Bindu KK, Bedford J, Nilles EJ.** 2019. Institutional trust and misinformation in the response to the 2018-19 Ebola outbreak in North Kivu, DR Congo: a population-based survey. *Lancet Infect Dis* **19**:529-536.
 9. **Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, Wang W, Eden JS, Shen JJ, Liu L, Holmes EC, Zhang YZ.** 2018. The evolutionary history of vertebrate RNA viruses. *Nature* **556**:197-202.
 10. **Booth TF, Rabb MJ, Beniac DR.** 2013. How do filovirus filaments bend without breaking? *Trends Microbiol* **21**:583-593.
 11. **Ellis DS, Bowen ET, Simpson DI, Stamford S.** 1978. Ebola virus: a comparison, at ultrastructural level, of the behaviour of the Sudan and Zaire strains in monkeys. *Br J Exp Pathol* **59**:584-593.
 12. **Ellis DS, Simpson IH, Francis DP, Knobloch J, Bowen ET, Lolik P, Deng IM.** 1978. Ultrastructure of Ebola virus particles in human liver. *J Clin Pathol* **31**:201-208.
 13. **Geisbert TW, Jahrling PB.** 1995. Differentiation of filoviruses by electron microscopy. *Virus Res* **39**:129-150.

14. **Beniac DR, Melito PL, Devarenes SL, Hiebert SL, Rabb MJ, Lamboo LL, Jones SM, Booth TF.** 2012. The organisation of Ebola virus reveals a capacity for extensive, modular polyploidy. *PLoS One* **7**:e29608.
15. **Bharat TA, Noda T, Riches JD, Kraehling V, Kolesnikova L, Becker S, Kawaoka Y, Briggs JA.** 2012. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. *Proc Natl Acad Sci U S A* **109**:4275-4280.
16. **Sanchez A, Kiley MP, Holloway BP, Auperin DD.** 1993. Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res* **29**:215-240.
17. **Takada A, Robison C, Goto H, Sanchez A, Murti KG, Whitt MA, Kawaoka Y.** 1997. A system for functional analysis of Ebola virus glycoprotein. *Proc Natl Acad Sci U S A* **94**:14764-14769.
18. **Volchkova VA, Klenk HD, Volchkov VE.** 1999. Delta-peptide is the carboxy-terminal cleavage fragment of the nonstructural small glycoprotein sGP of Ebola virus. *Virology* **265**:164-171.
19. **Ito H, Watanabe S, Takada A, Kawaoka Y.** 2001. Ebola virus glycoprotein: proteolytic processing, acylation, cell tropism, and detection of neutralizing antibodies. *J Virol* **75**:1576-1580.
20. **Mehedi M, Falzarano D, Seebach J, Hu X, Carpenter MS, Schnittler HJ, Feldmann H.** 2011. A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J Virol* **85**:5406-5414.
21. **He J, Melnik LI, Komin A, Wiedman G, Fuselier T, Morris CF, Starr CG, Searson PC, Gallaher WR, Hristova K, Garry RF, Wimley WC.** 2017. Ebola Virus Delta Peptide is a Viroporin. *J Virol* doi:10.1128/JVI.00438-17.
22. **Radoshitzky SR, Warfield KL, Chi X, Dong L, Kota K, Bradfute SB, Gearhart JD, Retterer C, Kranzusch PJ, Misasi JN, Hogenbirk MA, Wahl-**

- Jensen V, Volchkov VE, Cunningham JM, Jahrling PB, Aman MJ, Bavari S, Farzan M, Kuhn JH.** 2011. Ebola virus delta-peptide immunoadhesins inhibit marburgvirus and ebolavirus cell entry. *J Virol* **85**:8502-8513.
23. **Watanabe S, Noda T, Kawaoka Y.** 2006. Functional mapping of the nucleoprotein of Ebola virus. *J Virol* **80**:3743-3751.
24. **Kirchdoerfer RN, Abelson DM, Li S, Wood MR, Saphire EO.** 2015. Assembly of the Ebola Virus Nucleoprotein from a Chaperoned VP35 Complex. *Cell Rep* **12**:140-149.
25. **Kirchdoerfer RN, Saphire EO, Ward AB.** 2019. Cryo-EM structure of the Ebola virus nucleoprotein-RNA complex. *Acta Crystallogr F Struct Biol Commun* **75**:340-347.
26. **Noda T, Ebihara H, Muramoto Y, Fujii K, Takada A, Sagara H, Kim JH, Kida H, Feldmann H, Kawaoka Y.** 2006. Assembly and budding of Ebolavirus. *PLoS Pathog* **2**:e99.
27. **Muhlberger E, Lotfering B, Klenk HD, Becker S.** 1998. 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**:8756-8764.
28. **Su Z, Wu C, Shi L, Luthra P, Pintilie GD, Johnson B, Porter JR, Ge P, Chen M, Liu G, Frederick TE, Binning JM, Bowman GR, Zhou ZH, Basler CF, Gross ML, Leung DW, Chiu W, Amarasinghe GK.** 2018. Electron Cryo-microscopy Structure of Ebola Virus Nucleoprotein Reveals a Mechanism for Nucleocapsid-like Assembly. *Cell* **172**:966-978 e912.
29. **Muhlberger E, Weik M, Volchkov VE, Klenk HD, Becker S.** 1999. Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J Virol* **73**:2333-2342.

30. **Prins KC, Binning JM, Shabman RS, Leung DW, Amarasinghe GK, Basler CF.** 2010. Basic residues within the ebolavirus VP35 protein are required for its viral polymerase cofactor function. *J Virol* **84**:10581-10591.
31. **Bukreyev AA, Volchkov VE, Blinov VM, Netesov SV.** 1993. The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses. *FEBS Lett* **322**:41-46.
32. **Muhlberger E.** 2007. Filovirus replication and transcription. *Future Virol* **2**:205-215.
33. **Luthra P, Ramanan P, Mire CE, Weisend C, Tsuda Y, Yen B, Liu G, Leung DW, Geisbert TW, Ebihara H, Amarasinghe GK, Basler CF.** 2013. Mutual antagonism between the Ebola virus VP35 protein and the RIG-I activator PACT determines infection outcome. *Cell Host Microbe* **14**:74-84.
34. **Basler CF, Wang X, Muhlberger E, Volchkov V, Paragas J, Klenk HD, Garcia-Sastre A, Palese P.** 2000. The Ebola virus VP35 protein functions as a type I IFN antagonist. *Proc Natl Acad Sci U S A* **97**:12289-12294.
35. **Basler CF, Mikulasova A, Martinez-Sobrido L, Paragas J, Muhlberger E, Bray M, Klenk HD, Palese P, Garcia-Sastre A.** 2003. The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. *J Virol* **77**:7945-7956.
36. **Cardenas WB, Loo YM, Gale M, Jr., Hartman AL, Kimberlin CR, Martinez-Sobrido L, Saphire EO, Basler CF.** 2006. Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling. *J Virol* **80**:5168-5178.
37. **Prins KC, Cardenas WB, Basler CF.** 2009. Ebola virus protein VP35 impairs the function of interferon regulatory factor-activating kinases IKKepsilon and TBK-1. *J Virol* **83**:3069-3077.

38. **Haasnoot J, de Vries W, Geutjes EJ, Prins M, de Haan P, Berkhout B.** 2007. The Ebola virus VP35 protein is a suppressor of RNA silencing. *PLoS Pathog* **3**:e86.
39. **Fabozzi G, Nabel CS, Dolan MA, Sullivan NJ.** 2011. Ebolavirus proteins suppress the effects of small interfering RNA by direct interaction with the mammalian RNA interference pathway. *J Virol* **85**:2512-2523.
40. **John SP, Wang T, Steffen S, Longhi S, Schmaljohn CS, Jonsson CB.** 2007. Ebola virus VP30 is an RNA binding protein. *J Virol* **81**:8967-8976.
41. **Modrof J, Becker S, Muhlberger E.** 2003. Ebola virus transcription activator VP30 is a zinc-binding protein. *J Virol* **77**:3334-3338.
42. **Modrof J, Muhlberger E, Klenk HD, Becker S.** 2002. Phosphorylation of VP30 impairs ebola virus transcription. *J Biol Chem* **277**:33099-33104.
43. **Hartlieb B, Modrof J, Muhlberger E, Klenk HD, Becker S.** 2003. Oligomerization of Ebola virus VP30 is essential for viral transcription and can be inhibited by a synthetic peptide. *J Biol Chem* **278**:41830-41836.
44. **Martinez MJ, Biedenkopf N, Volchkova V, Hartlieb B, Alazard-Dany N, Reynard O, Becker S, Volchkov V.** 2008. Role of Ebola virus VP30 in transcription reinitiation. *J Virol* **82**:12569-12573.
45. **Ilinykh PA, Tigabu B, Ivanov A, Ammosova T, Obukhov Y, Garron T, Kumari N, Kovalskyy D, Platonov MO, Naumchik VS, Freiberg AN, Nekhai S, Bukreyev A.** 2014. 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**:22723-22738.
46. **Trunschke M, Conrad D, Enterlein S, Olejnik J, Brauburger K, Muhlberger E.** 2013. 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**:135-145.

47. **Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, Grant T, Grigorieff N, Harrison SC, Whelan SP.** 2015. Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell* **162**:314-327.
48. **Tchesnokov EP, Raesisimakiani P, Ngure M, Marchant D, Gotte M.** 2018. Recombinant RNA-Dependent RNA Polymerase Complex of Ebola Virus. *Sci Rep* **8**:3970.
49. **Noda T, Sagara H, Suzuki E, Takada A, Kida H, Kawaoka Y.** 2002. Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP. *J Virol* **76**:4855-4865.
50. **Noda T, Watanabe S, Sagara H, Kawaoka Y.** 2007. Mapping of the VP40-binding regions of the nucleoprotein of Ebola virus. *J Virol* **81**:3554-3562.
51. **Watanabe S, Watanabe T, Noda T, Takada A, Feldmann H, Jasenosky LD, Kawaoka Y.** 2004. Production of novel ebola virus-like particles from cDNAs: an alternative to ebola virus generation by reverse genetics. *J Virol* **78**:999-1005.
52. **Mateo M, Carbonnelle C, Martinez MJ, Reynard O, Page A, Volchkova VA, Volchkov VE.** 2011. Knockdown of Ebola virus VP24 impairs viral nucleocapsid assembly and prevents virus replication. *J Infect Dis* **204 Suppl 3**:S892-896.
53. **Reid SP, Leung LW, Hartman AL, Martinez O, Shaw ML, Carbonnelle C, Volchkov VE, Nichol ST, Basler CF.** 2006. Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. *J Virol* **80**:5156-5167.
54. **Reid SP, Valmas C, Martinez O, Sanchez FM, Basler CF.** 2007. Ebola virus VP24 proteins inhibit the interaction of NPI-1 subfamily karyopherin alpha proteins with activated STAT1. *J Virol* **81**:13469-13477.
55. **Watanabe S, Noda T, Halfmann P, Jasenosky L, Kawaoka Y.** 2007. Ebola virus (EBOV) VP24 inhibits transcription and replication of the EBOV genome. *J Infect Dis* **196 Suppl 2**:S284-290.

56. **Watt A, Moukambi F, Banadyga L, Groseth A, Callison J, Herwig A, Ebihara H, Feldmann H, Hoenen T.** 2014. A novel life cycle modeling system for Ebola virus shows a genome length-dependent role of VP24 in virus infectivity. *J Virol* **88**:10511-10524.
57. **Sanchez A, Trappier SG, Mahy BW, Peters CJ, Nichol ST.** 1996. 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**:3602-3607.
58. **Volchkov VE, Volchkova VA, Muhlberger E, Kolesnikova LV, Weik M, Dolnik O, Klenk HD.** 2001. Recovery of infectious Ebola virus from complementary DNA: RNA editing of the GP gene and viral cytotoxicity. *Science* **291**:1965-1969.
59. **Mehedi M, Hoenen T, Robertson S, Ricklefs S, Dolan MA, Taylor T, Falzarano D, Ebihara H, Porcella SF, Feldmann H.** 2013. Ebola virus RNA editing depends on the primary editing site sequence and an upstream secondary structure. *PLoS Pathog* **9**:e1003677.
60. **Mohan GS, Li W, Ye L, Compans RW, Yang C.** 2012. Antigenic subversion: a novel mechanism of host immune evasion by Ebola virus. *PLoS Pathog* **8**:e1003065.
61. **Ilinykh PA, Shen X, Flyak AI, Kuzmina N, Ksiazek TG, Crowe JE, Jr., Bukreyev A.** 2016. Chimeric Filoviruses for Identification and Characterization of Monoclonal Antibodies. *J Virol* **90**:3890-3901.
62. **Volchkov VE, Feldmann H, Volchkova VA, Klenk HD.** 1998. Processing of the Ebola virus glycoprotein by the proprotein convertase furin. *Proc Natl Acad Sci U S A* **95**:5762-5767.
63. **Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, Kuehne AI, Kranzusch PJ, Griffin AM, Ruthel G, Dal Cin P, Dye JM,**

- Whelan SP, Chandran K, Brummelkamp TR.** 2011. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* **477**:340-343.
64. **Cote M, Misasi J, Ren T, Bruchez A, Lee K, Filone CM, Hensley L, Li Q, Ory D, Chandran K, Cunningham J.** 2011. Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. *Nature* **477**:344-348.
65. **Lee JE, Fusco ML, Hessel AJ, Oswald WB, Burton DR, Saphire EO.** 2008. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* **454**:177-182.
66. **Dolnik O, Volchkova V, Garten W, Carbonnelle C, Becker S, Kahnt J, Stroher U, Klenk HD, Volchkov V.** 2004. Ectodomain shedding of the glycoprotein GP of Ebola virus. *EMBO J* **23**:2175-2184.
67. **Escudero-Perez B, Volchkova VA, Dolnik O, Lawrence P, Volchkov VE.** 2014. Shed GP of Ebola virus triggers immune activation and increased vascular permeability. *PLoS Pathog* **10**:e1004509.
68. **Kondratowicz AS, Lennemann NJ, Sinn PL, Davey RA, Hunt CL, Moller-Tank S, Meyerholz DK, Rennert P, Mullins RF, Brindley M, Sandersfeld LM, Quinn K, Weller M, McCray PB, Jr., Chiorini J, Maury W.** 2011. T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus. *Proc Natl Acad Sci U S A* **108**:8426-8431.
69. **Alvarez CP, Lasala F, Carrillo J, Muniz O, Corbi AL, Delgado R.** 2002. C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. *J Virol* **76**:6841-6844.
70. **Simmons G, Reeves JD, Grogan CC, Vandenberghe LH, Baribaud F, Whitbeck JC, Burke E, Buchmeier MJ, Soilleux EJ, Riley JL, Doms RW, Bates P, Pohlmann S.** 2003. DC-SIGN and DC-SIGNR bind ebola glycoproteins and enhance infection of macrophages and endothelial cells. *Virology* **305**:115-123.

71. **Aleksandrowicz P, Marzi A, Biedenkopf N, Beimforde N, Becker S, Hoenen T, Feldmann H, Schnittler HJ.** 2011. Ebola virus enters host cells by macropinocytosis and clathrin-mediated endocytosis. *J Infect Dis* **204 Suppl 3**:S957-967.
72. **Saeed MF, Kolokoltsov AA, Albrecht T, Davey RA.** 2010. Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes. *PLoS Pathog* **6**:e1001110.
73. **Chandran K, Sullivan NJ, Felbor U, Whelan SP, Cunningham JM.** 2005. Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* **308**:1643-1645.
74. **Brecher M, Schornberg KL, Delos SE, Fusco ML, Sapphire EO, White JM.** 2012. Cathepsin cleavage potentiates the Ebola virus glycoprotein to undergo a subsequent fusion-relevant conformational change. *J Virol* **86**:364-372.
75. **Schornberg K, Matsuyama S, Kabsch K, Delos S, Bouton A, White J.** 2006. Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J Virol* **80**:4174-4178.
76. **Marzi A, Reinheckel T, Feldmann H.** 2012. Cathepsin B & L are not required for ebola virus replication. *PLoS Negl Trop Dis* **6**:e1923.
77. **Miller EH, Obernosterer G, Raaben M, Herbert AS, Deffieu MS, Krishnan A, Ndungo E, Sandesara RG, Carette JE, Kuehne AI, Ruthel G, Pfeffer SR, Dye JM, Whelan SP, Brummelkamp TR, Chandran K.** 2012. Ebola virus entry requires the host-programmed recognition of an intracellular receptor. *EMBO J* **31**:1947-1960.
78. **Wong AC, Sandesara RG, Mulherkar N, Whelan SP, Chandran K.** 2010. A forward genetic strategy reveals destabilizing mutations in the Ebolavirus

- glycoprotein that alter its protease dependence during cell entry. *J Virol* **84**:163-175.
79. **Bale S, Liu T, Li S, Wang Y, Abelson D, Fusco M, Woods VL, Jr., Saphire EO.** 2011. Ebola virus glycoprotein needs an additional trigger, beyond proteolytic priming for membrane fusion. *PLoS Negl Trop Dis* **5**:e1395.
80. **Lee JE, Saphire EO.** 2009. Ebolavirus glycoprotein structure and mechanism of entry. *Future Virol* **4**:621-635.
81. **Nanbo A, Watanabe S, Halfmann P, Kawaoka Y.** 2013. The spatio-temporal distribution dynamics of Ebola virus proteins and RNA in infected cells. *Sci Rep* **3**:1206.
82. **Luthra P, Jordan DS, Leung DW, Amarasinghe GK, Basler CF.** 2015. Ebola virus VP35 interaction with dynein LC8 regulates viral RNA synthesis. *J Virol* **89**:5148-5153.
83. **Leung DW, Borek D, Luthra P, Binning JM, Anantpadma M, Liu G, Harvey IB, Su Z, Endlich-Frazier A, Pan J, Shabman RS, Chiu W, Davey RA, Otwinowski Z, Basler CF, Amarasinghe GK.** 2015. An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions. *Cell Rep* **11**:376-389.
84. **Modrof J, Moritz C, Kolesnikova L, Konakova T, Hartlieb B, Randolph A, Muhlberger E, Becker S.** 2001. Phosphorylation of Marburg virus VP30 at serines 40 and 42 is critical for its interaction with NP inclusions. *Virology* **287**:171-182.
85. **Kirchdoerfer RN, Moyer CL, Abelson DM, Saphire EO.** 2016. The Ebola Virus VP30-NP Interaction Is a Regulator of Viral RNA Synthesis. *PLoS Pathog* **12**:e1005937.

86. **Weik M, Enterlein S, Schlenz K, Muhlberger E.** 2005. The Ebola virus genomic replication promoter is bipartite and follows the rule of six. *J Virol* **79**:10660-10671.
87. **Green TJ, Cox R, Tsao J, Rowse M, Qiu S, Luo M.** 2014. Common mechanism for RNA encapsidation by negative-strand RNA viruses. *J Virol* **88**:3766-3775.
88. **Wan W, Kolesnikova L, Clarke M, Koehler A, Noda T, Becker S, Briggs JAG.** 2017. Structure and assembly of the Ebola virus nucleocapsid. *Nature* **551**:394-397.
89. **Deflube LR, Cressey TN, Hume AJ, Olejnik J, Haddock E, Feldmann F, Ebihara H, Fearn R, Muhlberger E.** 2019. Ebolavirus polymerase uses an unconventional genome replication mechanism. *Proc Natl Acad Sci U S A* **116**:8535-8543.
90. **Groseth A, Charton JE, Sauerborn M, Feldmann F, Jones SM, Hoenen T, Feldmann H.** 2009. The Ebola virus ribonucleoprotein complex: a novel VP30-L interaction identified. *Virus Res* **140**:8-14.
91. **Weik M, Modrof J, Klenk HD, Becker S, Muhlberger E.** 2002. Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J Virol* **76**:8532-8539.
92. **Kruse T, Biedenkopf N, Hertz EPT, Dietzel E, Stalman G, Lopez-Mendez B, Davey NE, Nilsson J, Becker S.** 2018. The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. *Mol Cell* **69**:136-145 e136.
93. **Lier C, Becker S, Biedenkopf N.** 2017. Dynamic phosphorylation of Ebola virus VP30 in NP-induced inclusion bodies. *Virology* **512**:39-47.
94. **Xu W, Luthra P, Wu C, Batra J, Leung DW, Basler CF, Amarasinghe GK.** 2017. Ebola virus VP30 and nucleoprotein interactions modulate viral RNA synthesis. *Nat Commun* **8**:15576.

95. **Brauburger K, Boehmann Y, Tsuda Y, Hoenen T, Olejnik J, Schumann M, Ebihara H, Muhlberger E.** 2014. Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation. *J Virol* **88**:12558-12571.
96. **Brauburger K, Boehmann Y, Krahling V, Muhlberger E.** 2015. Transcriptional Regulation in Ebola Virus: Effects of Gene Border Structure and Regulatory Elements on Gene Expression and Polymerase Scanning Behavior. *J Virol* **90**:1898-1909.
97. **Stillman EA, Whitt MA.** 1999. Transcript initiation and 5'-end modifications are separable events during vesicular stomatitis virus transcription. *J Virol* **73**:7199-7209.
98. **Shabman RS, Jabado OJ, Mire CE, Stockwell TB, Edwards M, Mahajan M, Geisbert TW, Basler CF.** 2014. Deep sequencing identifies noncanonical editing of Ebola and Marburg virus RNAs in infected cells. *MBio* **5**:e02011.
99. **Muhlberger E, Trommer S, Funke C, Volchkov V, Klenk HD, Becker S.** 1996. Termini of all mRNA species of Marburg virus: sequence and secondary structure. *Virology* **223**:376-380.
100. **Shabman RS, Hoenen T, Groseth A, Jabado O, Binning JM, Amarasinghe GK, Feldmann H, Basler CF.** 2013. An upstream open reading frame modulates ebola virus polymerase translation and virus replication. *PLoS Pathog* **9**:e1003147.
101. **Baskerville A, Fisher-Hoch SP, Neild GH, Dowsett AB.** 1985. Ultrastructural pathology of experimental Ebola haemorrhagic fever virus infection. *J Pathol* **147**:199-209.
102. **Hoenen T, Shabman RS, Groseth A, Herwig A, Weber M, Schudt G, Dolnik O, Basler CF, Becker S, Feldmann H.** 2012. Inclusion bodies are a site of ebolavirus replication. *J Virol* **86**:11779-11788.

103. **Bavari S, Bosio CM, Wiegand E, Ruthel G, Will AB, Geisbert TW, Hevey M, Schmaljohn C, Schmaljohn A, Aman MJ.** 2002. Lipid raft microdomains: a gateway for compartmentalized trafficking of Ebola and Marburg viruses. *J Exp Med* **195**:593-602.
104. **Votteler J, Sundquist WI.** 2013. Virus budding and the ESCRT pathway. *Cell Host Microbe* **14**:232-241.
105. **Younan P, Iampietro M, Santos RI, Ramanathan P, Popov VL, Bukreyev A.** 2018. Role of Transmembrane Protein 16F in the Incorporation of Phosphatidylserine Into Budding Ebola Virus Virions. *J Infect Dis* **218**:S335-S345.
106. **Centers for Disease Control.** 1976. MMWR October 16 1976. Centers for Disease Control,
107. **World Health Organization.** 1976. Weekly Epidemiological Record 15 October 1976. World Health Organization,
108. **Anonymous.** 1978. Ebola haemorrhagic fever in Zaire, 1976. *Bull World Health Organ* **56**:271-293.
109. **Centers for Disease Control.** 1976. MMWR October 29 1976. Centers for Disease Control,
110. **World Health Organization.** 1976. Weekly Epidemiological Record 22 October 1976. World Health Organization,
111. **Kissling RE, Robinson RQ, Murphy FA, Whitfield SG.** 1968. Agent of disease contracted from green monkeys. *Science* **160**:888-890.
112. **Bowen ET, Lloyd G, Harris WJ, Platt GS, Baskerville A, Vella EE.** 1977. Viral haemorrhagic fever in southern Sudan and northern Zaire. Preliminary studies on the aetiological agent. *Lancet* **1**:571-573.

113. **Johnson KM, Lange JV, Webb PA, Murphy FA.** 1977. Isolation and partial characterisation of a new virus causing acute haemorrhagic fever in Zaire. *Lancet* **1**:569-571.
114. **Nanclares C, Kapetshi J, Lionetto F, de la Rosa O, Tamfun JJ, Alia M, Kobinger G, Bernasconi A.** 2016. Ebola Virus Disease, Democratic Republic of the Congo, 2014. *Emerg Infect Dis* **22**:1579-1586.
115. **Khan AS, Tshioko FK, Heymann DL, Le Guenno B, Nabeth P, Kerstiens B, Flerackers Y, Kilmarx PH, Rodier GR, Nkuku O, Rollin PE, Sanchez A, Zaki SR, Swanepoel R, Tomori O, Nichol ST, Peters CJ, Muyembe-Tamfum JJ, Ksiazek TG.** 1999. The reemergence of Ebola hemorrhagic fever, Democratic Republic of the Congo, 1995. Commission de Lutte contre les Epidemies a Kikwit. *J Infect Dis* **179 Suppl 1**:S76-86.
116. **Leroy EM, Epelboin A, Mondonge V, Pourrut X, Gonzalez JP, Muyembe-Tamfum JJ, Formenty P.** 2009. Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector Borne Zoonotic Dis* **9**:723-728.
117. **Georges-Courbot MC, Sanchez A, Lu CY, Baize S, Leroy E, Lansout-Soukate J, Tevi-Benissan C, Georges AJ, Trappier SG, Zaki SR, Swanepoel R, Leman PA, Rollin PE, Peters CJ, Nichol ST, Ksiazek TG.** 1997. Isolation and phylogenetic characterization of Ebola viruses causing different outbreaks in Gabon. *Emerg Infect Dis* **3**:59-62.
118. **Leroy EM, Souquiere S, Rouquet P, Drevet D.** 2002. Re-emergence of ebola haemorrhagic fever in Gabon. *Lancet* **359**:712.
119. **Leroy EM, Rouquet P, Formenty P, Souquiere S, Kilbourne A, Froment JM, Bermejo M, Smit S, Karesh W, Swanepoel R, Zaki SR, Rollin PE.** 2004. Multiple Ebola virus transmission events and rapid decline of central African wildlife. *Science* **303**:387-390.

120. **Georges AJ, Leroy EM, Renaut AA, Benissan CT, Nabias RJ, Ngoc MT, Obiang PI, Lepage JP, Bertherat EJ, Benoni DD, Wickings EJ, Amblard JP, Lansoud-Soukate JM, Milleliri JM, Baize S, Georges-Courbot MC.** 1999. Ebola hemorrhagic fever outbreaks in Gabon, 1994-1997: epidemiologic and health control issues. *J Infect Dis* **179 Suppl 1**:S65-75.
121. **Baize S, Pannetier D, Oestereich L, Rieger T, Koivogui L, Magassouba N, Soropogui B, Sow MS, Keita S, De Clerck H, Tiffany A, Dominguez G, Loua M, Traore A, Kolie M, Malano ER, Heleze E, Bocquin A, Mely S, Raoul H, Caro V, Cadar D, Gabriel M, Pahlmann M, Tappe D, Schmidt-Chanasit J, Impouma B, Diallo AK, Formenty P, Van Herp M, Gunther S.** 2014. Emergence of Zaire Ebola virus disease in Guinea. *N Engl J Med* **371**:1418-1425.
122. **Mari Saez A, Weiss S, Nowak K, Lapeyre V, Zimmermann F, Dux A, Kuhl HS, Kaba M, Regnaut S, Merkel K, Sachse A, Thiesen U, Villanyi L, Boesch C, Dabrowski PW, Radonic A, Nitsche A, Leendertz SA, Petterson S, Becker S, Krahling V, Couacy-Hymann E, Akoua-Koffi C, Weber N, Schaade L, Fahr J, Borchert M, Gogarten JF, Calvignac-Spencer S, Leendertz FH.** 2015. Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol Med* **7**:17-23.
123. **Goba A, Khan SH, Fonnio M, Fullah M, Moigboi A, Kovoma A, Sinnah V, Yoko N, Rogers H, Safai S, Momoh M, Koroma V, Kamara FK, Konowu E, Yillah M, French I, Mustapha I, Kanneh F, Foday M, McCarthy H, Kallon T, Kallon M, Naiebu J, Sellu J, Jalloh AA, Gbakie M, Kanneh L, Massaly JL, Kargbo D, Kargbo B, Vandi M, Gbetuwa M, Gevao SM, Sandi JD, Jalloh SC, Grant DS, Blyden SO, Crozier I, Schieffelin JS, McLellan SL, Jacob ST, Boisen ML, Hartnett JN, Cross RW, Branco LM, Andersen KG, Yozwiak NL, Gire SK, Tariyal R, Park DJ, et al.** 2016. An Outbreak of Ebola Virus Disease in the Lassa Fever Zone. *J Infect Dis* **214**:S110-S121.

124. **Shultz JM, Cooper JL, Baingana F, Oquendo MA, Espinel Z, Althouse BM, Marcelin LH, Towers S, Espinola M, McCoy CB, Mazurik L, Wainberg ML, Neria Y, Rechkemmer A.** 2016. The Role of Fear-Related Behaviors in the 2013-2016 West Africa Ebola Virus Disease Outbreak. *Curr Psychiatry Rep* **18**:104.
125. **Henao-Restrepo AM, Camacho A, Longini IM, Watson CH, Edmunds WJ, Egger M, Carroll MW, Dean NE, Diatta I, Doumbia M, Draguez B, Duraffour S, Enwere G, Grais R, Gunther S, Gsell PS, Hossmann S, Watle SV, Konde MK, Keita S, Kone S, Kuisma E, Levine MM, Mandal S, Mauget T, Norheim G, Riveros X, Soumah A, Trelle S, Vicari AS, Rottingen JA, Kieny MP.** 2017. Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ca Suffit!). *Lancet* **389**:505-518.
126. **Reiter P, Turell M, Coleman R, Miller B, Maupin G, Liz J, Kuehne A, Barth J, Geisbert J, Dohm D, Glick J, Pecor J, Robbins R, Jahrling P, Peters C, Ksiazek T.** 1999. Field investigations of an outbreak of Ebola hemorrhagic fever, Kikwit, Democratic Republic of the Congo, 1995: arthropod studies. *J Infect Dis* **179 Suppl 1**:S148-154.
127. **Palich R, Irengue LM, Barte de Sainte Fare E, Augier A, Malvy D, Gala JL.** 2017. Ebola virus RNA detection on fomites in close proximity to confirmed Ebola patients; N'Zerekore, Guinea, 2015. *PLoS One* **12**:e0177350.
128. **Vetter P, Fischer WA, 2nd, Schibler M, Jacobs M, Bausch DG, Kaiser L.** 2016. Ebola Virus Shedding and Transmission: Review of Current Evidence. *J Infect Dis* **214**:S177-S184.
129. **Jaax NK, Davis KJ, Geisbert TJ, Vogel P, Jaax GP, Topper M, Jahrling PB.** 1996. Lethal experimental infection of rhesus monkeys with Ebola-Zaire

(Mayinga) virus by the oral and conjunctival route of exposure. *Arch Pathol Lab Med* **120**:140-155.

130. **Mire CE, Geisbert JB, Agans KN, Deer DJ, Fenton KA, Geisbert TW.** 2016. Oral and Conjunctival Exposure of Nonhuman Primates to Low Doses of Ebola Makona Virus. *J Infect Dis* **214**:S263-S267.
131. **Osterholm MT, Moore KA, Kelley NS, Brosseau LM, Wong G, Murphy FA, Peters CJ, LeDuc JW, Russell PK, Van Herp M, Kapetshi J, Muyembe JJ, Ilunga BK, Strong JE, Grolla A, Wolz A, Kargbo B, Kargbo DK, Formenty P, Sanders DA, Kobinger GP.** 2015. Transmission of Ebola viruses: what we know and what we do not know. *MBio* **6**:e00137.
132. **Park DJ, Dudas G, Wohl S, Goba A, Whitmer SL, Andersen KG, Sealfon RS, Ladner JT, Kugelman JR, Matranga CB, Winnicki SM, Qu J, Gire SK, Gladden-Young A, Jalloh S, Nosamiefan D, Yozwiak NL, Moses LM, Jiang PP, Lin AE, Schaffner SF, Bird B, Towner J, Mamoh M, Gbakie M, Kanneh L, Kargbo D, Massally JL, Kamara FK, Konuwa E, Sellu J, Jalloh AA, Mustapha I, Foday M, Yillah M, Erickson BR, Sealy T, Blau D, Paddock C, Brault A, Amman B, Basile J, Bearden S, Belser J, Bergeron E, Campbell S, Chakrabarti A, Dodd K, Flint M, Gibbons A, et al.** 2015. Ebola Virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone. *Cell* **161**:1516-1526.
133. **Becquart P, Wauquier N, Mahlakoiv T, Nkoghe D, Padilla C, Souris M, Ollomo B, Gonzalez JP, De Lamballerie X, Kazanji M, Leroy EM.** 2010. High prevalence of both humoral and cellular immunity to Zaire ebolavirus among rural populations in Gabon. *PLoS One* **5**:e9126.
134. **Richardson ET, Kelly JD, Barrie MB, Mesman AW, Karku S, Quiwa K, Marsh RH, Koedoyoma S, Daboh F, Barron KP, Grady M, Tucker E, Dierberg KL, Rutherford GW, Barry M, Jones JH, Murray MB, Farmer PE.**

2016. Minimally Symptomatic Infection in an Ebola 'Hotspot': A Cross-Sectional Serosurvey. *PLoS Negl Trop Dis* **10**:e0005087.
135. **Hoff NA, Mukadi P, Doshi RH, Bramble MS, Lu K, Gadoth A, Sinai C, Spencer D, Nicholson BP, Williams R, Mossoko M, Ilunga-Kebela B, Wasiswa J, Okitolonda-Wemakoy E, Alfonso VH, Steffen I, Muyembe-Tamfum JJ, Simmons G, Rimoin AW.** 2019. Serologic Markers for Ebolavirus Among Healthcare Workers in the Democratic Republic of the Congo. *J Infect Dis* **219**:517-525.
136. **Bower H, Glynn JR.** 2017. A systematic review and meta-analysis of seroprevalence surveys of ebolavirus infection. *Sci Data* **4**:160133.
137. **Hewlett BS, Hewlett BL.** 2008. Ebola, culture, and politics : the anthropology of an emerging disease. Thomson, Belmont, CA.
138. **Feldmann H.** 2018. Virus in Semen and the Risk of Sexual Transmission. *N Engl J Med* **378**:1440-1441.
139. **Guerrier G, D'Ortenzio E.** 2014. Africa: Ebola virus control needs local buy-in. *Nature* **513**:315.
140. **Cancedda C, Davis SM, Dierberg KL, Lascher J, Kelly JD, Barrie MB, Koroma AP, George P, Kamara AA, Marsh R, Sumbuya MS, Nutt CT, Scott KW, Thomas E, Bollbach K, Sesay A, Barrie A, Barrera E, Barron K, Welch J, Bhadelia N, Frankfurter RG, Dahl OM, Das S, Rollins RE, Eustis B, Schwartz A, Pertile P, Pavlopoulos I, Mayfield A, Marsh RH, Dibba Y, Kloepper D, Hall A, Huster K, Grady M, Spray K, Walton DA, Daboh F, Nally C, James S, Warren GS, Chang J, Drasher M, Lamin G, Bangura S, Miller AC, Michaelis AP, McBain R, Broadhurst MJ, et al.** 2016. Strengthening Health Systems While Responding to a Health Crisis: Lessons Learned by a Nongovernmental Organization During the Ebola Virus Disease Epidemic in Sierra Leone. *J Infect Dis* **214**:S153-S163.

141. **Coller BG, Blue J, Das R, Dubey S, Finelli L, Gupta S, Helmond F, Grant-Klein RJ, Liu K, Simon J, Troth S, VanRheenen S, Waterbury J, Wivel A, Wolf J, Heppner DG, Kemp T, Nichols R, Monath TP.** 2017. Clinical development of a recombinant Ebola vaccine in the midst of an unprecedented epidemic. *Vaccine* **35**:4465-4469.
142. **Regules JA, Beigel JH, Paolino KM, Voell J, Castellano AR, Hu Z, Munoz P, Moon JE, Ruck RC, Bennett JW, Twomey PS, Gutierrez RL, Remich SA, Hack HR, Wisniewski ML, Josleyn MD, Kwilas SA, Van Deusen N, Mbaya OT, Zhou Y, Stanley DA, Jing W, Smith KS, Shi M, Ledgerwood JE, Graham BS, Sullivan NJ, Jagodzinski LL, Peel SA, Alimonti JB, Hooper JW, Silvera PM, Martin BK, Monath TP, Ramsey WJ, Link CJ, Lane HC, Michael NL, Davey RT, Jr., Thomas SJ, r V-Z-GPSG.** 2017. A Recombinant Vesicular Stomatitis Virus Ebola Vaccine. *N Engl J Med* **376**:330-341.
143. **U.S. Food and Drug Administration.** 2019. First FDA-approved vaccine for the prevention of Ebola virus disease, marking a critical milestone in public health preparedness and response. U.S. Food and Drug Administration.
144. **Stehling-Ariza T, Rosewell A, Moiba SA, Yorpie BB, Ndomaina KD, Jimissa KS, Leidman E, Rijken DJ, Basler C, Wood J, Manso D.** 2016. The impact of active surveillance and health education on an Ebola virus disease cluster - Kono District, Sierra Leone, 2014-2015. *BMC Infect Dis* **16**:611.
145. **Cohen J.** 2018. Congo rapidly curtails Ebola. *Science* **361**:211-212.
146. **Mbala-Kingebeni P, Aziza A, Di Paola N, Wiley MR, Makiala-Mandanda S, Caviness K, Pratt CB, Ladner JT, Kugelman JR, Prieto K, Chitty JA, Larson PA, Beitzel B, Ayouba A, Vidal N, Karhemere S, Diop M, Diagne MM, Faye M, Faye O, Aruna A, Nsio J, Mulangu F, Mukadi D, Mukadi P, Kombe J, Mulumba A, Villabona-Arenas CJ, Pukuta E, Gonzalez J, Bartlett ML, Sozhamannan S, Gross SM, Schroth GP, Tim R, Zhao JJ, Kuhn JH,**

- Diallo B, Yao M, Fall IS, Ndjoloko B, Mossoko M, Lacroix A, Delaporte E, Sanchez-Lockhart M, Sall AA, Muyembe-Tamfum JJ, Peeters M, Palacios G, Ahuka-Mundeke S.** 2019. Medical countermeasures during the 2018 Ebola virus disease outbreak in the North Kivu and Ituri Provinces of the Democratic Republic of the Congo: a rapid genomic assessment. *Lancet Infect Dis* doi:10.1016/S1473-3099(19)30118-5.
147. **Holt K, Ratcliffe R.** 2/12/19 2019. Ebola vaccine offered in exchange for sex, Congo taskforce meeting told, p *In* The Guardian. Manchester, UK.
148. **Geisbert TW, Hensley LE, Larsen T, Young HA, Reed DS, Geisbert JB, Scott DP, Kagan E, Jahrling PB, Davis KJ.** 2003. Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. *Am J Pathol* **163**:2347-2370.
149. **Bray M, Geisbert TW.** 2005. Ebola virus: the role of macrophages and dendritic cells in the pathogenesis of Ebola hemorrhagic fever. *Int J Biochem Cell Biol* **37**:1560-1566.
150. **Martinez O, Johnson JC, Honko A, Yen B, Shabman RS, Hensley LE, Olinger GG, Basler CF.** 2013. Ebola virus exploits a monocyte differentiation program to promote its entry. *J Virol* **87**:3801-3814.
151. **Bosio CM, Aman MJ, Grogan C, Hogan R, Ruthel G, Negley D, Mohamadzadeh M, Bavari S, Schmaljohn A.** 2003. Ebola and Marburg viruses replicate in monocyte-derived dendritic cells without inducing the production of cytokines and full maturation. *J Infect Dis* **188**:1630-1638.
152. **Mahanty S, Hutchinson K, Agarwal S, McRae M, Rollin PE, Pulendran B.** 2003. Cutting edge: impairment of dendritic cells and adaptive immunity by Ebola and Lassa viruses. *J Immunol* **170**:2797-2801.
153. **Lubaki NM, Ilinykh P, Pietzsch C, Tigabu B, Freiberg AN, Koup RA, Bukreyev A.** 2013. The lack of maturation of Ebola virus-infected dendritic cells

- results from the cooperative effect of at least two viral domains. *J Virol* **87**:7471-7485.
154. **Misasi J, Sullivan NJ.** 2014. Camouflage and misdirection: the full-on assault of ebola virus disease. *Cell* **159**:477-486.
155. **Reed DS, Hensley LE, Geisbert JB, Jahrling PB, Geisbert TW.** 2004. Depletion of peripheral blood T lymphocytes and NK cells during the course of ebola hemorrhagic Fever in cynomolgus macaques. *Viral Immunol* **17**:390-400.
156. **Bradfute SB, Braun DR, Shamblin JD, Geisbert JB, Paragas J, Garrison A, Hensley LE, Geisbert TW.** 2007. Lymphocyte death in a mouse model of Ebola virus infection. *J Infect Dis* **196 Suppl 2**:S296-304.
157. **Younan P, Santos RI, Ramanathan P, Iampietro M, Nishida A, Dutta M, Ammosova T, Meyer M, Katze MG, Popov VL, Nekhai S, Bukreyev A.** 2019. Ebola virus-mediated T-lymphocyte depletion is the result of an abortive infection. *PLoS Pathog* **15**:e1008068.
158. **Iampietro M, Younan P, Nishida A, Dutta M, Lubaki NM, Santos RI, Koup RA, Katze MG, Bukreyev A.** 2017. Ebola virus glycoprotein directly triggers T lymphocyte death despite of the lack of infection. *PLoS Pathog* **13**:e1006397.
159. **Younan P, Iampietro M, Nishida A, Ramanathan P, Santos RI, Dutta M, Lubaki NM, Koup RA, Katze MG, Bukreyev A.** 2017. Ebola Virus Binding to Tim-1 on T Lymphocytes Induces a Cytokine Storm. *MBio* **8**.
160. **Wolf K, Beimforde N, Falzarano D, Feldmann H, Schnittler HJ.** 2011. The Ebola virus soluble glycoprotein (sGP) does not affect lymphocyte apoptosis and adhesion to activated endothelium. *J Infect Dis* **204 Suppl 3**:S947-952.
161. **Lubaki NM, Younan P, Santos RI, Meyer M, Iampietro M, Koup RA, Bukreyev A.** 2016. The Ebola Interferon Inhibiting Domains Attenuate and Dysregulate Cell-Mediated Immune Responses. *PLoS Pathog* **12**:e1006031.

162. **Younan P, Iampietro M, Bukreyev A.** 2018. Disabling of lymphocyte immune response by Ebola virus. *PLoS Pathog* **14**:e1006932.
163. **Sharma N, Cappell MS.** 2015. Gastrointestinal and Hepatic Manifestations of Ebola Virus Infection. *Dig Dis Sci* **60**:2590-2603.
164. **Baskerville A, Bowen ET, Platt GS, McArdell LB, Simpson DI.** 1978. The pathology of experimental Ebola virus infection in monkeys. *J Pathol* **125**:131-138.
165. **Yeh S, Varkey JB, Crozier I.** 2015. Persistent Ebola Virus in the Eye. *N Engl J Med* **373**:1982-1983.
166. **Olejniak J, Forero A, Deflube LR, Hume AJ, Manhart WA, Nishida A, Marzi A, Katze MG, Ebihara H, Rasmussen AL, Muhlberger E.** 2017. Ebolaviruses Associated with Differential Pathogenicity Induce Distinct Host Responses in Human Macrophages. *J Virol* **91**.
167. **Zampieri CA, Sullivan NJ, Nabel GJ.** 2007. Immunopathology of highly virulent pathogens: insights from Ebola virus. *Nat Immunol* **8**:1159-1164.
168. **Kreuels B, Wichmann D, Emmerich P, Schmidt-Chanasit J, de Heer G, Kluge S, Sow A, Renne T, Gunther S, Lohse AW, Addo MM, Schmiedel S.** 2014. A case of severe Ebola virus infection complicated by gram-negative septicemia. *N Engl J Med* **371**:2394-2401.
169. **Liddell AM, Davey RT, Jr., Mehta AK, Varkey JB, Kraft CS, Tseggay GK, Badidi O, Faust AC, Brown KV, Suffredini AF, Barrett K, Wolcott MJ, Marconi VC, Lyon GM, 3rd, Weinstein GL, Weinmeister K, Sutton S, Hazbun M, Albarino CG, Reed Z, Cannon D, Stroher U, Feldman M, Ribner BS, Lane HC, Fauci AS, Uyeki TM.** 2015. Characteristics and Clinical Management of a Cluster of 3 Patients With Ebola Virus Disease, Including the First Domestically Acquired Cases in the United States. *Ann Intern Med* **163**:81-90.

170. **Geisbert TW, Young HA, Jahrling PB, Davis KJ, Kagan E, Hensley LE.** 2003. Mechanisms underlying coagulation abnormalities in ebola hemorrhagic fever: overexpression of tissue factor in primate monocytes/macrophages is a key event. *J Infect Dis* **188**:1618-1629.
171. **Geisbert TW, Young HA, Jahrling PB, Davis KJ, Larsen T, Kagan E, Hensley LE.** 2003. Pathogenesis of Ebola hemorrhagic fever in primate models: evidence that hemorrhage is not a direct effect of virus-induced cytolysis of endothelial cells. *Am J Pathol* **163**:2371-2382.
172. **Boral BM, Williams DJ, Boral LI.** 2016. Disseminated Intravascular Coagulation. *Am J Clin Pathol* **146**:670-680.
173. **Schieffelin JS, Shaffer JG, Goba A, Gbakie M, Gire SK, Colubri A, Sealfon RS, Kanneh L, Moigboi A, Momoh M, Fullah M, Moses LM, Brown BL, Andersen KG, Winnicki S, Schaffner SF, Park DJ, Yozwiak NL, Jiang PP, Kargbo D, Jalloh S, Fonnies M, Sinnah V, French I, Kovoma A, Kamara FK, Tucker V, Konuwa E, Sellu J, Mustapha I, Foday M, Yillah M, Kanneh F, Saffa S, Massally JL, Boisen ML, Branco LM, Vandi MA, Grant DS, Happi C, Gevao SM, Fletcher TE, Fowler RA, Bausch DG, Sabeti PC, Khan SH, Garry RF, Program KGHLE, Viral Hemorrhagic Fever C, Team WHO-CR.** 2014. Clinical illness and outcomes in patients with Ebola in Sierra Leone. *N Engl J Med* **371**:2092-2100.
174. **Cross RW, Mire CE, Borisevich V, Geisbert JB, Fenton KA, Geisbert TW.** 2016. The Domestic Ferret (*Mustela putorius furo*) as a Lethal Infection Model for 3 Species of Ebolavirus. *J Infect Dis* **214**:565-569.
175. **Hensley LE, Alves DA, Geisbert JB, Fritz EA, Reed C, Larsen T, Geisbert TW.** 2011. Pathogenesis of Marburg hemorrhagic fever in cynomolgus macaques. *J Infect Dis* **204 Suppl 3**:S1021-1031.
176. **Haas CN.** 2014. On the quarantine period for ebola virus. *PLoS Curr* **6**.

177. **Lado M, Walker NF, Baker P, Haroon S, Brown CS, Youkee D, Studd N, Kessete Q, Maini R, Boyles T, Hanciles E, Wurie A, Kamara TB, Johnson O, Leather AJM.** 2015. Clinical features of patients isolated for suspected Ebola virus disease at Connaught Hospital, Freetown, Sierra Leone: a retrospective cohort study. *Lancet Infect Dis* **15**:1024-1033.
178. **Hartley MA, Young A, Tran AM, Okoni-Williams HH, Suma M, Mancuso B, Al-Dikhari A, Faouzi M.** 2017. Predicting Ebola Severity: A Clinical Prioritization Score for Ebola Virus Disease. *PLoS Negl Trop Dis* **11**:e0005265.
179. **Uyeki TM, Mehta AK, Davey RT, Jr., Liddell AM, Wolf T, Vetter P, Schmiedel S, Grunewald T, Jacobs M, Arribas JR, Evans L, Hewlett AL, Brantsaeter AB, Ippolito G, Rapp C, Hoepelman AI, Gutman J, Working Group of the USECNoCMoEVDpitUS, Europe.** 2016. Clinical Management of Ebola Virus Disease in the United States and Europe. *N Engl J Med* **374**:636-646.
180. **Burki TK.** 2016. Post-Ebola syndrome. *Lancet Infect Dis* **16**:780-781.
181. **Jacobs M, Rodger A, Bell DJ, Bhagani S, Cropley I, Filipe A, Gifford RJ, Hopkins S, Hughes J, Jabeen F, Johannessen I, Karageorgopoulos D, Lackenby A, Lester R, Liu RS, MacConnachie A, Mahungu T, Martin D, Marshall N, Mephram S, Orton R, Palmarini M, Patel M, Perry C, Peters SE, Porter D, Ritchie D, Ritchie ND, Seaton RA, Sreenu VB, Templeton K, Warren S, Wilkie GS, Zambon M, Gopal R, Thomson EC.** 2016. Late Ebola virus relapse causing meningoencephalitis: a case report. *Lancet* **388**:498-503.
182. **Boisen ML, Schieffelin JS, Goba A, Oottamasathien D, Jones AB, Shaffer JG, Hastie KM, Hartnett JN, Momoh M, Fullah M, Gabiki M, Safa S, Zandonatti M, Fusco M, Bornholdt Z, Abelson D, Gire SK, Andersen KG, Tariyal R, Stremlau M, Cross RW, Geisbert JB, Pitts KR, Geisbert TW, Kulakoski P, Wilson RB, Henderson L, Sabeti PC, Grant DS, Garry RF, Sapphire EO, Branco LM, Khan SH, Viral Hemorrhagic Fever C.** 2015.

Multiple circulating infections can mimic the early stages of viral hemorrhagic fevers and possible human exposure to filoviruses in Sierra Leone prior to the 2014 outbreak. *Viral Immunol* **28**:19-31.

183. **de Wit E, Falzarano D, Onyango C, Rosenke K, Marzi A, Ochieng M, Juma B, Fischer RJ, Prescott JB, Safronetz D, Omballa V, Owuor C, Hoenen T, Groseth A, van Doremalen N, Zemtsova G, Self J, Bushmaker T, McNally K, Rowe T, Emery SL, Feldmann F, Williamson B, Nyenswah TG, Grolla A, Strong JE, Kobinger G, Strocher U, Rayfield M, Bolay FK, Zoon KC, Stassijns J, Tampellini L, de Smet M, Nichol ST, Fields B, Sprecher A, Feldmann H, Massaquoi M, Munster VJ.** 2016. The Merits of Malaria Diagnostics during an Ebola Virus Disease Outbreak. *Emerg Infect Dis* **22**:323-326.
184. **de Wit E, Rosenke K, Fischer RJ, Marzi A, Prescott J, Bushmaker T, van Doremalen N, Emery SL, Falzarano D, Feldmann F, Groseth A, Hoenen T, Juma B, McNally KL, Ochieng M, Omballa V, Onyango CO, Owuor C, Rowe T, Safronetz D, Self J, Williamson BN, Zemtsova G, Grolla A, Kobinger G, Rayfield M, Stroher U, Strong JE, Best SM, Ebihara H, Zoon KC, Nichol ST, Nyenswah TG, Bolay FK, Massaquoi M, Feldmann H, Fields B.** 2016. Ebola Laboratory Response at the Eternal Love Winning Africa Campus, Monrovia, Liberia, 2014-2015. *J Infect Dis* **214**:S169-S176.
185. **Diarra B, Safronetz D, Sarro YD, Kone A, Sanogo M, Tounkara S, Togo AC, Daou F, Maiga AI, Dao S, Rosenke K, Falzarano D, Doumbia S, Zoon KC, Polis M, Siddiqui S, Sow S, Schwan TG, Feldmann H, Diallo S, Koita OA.** 2016. Laboratory Response to 2014 Ebola Virus Outbreak in Mali. *J Infect Dis* **214**:S164-S168.
186. **Kilgore PE, Grabenstein JD, Salim AM, Rybak M.** 2015. Treatment of ebola virus disease. *Pharmacotherapy* **35**:43-53.

187. **Rojek A, Horby P, Dunning J.** 2017. Insights from clinical research completed during the west Africa Ebola virus disease epidemic. *Lancet Infect Dis* **17**:e280-e292.
188. **Dunning J, Sahr F, Rojek A, Gannon F, Carson G, Idriss B, Massaquoi T, Gandhi R, Joseph S, Osman HK, Brooks TJ, Simpson AJ, Goodfellow I, Thorne L, Arias A, Merson L, Castle L, Howell-Jones R, Pardinaz-Solis R, Hope-Gill B, Ferri M, Grove J, Kowalski M, Stepniewska K, Lang T, Whitehead J, Oliaro P, Samai M, Horby PW, team R-Tt.** 2016. Experimental Treatment of Ebola Virus Disease with TKM-130803: A Single-Arm Phase 2 Clinical Trial. *PLoS Med* **13**:e1001997.
189. **Furuyama W, Marzi A, Nanbo A, Haddock E, Maruyama J, Miyamoto H, Igarashi M, Yoshida R, Noyori O, Feldmann H, Takada A.** 2016. Discovery of an antibody for pan-ebolavirus therapy. *Sci Rep* **6**:20514.
190. **Group PIW, Multi-National PIIST, Davey RT, Jr., Dodd L, Proschan MA, Neaton J, Neuhaus Nordwall J, Koopmeiners JS, Beigel J, Tierney J, Lane HC, Fauci AS, Massaquoi MBF, Sahr F, Malvy D.** 2016. A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection. *N Engl J Med* **375**:1448-1456.
191. **Qiu X, Audet J, Lv M, He S, Wong G, Wei H, Luo L, Fernando L, Kroeker A, Fausther Bovendo H, Bello A, Li F, Ye P, Jacobs M, Ippolito G, Sapphire EO, Bi S, Shen B, Gao GF, Zeitlin L, Feng J, Zhang B, Kobinger GP.** 2016. Two-mAb cocktail protects macaques against the Makona variant of Ebola virus. *Sci Transl Med* **8**:329ra333.
192. **Flyak AI, Kuzmina N, Murin CD, Bryan C, Davidson E, Gilchuk P, Gulka CP, Ilinykh PA, Shen X, Huang K, Ramanathan P, Turner H, Fusco ML, Lampley R, Kose N, King H, Sapparapu G, Doranz BJ, Ksiazek TG, Wright DW, Sapphire EO, Ward AB, Bukreyev A, Crowe JE, Jr.** 2018. Broadly

neutralizing antibodies from human survivors target a conserved site in the Ebola virus glycoprotein HR2-MPER region. *Nat Microbiol* **3**:670-677.

193. **Parren PW, Geisbert TW, Maruyama T, Jahrling PB, Burton DR.** 2002. Pre- and postexposure prophylaxis of Ebola virus infection in an animal model by passive transfer of a neutralizing human antibody. *J Virol* **76**:6408-6412.
194. **Jahrling PB, Geisbert JB, Swearingen JR, Larsen T, Geisbert TW.** 2007. Ebola hemorrhagic fever: evaluation of passive immunotherapy in nonhuman primates. *J Infect Dis* **196 Suppl 2**:S400-403.
195. **Oswald WB, Geisbert TW, Davis KJ, Geisbert JB, Sullivan NJ, Jahrling PB, Parren PW, Burton DR.** 2007. Neutralizing antibody fails to impact the course of Ebola virus infection in monkeys. *PLoS Pathog* **3**:e9.
196. **Takada A, Ebihara H, Jones S, Feldmann H, Kawaoka Y.** 2007. Protective efficacy of neutralizing antibodies against Ebola virus infection. *Vaccine* **25**:993-999.
197. **Basler CF.** 2014. New hope in the search for Ebola virus treatments. *Immunity* **41**:515-517.
198. **Mire CE, Geisbert JB, Agans KN, Thi EP, Lee AC, Fenton KA, Geisbert TW.** 2016. Passive Immunotherapy: Assessment of Convalescent Serum Against Ebola Virus Makona Infection in Nonhuman Primates. *J Infect Dis* **214**:S367-S374.
199. **Centers for Disease Control.** 2014. Guidance on Personal Protective Equipment (PPE) To Be Used By Healthcare Workers during Management of Patients with Confirmed Ebola or Persons under Investigation (PUIs) for Ebola who are Clinically Unstable or Have Bleeding, Vomiting, or Diarrhea in U.S. Hospitals, Including Procedures for Donning and Doffing PPE. Services HaH, Centers for Disease Control,, Atlanta, Georgia.

200. **Wichmann D, Schmiedel S, Kluge S.** 2015. Isolation in patients with Ebola virus disease. *Intensive Care Med* **41**:511-513.
201. **Rouquet P, Froment JM, Bermejo M, Kilbourn A, Karesh W, Reed P, Kumulungui B, Yaba P, Delicat A, Rollin PE, Leroy EM.** 2005. Wild animal mortality monitoring and human Ebola outbreaks, Gabon and Republic of Congo, 2001-2003. *Emerg Infect Dis* **11**:283-290.
202. **Swanepoel R, Leman PA, Burt FJ, Zachariades NA, Braack LE, Ksiazek TG, Rollin PE, Zaki SR, Peters CJ.** 1996. Experimental inoculation of plants and animals with Ebola virus. *Emerg Infect Dis* **2**:321-325.
203. **Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, Delicat A, Paweska JT, Gonzalez JP, Swanepoel R.** 2005. Fruit bats as reservoirs of Ebola virus. *Nature* **438**:575-576.
204. **Hayman DT, Yu M, Crameri G, Wang LF, Suu-Ire R, Wood JL, Cunningham AA.** 2012. Ebola virus antibodies in fruit bats, Ghana, West Africa. *Emerg Infect Dis* **18**:1207-1209.
205. **Olival KJ, Hayman DT.** 2014. Filoviruses in bats: current knowledge and future directions. *Viruses* **6**:1759-1788.
206. **Ogawa H, Miyamoto H, Nakayama E, Yoshida R, Nakamura I, Sawa H, Ishii A, Thomas Y, Nakagawa E, Matsuno K, Kajihara M, Maruyama J, Nao N, Muramatsu M, Kuroda M, Simulundu E, Changula K, Hang'ombe B, Namangala B, Nambota A, Katampi J, Igarashi M, Ito K, Feldmann H, Sugimoto C, Moonga L, Mweene A, Takada A.** 2015. Seroepidemiological Prevalence of Multiple Species of Filoviruses in Fruit Bats (*Eidolon helvum*) Migrating in Africa. *J Infect Dis* **212 Suppl 2**:S101-108.
207. **Leendertz SA, Gogarten JF, Dux A, Calvignac-Spencer S, Leendertz FH.** 2016. Assessing the Evidence Supporting Fruit Bats as the Primary Reservoirs for Ebola Viruses. *Ecohealth* **13**:18-25.

208. **Ramos AM.** 2019. ECOHEALTH ALLIANCE SCIENTISTS DISCOVER THE DEADLY ZAIRE EBOLAVIRUS IN WEST AFRICAN BAT. EcoHealth Alliance, New York.
209. **Towner JS, Amman BR, Sealy TK, Carroll SA, Comer JA, Kemp A, Swanepoel R, Paddock CD, Balinandi S, Khristova ML, Formenty PB, Albarino CG, Miller DM, Reed ZD, Kayiwa JT, Mills JN, Cannon DL, Greer PW, Byaruhanga E, Farnon EC, Atimnedi P, Okware S, Katongole-Mbidde E, Downing R, Tappero JW, Zaki SR, Ksiazek TG, Nichol ST, Rollin PE.** 2009. Isolation of genetically diverse Marburg viruses from Egyptian fruit bats. PLoS Pathog **5**:e1000536.
210. **Jones ME, Schuh AJ, Amman BR, Sealy TK, Zaki SR, Nichol ST, Towner JS.** 2015. Experimental Inoculation of Egyptian Rousette Bats (*Rousettus aegyptiacus*) with Viruses of the Ebolavirus and Marburgvirus Genera. Viruses **7**:3420-3442.
211. **Paweska JT, Storm N, Grobbelaar AA, Markotter W, Kemp A, Jansen van Vuren P.** 2016. Experimental Inoculation of Egyptian Fruit Bats (*Rousettus aegyptiacus*) with Ebola Virus. Viruses **8**.
212. **Kuzmin IV, Schwarz TM, Ilinykh PA, Jordan I, Ksiazek TG, Sachidanandam R, Basler CF, Bukreyev A.** 2017. Innate Immune Responses of Bat and Human Cells to Filoviruses: Commonalities and Distinctions. J Virol **91**.
213. **Bratsch S, Wertz N, Chaloner K, Kunz TH, Butler JE.** 2011. The little brown bat, *M. lucifugus*, displays a highly diverse V H, D H and J H repertoire but little evidence of somatic hypermutation. Dev Comp Immunol **35**:421-430.
214. **Kunz TH, Racey PA.** 1998. Bat biology and conservation. Smithsonian Institution Press, Washington.
215. **Kunz TH, Fenton MB.** 2003. Bat ecology. University of Chicago Press, Chicago, Ill.

216. **Ng M, Ndungo E, Kaczmarek ME, Herbert AS, Binger T, Kuehne AI, Jangra RK, Hawkins JA, Gifford RJ, Biswas R, Demogines A, James RM, Yu M, Brummelkamp TR, Drosten C, Wang LF, Kuhn JH, Muller MA, Dye JM, Sawyer SL, Chandran K.** 2015. Filovirus receptor NPC1 contributes to species-specific patterns of ebolavirus susceptibility in bats. *Elife* **4**.
217. **Olson SH, Bounga G, Ondzie A, Bushmaker T, Seifert SN, Kuisma E, Taylor DW, Munster VJ, Walzer C.** 2019. Lek-associated movement of a putative Ebolavirus reservoir, the hammer-headed fruit bat (*Hypsignathus monstrosus*), in northern Republic of Congo. *PLoS One* **14**:e0223139.
218. **Amman BR, Jones ME, Sealy TK, Uebelhoer LS, Schuh AJ, Bird BH, Coleman-McCray JD, Martin BE, Nichol ST, Towner JS.** 2015. Oral shedding of Marburg virus in experimentally infected Egyptian fruit bats (*Rousettus aegyptiacus*). *J Wildl Dis* **51**:113-124.
219. **Paweska JT, Jansen van Vuren P, Fenton KA, Graves K, Grobbelaar AA, Moolla N, Leman P, Weyer J, Storm N, McCulloch SD, Scott TP, Markotter W, Odendaal L, Clift SJ, Geisbert TW, Hale MJ, Kemp A.** 2015. Lack of Marburg Virus Transmission From Experimentally Infected to Susceptible In-Contact Egyptian Fruit Bats. *J Infect Dis* **212 Suppl 2**:S109-118.
220. **Hayman DT.** 2015. Biannual birth pulses allow filoviruses to persist in bat populations. *Proc Biol Sci* **282**:20142591.
221. **Olivero J, Fa JE, Real R, Marquez AL, Farfan MA, Vargas JM, Gaveau D, Salim MA, Park D, Suter J, King S, Leendertz SA, Sheil D, Nasi R.** 2017. Recent loss of closed forests is associated with Ebola virus disease outbreaks. *Sci Rep* **7**:14291.
222. **Holmes EC.** 2003. Molecular clocks and the puzzle of RNA virus origins. *J Virol* **77**:3893-3897.

223. **Carroll SA, Towner JS, Sealy TK, McMullan LK, Khristova ML, Burt FJ, Swanepoel R, Rollin PE, Nichol ST.** 2013. Molecular evolution of viruses of the family Filoviridae based on 97 whole-genome sequences. *J Virol* **87**:2608-2616.
224. **Belyi VA, Levine AJ, Skalka AM.** 2010. Unexpected inheritance: multiple integrations of ancient bornavirus and ebolavirus/marburgvirus sequences in vertebrate genomes. *PLoS Pathog* **6**:e1001030.
225. **Taylor DJ, Leach RW, Bruenn J.** 2010. Filoviruses are ancient and integrated into mammalian genomes. *BMC Evol Biol* **10**:193.
226. **Taylor DJ, Dittmar K, Ballinger MJ, Bruenn JA.** 2011. Evolutionary maintenance of filovirus-like genes in bat genomes. *BMC Evol Biol* **11**:336.
227. **Taylor DJ, Ballinger MJ, Zhan JJ, Hanzly LE, Bruenn JA.** 2014. Evidence that ebolaviruses and cuevaviruses have been diverging from marburgviruses since the Miocene. *PeerJ* **2**:e556.
228. **Wertheim JO, Kosakovsky Pond SL.** 2011. Purifying selection can obscure the ancient age of viral lineages. *Mol Biol Evol* **28**:3355-3365.
229. **Edwards MR, Liu H, Shabman RS, Ginell GM, Luthra P, Ramanan P, Keefe LJ, Kollner B, Amarasinghe GK, Taylor DJ, Leung DW, Basler CF.** 2018. Conservation of Structure and Immune Antagonist Functions of Filoviral VP35 Homologs Present in Microbat Genomes. *Cell Rep* **24**:861-872 e866.
230. **Holmes EC, Dudas G, Rambaut A, Andersen KG.** 2016. The evolution of Ebola virus: Insights from the 2013-2016 epidemic. *Nature* **538**:193-200.
231. **Holmes EC.** 2009. *The Evolution and Emergence of RNA Viruses.* Oxford University Press, Oxford.
232. **Rodriguez LL, De Roo A, Guimard Y, Trappier SG, Sanchez A, Bressler D, Williams AJ, Rowe AK, Bertolli J, Khan AS, Ksiazek TG, Peters CJ, Nichol ST.** 1999. Persistence and genetic stability of Ebola virus during the outbreak in

- Kikwit, Democratic Republic of the Congo, 1995. *J Infect Dis* **179 Suppl 1**:S170-176.
233. **Urbanowicz RA, McClure CP, Sakuntabhai A, Sall AA, Kobinger G, Muller MA, Holmes EC, Rey FA, Simon-Loriere E, Ball JK.** 2016. Human Adaptation of Ebola Virus during the West African Outbreak. *Cell* **167**:1079-1087 e1075.
234. **Dudas G, Carvalho LM, Bedford T, Tatem AJ, Baele G, Faria NR, Park DJ, Ladner JT, Arias A, Asogun D, Bielejec F, Caddy SL, Cotten M, D'Ambrozio J, Dellicour S, Di Caro A, DiClaro JW, Duraffour S, Elmore MJ, Fakoli LS, Faye O, Gilbert ML, Gevao SM, Gire S, Gladden-Young A, Gnirke A, Goba A, Grant DS, Haagmans BL, Hiscox JA, Jah U, Kugelman JR, Liu D, Lu J, Malboeuf CM, Mate S, Matthews DA, Matranga CB, Meredith LW, Qu J, Quick J, Pas SD, Phan MVT, Pollakis G, Reusken CB, Sanchez-Lockhart M, Schaffner SF, Schieffelin JS, Sealfon RS, Simon-Loriere E, et al.** 2017. Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* **544**:309-315.
235. **Wittmann TJ, Biek R, Hassanin A, Rouquet P, Reed P, Yaba P, Pourrut X, Real LA, Gonzalez JP, Leroy EM.** 2007. Isolates of Zaire ebolavirus from wild apes reveal genetic lineage and recombinants. *Proc Natl Acad Sci U S A* **104**:17123-17127.
236. **Dudas G, Rambaut A.** 2014. Phylogenetic Analysis of Guinea 2014 EBOV Ebolavirus Outbreak. *PLoS Curr* **6**.
237. **Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P.** 2005. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **309**:1577-1581.
238. **Buck AH, Perot J, Chisholm MA, Kumar DS, Tuddenham L, Cognat V, Marcinowski L, Dolken L, Pfeffer S.** 2010. Post-transcriptional regulation of miR-27 in murine cytomegalovirus infection. *RNA* **16**:307-315.

239. **Cazalla D, Yario T, Steitz JA.** 2010. Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA. *Science* **328**:1563-1566.
240. **Backes S, Shapiro JS, Sabin LR, Pham AM, Reyes I, Moss B, Cherry S, tenOever BR.** 2012. Degradation of host microRNAs by poxvirus poly(A) polymerase reveals terminal RNA methylation as a protective antiviral mechanism. *Cell Host Microbe* **12**:200-210.
241. **Liang H, Zhou Z, Zhang S, Zen K, Chen X, Zhang C.** 2014. Identification of Ebola virus microRNAs and their putative pathological function. *Sci China Life Sci* **57**:973-981.
242. **Teng Y, Wang Y, Zhang X, Liu W, Fan H, Yao H, Lin B, Zhu P, Yuan W, Tong Y, Cao W.** 2015. Systematic Genome-wide Screening and Prediction of microRNAs in EBOV During the 2014 Ebolavirus Outbreak. *Sci Rep* **5**:9912.
243. **Chen Z, Liang H, Chen X, Ke Y, Zhou Z, Yang M, Zen K, Yang R, Liu C, Zhang CY.** 2016. An Ebola virus-encoded microRNA-like fragment serves as a biomarker for early diagnosis of Ebola virus disease. *Cell Res* **26**:380-383.
244. **Liu Y, Sun J, Zhang H, Wang M, Gao GF, Li X.** 2016. Ebola virus encodes a miR-155 analog to regulate importin-alpha5 expression. *Cell Mol Life Sci* **73**:3733-3744.
245. **Duy J, Honko AN, Altamura LA, Bixler SL, Wollen-Roberts S, Wauquier N, O'Hearn A, Mucker EM, Johnson JC, Shamblin JD, Zelko J, Botto MA, Bangura J, Coomber M, Pitt ML, Gonzalez J-P, Schoepp RJ, Goff AJ, Minogue TD.** 2018. Virus-encoded miRNAs in Ebola virus disease. *Scientific Reports* **8**:6480.
246. **Hsu P-C, Chiou B-H, Huang C-M.** 2018. On revealing the gene targets of Ebola virus microRNAs involved in the human skin microbiome. *PeerJ* **6**:e4138.
247. **Samuel CE.** 2012. ADARs: viruses and innate immunity. *Curr Top Microbiol Immunol* **353**:163-195.

248. **Harris RS, Dudley JP.** 2015. APOBECs and virus restriction. *Virology* **479-480**:131-145.
249. **Coffey LL, Vasilakis N, Brault AC, Powers AM, Tripet F, Weaver SC.** 2008. Arbovirus evolution in vivo is constrained by host alternation. *Proc Natl Acad Sci U S A* **105**:6970-6975.
250. **He L, Hannon GJ.** 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* **5**:522-531.
251. **Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN.** 2004. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**:4051-4060.
252. **Faller M, Guo F.** 2008. MicroRNA biogenesis: there's more than one way to skin a cat. *Biochim Biophys Acta* **1779**:663-667.
253. **Ha M, Kim VN.** 2014. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* **15**:509-524.
254. **Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN.** 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**:415-419.
255. **Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN.** 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **18**:3016-3027.
256. **Winter J, Jung S, Keller S, Gregory RI, Diederichs S.** 2009. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* **11**:228-234.
257. **Murchison EP, Hannon GJ.** 2004. miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr Opin Cell Biol* **16**:223-229.
258. **Lund E, Dahlberg JE.** 2006. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb Symp Quant Biol* **71**:59-66.

259. **Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN.** 2011. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **475**:201-205.
260. **Tijsterman M, Plasterk RH.** 2004. Dicers at RISC; the mechanism of RNAi. *Cell* **117**:1-3.
261. **Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD.** 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**:199-208.
262. **Pratt AJ, MacRae IJ.** 2009. The RNA-induced silencing complex: a versatile gene-silencing machine. *J Biol Chem* **284**:17897-17901.
263. **Djuranovic S, Nahvi A, Green R.** 2012. miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* **336**:237-240.
264. **Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, Wolfe SA, Giraldez AJ.** 2010. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**:1694-1698.
265. **Femminella GD, Ferrara N, Rengo G.** 2015. The emerging role of microRNAs in Alzheimer's disease. *Front Physiol* **6**:40.
266. **Peng Y, Croce CM.** 2016. The role of MicroRNAs in human cancer. *Signal Transduction And Targeted Therapy* **1**:15004.
267. **Henke JI, Goergen D, Zheng J, Song Y, Schuttler CG, Fehr C, Junemann C, Niepmann M.** 2008. microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* **27**:3300-3310.
268. **Machlin ES, Sarnow P, Sagan SM.** 2011. Masking the 5' terminal nucleotides of the hepatitis C virus genome by an unconventional microRNA-target RNA complex. *Proc Natl Acad Sci U S A* **108**:3193-3198.

269. **Shimakami T, Yamane D, Jangra RK, Kempf BJ, Spaniel C, Barton DJ, Lemon SM.** 2012. Stabilization of hepatitis C virus RNA by an Ago2-miR-122 complex. *Proc Natl Acad Sci U S A* **109**:941-946.
270. **Amador-Cañizares Y, Bernier A, Wilson JA, Sagan SM.** 2018. miR-122 does not impact recognition of the HCV genome by innate sensors of RNA but rather protects the 5' end from the cellular pyrophosphatases, DOM3Z and DUSP11. *Nucleic Acids Research* **46**:5139-5158.
271. **Grundhoff A, Sullivan CS.** 2011. Virus-encoded microRNAs. *Virology* **411**:325-343.
272. **Kincaid RP, Sullivan CS.** 2012. Virus-encoded microRNAs: an overview and a look to the future. *PLoS Pathog* **8**:e1003018.
273. **Cullen BR.** 2010. Five questions about viruses and microRNAs. *PLoS Pathog* **6**:e1000787.
274. **Asgari S.** 2014. Role of microRNAs in arbovirus/vector interactions. *Viruses* **6**:3514-3534.
275. **Hussain M, Torres S, Schnettler E, Funk A, Grundhoff A, Pijlman GP, Khromykh AA, Asgari S.** 2011. West Nile virus encodes a microRNA-like small RNA in the 3' untranslated region which up-regulates GATA4 mRNA and facilitates virus replication in mosquito cells. *Nucleic Acids Res* doi:10.1093/nar/gkr848.
276. **Hussain M, Asgari S.** 2014. MicroRNA-like viral small RNA from Dengue virus 2 autoregulates its replication in mosquito cells. *Proc Natl Acad Sci U S A* **111**:2746-2751.
277. **Weng KF, Hung CT, Hsieh PT, Li ML, Chen GW, Kung YA, Huang PN, Kuo RL, Chen LL, Lin JY, Wang RY, Chen SJ, Tang P, Horng JT, Huang HI, Wang JR, Ojcius DM, Brewer G, Shih SR.** 2014. A cytoplasmic RNA virus

- generates functional viral small RNAs and regulates viral IRES activity in mammalian cells. *Nucleic Acids Res* **42**:12789-12805.
278. **Morales L, Oliveros JC, Fernandez-Delgado R, tenOever BR, Enjuanes L, Sola I.** 2017. SARS-CoV-Encoded Small RNAs Contribute to Infection-Associated Lung Pathology. *Cell Host Microbe* **21**:344-355.
279. **Bogerd HP, Skalsky RL, Kennedy EM, Furuse Y, Whisnant AW, Flores O, Schultz KL, Putnam N, Barrows NJ, Sherry B, Scholle F, Garcia-Blanco MA, Griffin DE, Cullen BR.** 2014. Replication of many human viruses is refractory to inhibition by endogenous cellular microRNAs. *J Virol* **88**:8065-8076.
280. **Skalsky RL, Olson KE, Blair CD, Garcia-Blanco MA, Cullen BR.** 2014. A "microRNA-like" small RNA expressed by Dengue virus? *Proc Natl Acad Sci U S A* **111**:E2359.
281. **Perez JT, Varble A, Sachidanandam R, Zlatev I, Manoharan M, Garcia-Sastre A, tenOever BR.** 2010. Influenza A virus-generated small RNAs regulate the switch from transcription to replication. *Proc Natl Acad Sci U S A* **107**:11525-11530.
282. **Umbach JL, Yen HL, Poon LL, Cullen BR.** 2010. Influenza A virus expresses high levels of an unusual class of small viral leader RNAs in infected cells. *MBio* **1**.
283. **Perez JT, Zlatev I, Aggarwal S, Subramanian S, Sachidanandam R, Kim B, Manoharan M, tenOever BR.** 2012. A small-RNA enhancer of viral polymerase activity. *J Virol* **86**:13475-13485.
284. **Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC.** 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**:806-811.

285. **Andino R.** 2003. RNAi puts a lid on virus replication. *Nat Biotechnol* **21**:629-630.
286. **Wilson RC, Doudna JA.** 2013. Molecular mechanisms of RNA interference. *Annu Rev Biophys* **42**:217-239.
287. **O'Neal ST, Samuel GH, Adelman ZN, Myles KM.** 2014. Mosquito-borne viruses and suppressors of invertebrate antiviral RNA silencing. *Viruses* **6**:4314-4331.
288. **Zhao JH, Hua CL, Fang YY, Guo HS.** 2016. The dual edge of RNA silencing suppressors in the virus-host interactions. *Curr Opin Virol* **17**:39-44.
289. **Backes S, Langlois RA, Schmid S, Varble A, Shim JV, Sachs D, tenOever BR.** 2014. The Mammalian response to virus infection is independent of small RNA silencing. *Cell Rep* **8**:114-125.
290. **Cullen BR.** 2014. Viruses and RNA interference: issues and controversies. *J Virol* **88**:12934-12936.
291. **Ding SW, Voinnet O.** 2014. Antiviral RNA silencing in mammals: no news is not good news. *Cell Rep* **9**:795-797.
292. **tenOever BR.** 2014. Response to Voinnet et al. *Cell Rep* **9**:798-799.
293. **Cullen BR.** 2017. RNA Interference in Mammals: The Virus Strikes Back. *Immunity* **46**:970-972.
294. **Jeffrey KL, Li Y, Ding SW.** 2017. Reply to 'Questioning antiviral RNAi in mammals'. *Nat Microbiol* **2**:17053.
295. **tenOever BR.** 2017. Questioning antiviral RNAi in mammals. *Nat Microbiol* **2**:17052.
296. **Umbach JL, Cullen BR.** 2009. The role of RNAi and microRNAs in animal virus replication and antiviral immunity. *Genes Dev* **23**:1151-1164.
297. **Ma E, MacRae IJ, Kirsch JF, Doudna JA.** 2008. Autoinhibition of human dicer by its internal helicase domain. *J Mol Biol* **380**:237-243.

298. **Flemr M, Malik R, Franke V, Nejepinska J, Sedlacek R, Vlahovicek K, Svoboda P.** 2013. A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. *Cell* **155**:807-816.
299. **Taylor DW, Ma E, Shigematsu H, Cianfrocco MA, Noland CL, Nagayama K, Nogales E, Doudna JA, Wang HW.** 2013. Substrate-specific structural rearrangements of human Dicer. *Nat Struct Mol Biol* **20**:662-670.
300. **Garcia-Sastre A, Egorov A, Matassov D, Brandt S, Levy DE, Durbin JE, Palese P, Muster T.** 1998. Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. *Virology* **252**:324-330.
301. **Parameswaran P, Sklan E, Wilkins C, Burgon T, Samuel MA, Lu R, Ansel KM, Heissmeyer V, Einav S, Jackson W, Doukas T, Paranjape S, Polacek C, dos Santos FB, Jalili R, Babrzadeh F, Gharizadeh B, Grimm D, Kay M, Koike S, Sarnow P, Ronaghi M, Ding SW, Harris E, Chow M, Diamond MS, Kirkegaard K, Glenn JS, Fire AZ.** 2010. Six RNA viruses and forty-one hosts: viral small RNAs and modulation of small RNA repertoires in vertebrate and invertebrate systems. *PLoS Pathog* **6**:e1000764.
302. **tenOever BR.** 2013. RNA viruses and the host microRNA machinery. *Nat Rev Microbiol* **11**:169-180.
303. **Li Y, Lu J, Han Y, Fan X, Ding SW.** 2013. RNA Interference Functions as an Antiviral Immunity Mechanism in Mammals. *Science* **342**:231-234.
304. **Maillard PV, Ciaudo C, Marchais A, Li Y, Jay F, Ding SW, Voinnet O.** 2013. Antiviral RNA Interference in Mammalian Cells. *Science* **342**:235-238.
305. **Li Y, Basavappa M, Lu J, Dong S, Cronkite DA, Prior JT, Reinecker HC, Hertzog P, Han Y, Li WX, Cheloufi S, Karginov FV, Ding SW, Jeffrey KL.** 2016. Induction and suppression of antiviral RNA interference by influenza A virus in mammalian cells. *Nat Microbiol* **2**:16250.

306. **Qiu Y, Xu Y, Zhang Y, Zhou H, Deng Y-Q, Li X-F, Miao M, Zhang Q, Zhong B, Hu Y, Zhang F-C, Wu L, Qin C-F, Zhou X.** 2017. Human Virus-Derived Small RNAs Can Confer Antiviral Immunity in Mammals. *Immunity* **46**:992-1004.e1005.
307. **Zhu Y, Cherukuri NC, Jackel JN, Wu Z, Crary M, Buckley KJ, Bisaro DM, Parris DS.** 2012. Characterization of the RNA silencing suppression activity of the Ebola virus VP35 protein in plants and mammalian cells. *J Virol* **86**:3038-3049.
308. **Albarino CG, Uebelhoer LS, Vincent JP, Khristova ML, Chakrabarti AK, McElroy A, Nichol ST, Towner JS.** 2013. Development of a reverse genetics system to generate recombinant Marburg virus derived from a bat isolate. *Virology* **446**:230-237.
309. **Langmead B, Trapnell C, Pop M, Salzberg SL.** 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**:R25.
310. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.** 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**:2078-2079.
311. **Watson M, Schnettler E, Kohl A.** 2013. viRome: an R package for the visualization and analysis of viral small RNA sequence datasets. *Bioinformatics* **29**:1902-1903.
312. **Johnston S, Gallaher Z, Czaja K.** 2012. Exogenous reference gene normalization for real-time reverse transcription-polymerase chain reaction analysis under dynamic endogenous transcription. *Neural Regen Res* **7**:1064-1072.

313. **Roberts TC, Coenen-Stass AML, Wood MJA.** 2014. Assessment of RT-qPCR Normalization Strategies for Accurate Quantification of Extracellular microRNAs in Murine Serum. *PLOS ONE* **9**:e89237.
314. **Kawakami E, Watanabe T, Fujii K, Goto H, Watanabe S, Noda T, Kawaoka Y.** 2011. Strand-specific real-time RT-PCR for distinguishing influenza vRNA, cRNA, and mRNA. *J Virol Methods* **173**:1-6.
315. **Afonina I, Ankoudinova I, Mills A, Lokhov S, Huynh P, Mahoney W.** 2007. Primers with 5' flaps improve real-time PCR. *Biotechniques* **43**:770, 772, 774.
316. **Jordan I, Horn D, Oehmke S, Leendertz FH, Sandig V.** 2009. Cell lines from the Egyptian fruit bat are permissive for modified vaccinia Ankara. *Virus Res* **145**:54-62.
317. **Kuhl A, Hoffmann M, Muller MA, Munster VJ, Gnirss K, Kiene M, Tsegaye TS, Behrens G, Herrler G, Feldmann H, Drosten C, Pohlmann S.** 2011. Comparative analysis of Ebola virus glycoprotein interactions with human and bat cells. *J Infect Dis* **204 Suppl 3**:S840-849.
318. **Krahling V, Dolnik O, Kolesnikova L, Schmidt-Chanasit J, Jordan I, Sandig V, Gunther S, Becker S.** 2010. Establishment of fruit bat cells (*Rousettus aegyptiacus*) as a model system for the investigation of filoviral infection. *PLoS Negl Trop Dis* **4**:e802.
319. **Hartman AL, Towner JS, Nichol ST.** 2004. 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**:177-184.
320. **Ramanan P, Edwards MR, Shabman RS, Leung DW, Endlich-Frazier AC, Borek DM, Otwinowski Z, Liu G, Huh J, Basler CF, Amarasinghe GK.** 2012. Structural basis for Marburg virus VP35-mediated immune evasion mechanisms.

Proceedings of the National Academy of Sciences of the United States of America
109:20661-20666.

321. **Towner JS, Paragas J, Dover JE, Gupta M, Goldsmith CS, Huggins JW, Nichol ST.** 2005. Generation of eGFP expressing recombinant Zaire ebolavirus for analysis of early pathogenesis events and high-throughput antiviral drug screening. *Virology* **332**:20-27.
322. **Bukreyev AA, Volchkov VE, Blinov VM, Dryga SA, Netesov SV.** 1995. The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain. *Arch Virol* **140**:1589-1600.
323. **Kerpedjiev P, Hammer S, Hofacker IL.** 2015. Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. *Bioinformatics* **31**:3377-3379.
324. **Bogerd HP, Whisnant AW, Kennedy EM, Flores O, Cullen BR.** 2014. Derivation and characterization of Dicer- and microRNA-deficient human cells. *RNA* **20**:923-937.
325. **Rouha H, Thurner C, Mandl CW.** 2010. Functional microRNA generated from a cytoplasmic RNA virus. *Nucleic Acids Res* **38**:8328-8337.
326. **Shapiro JS, Varble A, Pham AM, Tenover BR.** 2010. Noncanonical cytoplasmic processing of viral microRNAs. *RNA* **16**:2068-2074.
327. **Varble A, Chua MA, Perez JT, Manicassamy B, Garcia-Sastre A, tenOver BR.** 2010. Engineered RNA viral synthesis of microRNAs. *Proc Natl Acad Sci U S A* **107**:11519-11524.
328. **Langlois RA, Shapiro JS, Pham AM, tenOver BR.** 2012. In vivo delivery of cytoplasmic RNA virus-derived miRNAs. *Mol Ther* **20**:367-375.
329. **Aguado LC, tenOver B.** 2018. RNA virus building blocks—miRNAs not included. *PLOS Pathogens* **14**:e1006963.

330. **Schountz T.** 2014. Immunology of bats and their viruses: challenges and opportunities. *Viruses* **6**:4880-4901.
331. **Nayak A, Tassetto M, Kunitomi M, Andino R.** 2013. RNA interference-mediated intrinsic antiviral immunity in invertebrates. *Curr Top Microbiol Immunol* **371**:183-200.
332. **Pomerantz RT, Temiakov D, Anikin M, Vassylyev DG, McAllister WT.** 2006. A mechanism of nucleotide misincorporation during transcription due to template-strand misalignment. *Mol Cell* **24**:245-255.
333. **Duy J, Koehler JW, Honko AN, Schoepp RJ, Wauquier N, Gonzalez JP, Pitt ML, Mucker EM, Johnson JC, O'Hearn A, Bangura J, Coomber M, Minogue TD.** 2016. Circulating microRNA profiles of Ebola virus infection. *Sci Rep* **6**:24496.
334. **Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ.** 2010. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**:584-589.
335. **Li X, Fu Z, Liang H, Wang Y, Qi X, Ding M, Sun X, Zhou Z, Huang Y, Gu H, Li L, Chen X, Li D, Zhao Q, Liu F, Wang H, Wang J, Zen K, Zhang CY.** 2018. H5N1 influenza virus-specific miRNA-like small RNA increases cytokine production and mouse mortality via targeting poly(rC)-binding protein 2. *Cell Res* **28**:157-171.
336. **Girardi E, Chane-Woon-Ming B, Messmer M, Kaukinen P, Pfeffer S.** 2013. Identification of RNase L-dependent, 3'-end-modified, viral small RNAs in Sindbis virus-infected mammalian cells. *MBio* **4**:e00698-00613.
337. **Lykke-Andersen S, Tomecki R, Jensen TH, Dziembowski A.** 2011. The eukaryotic RNA exosome: same scaffold but variable catalytic subunits. *RNA Biol* **8**:61-66.

338. **Cazalla D, Xie M, Steitz Joan A.** 2011. A Primate Herpesvirus Uses the Integrator Complex to Generate Viral MicroRNAs. *Molecular Cell* **43**:982-992.
339. **Xie M, Zhang W, Shu MD, Xu A, Lenis DA, DiMaio D, Steitz JA.** 2015. The host Integrator complex acts in transcription-independent maturation of herpesvirus microRNA 3' ends. *Genes Dev* **29**:1552-1564.
340. **Shapiro JS, Schmid S, Aguado LC, Sabin LR, Yasunaga A, Shim JV, Sachs D, Cherry S, tenOever BR.** 2014. Drosha as an interferon-independent antiviral factor. *Proc Natl Acad Sci U S A* **111**:7108-7113.
341. **Dai L, Chen K, Youngren B, Kulina J, Yang A, Guo Z, Li J, Yu P, Gu S.** 2016. Cytoplasmic Drosha activity generated by alternative splicing. *Nucleic Acids Res* **44**:10454-10466.
342. **Han J, Lee Y, Yeom K-H, Nam J-W, Heo I, Rhee J-K, Sohn SY, Cho Y, Zhang B-T, Kim VN.** 2006. Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. *Cell* **125**:887-901.
343. **Zeng Y, Yi R, Cullen BR.** 2005. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J* **24**:138-148.
344. **Jodoin JN, Sitaram P, Albrecht TR, May SB, Shboul M, Lee E, Reversade B, Wagner EJ, Lee LA.** 2013. Nuclear-localized Asunder regulates cytoplasmic dynein localization via its role in the integrator complex. *Mol Biol Cell* **24**:2954-2965.
345. **Janas MM, Wang B, Harris AS, Aguiar M, Shaffer JM, Subrahmanyam YV, Behlke MA, Wucherpfennig KW, Gygi SP, Gagnon E, Novina CD.** 2012. Alternative RISC assembly: binding and repression of microRNA-mRNA duplexes by human Ago proteins. *RNA* **18**:2041-2055.
346. **Stalder L, Heusermann W, Sokol L, Trojer D, Wirz J, Hean J, Fritzsche A, Aeschimann F, Pfanzagl V, Basselet P, Weiler J, Hintersteiner M, Morrissey**

- DV, Meisner-Kober NC.** 2013. The rough endoplasmatic reticulum is a central nucleation site of siRNA-mediated RNA silencing. *EMBO J* **32**:1115-1127.
347. **Flores O, Kennedy EM, Skalsky RL, Cullen BR.** 2014. Differential RISC association of endogenous human microRNAs predicts their inhibitory potential. *Nucleic Acids Res* **42**:4629-4639.
348. **Mullochandov G, Baccarini A, Ruzo A, Jayaprakash AD, Tung N, Israelow B, Evans MJ, Sachidanandam R, Brown BD.** 2012. High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries. *Nat Methods* **9**:840-846.
349. **Tekeş G, Rahmeh AA, Whelan SP.** 2011. A freeze frame view of vesicular stomatitis virus transcription defines a minimal length of RNA for 5' processing. *PLoS Pathog* **7**:e1002073.
350. **Li J, Rahmeh A, Brusic V, Whelan SP.** 2009. Opposing effects of inhibiting cap addition and cap methylation on polyadenylation during vesicular stomatitis virus mRNA synthesis. *J Virol* **83**:1930-1940.
351. **Ogino T.** 2014. Capping of vesicular stomatitis virus pre-mRNA is required for accurate selection of transcription stop–start sites and virus propagation. *Nucleic Acids Research* **42**:12112-12125.
352. **Liuzzi M, Mason SW, Cartier M, Lawetz C, McCollum RS, Dansereau N, Bolger G, Lapeyre N, Gaudette Y, Lagace L, Massariol MJ, Do F, Whitehead P, Lamarre L, Scouten E, Bordeleau J, Landry S, Rancourt J, Fazal G, Simoneau B.** 2005. Inhibitors of respiratory syncytial virus replication target cotranscriptional mRNA guanylation by viral RNA-dependent RNA polymerase. *J Virol* **79**:13105-13115.
353. **Braun MR, Deflubé LR, Noton SL, Mawhorter ME, Tremaglio CZ, Fearn R.** 2017. RNA elongation by respiratory syncytial virus polymerase is calibrated by conserved region V. *PLOS Pathogens* **13**:e1006803.

354. **Leppert M, Rittenhouse L, Perrault J, Summers DF, Kolakofsky D.** 1979. Plus and minus strand leader RNAs in negative strand virus-infected cells. *Cell* **18**:735-747.
355. **Kurilla MG, Cabradilla CD, Holloway BP, Keene JD.** 1984. Nucleotide sequence and host La protein interactions of rabies virus leader RNA. *J Virol* **50**:773-778.
356. **Messaoudi I, Amarasinghe GK, Basler CF.** 2015. Filovirus pathogenesis and immune evasion: insights from Ebola virus and Marburg virus. *Nat Rev Microbiol* **13**:663-676.
357. **Gould SJ, Lewontin RC.** 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc R Soc Lond B Biol Sci* **205**:581-598.
358. **Gould SJ.** 2002. The structure of evolutionary theory. Belknap Press of Harvard University Press, Cambridge, Mass.
359. **Wright S.** 1929. The evolution of dominance. *The American Naturalist* **63**:566-561.
360. **Darwin C, Kebler L.** 1859. On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life. J. Murray, London.
361. **Muller HJ.** 1964. The Relation of Recombination to Mutational Advance. *Mutat Res* **106**:2-9.
362. **Templeton AR.** 1980. The theory of speciation via the founder principle. *Genetics* **94**:1011-1038.
363. **Simon-Loriere E, Holmes EC.** 2011. Why do RNA viruses recombine? *Nat Rev Microbiol* **9**:617-626.
364. **Lai MM.** 1992. RNA recombination in animal and plant viruses. *Microbiol Rev* **56**:61-79.

365. **Boni MF, Zhou Y, Taubenberger JK, Holmes EC.** 2008. Homologous recombination is very rare or absent in human influenza A virus. *J Virol* **82**:4807-4811.
366. **Han GZ, Worobey M.** 2011. Homologous recombination in negative sense RNA viruses. *Viruses* **3**:1358-1373.
367. **Vijaykrishna D, Mukerji R, Smith GJ.** 2015. RNA Virus Reassortment: An Evolutionary Mechanism for Host Jumps and Immune Evasion. *PLoS Pathog* **11**:e1004902.
368. **Holmes EC.** 2003. Error thresholds and the constraints to RNA virus evolution. *Trends Microbiol* **11**:543-546.
369. **Denison MR, Graham RL, Donaldson EF, Eckerle LD, Baric RS.** 2011. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol* **8**:270-279.
370. **Piontkivska H, Frederick M, Miyamoto MM, Wayne ML.** 2017. RNA editing by the host ADAR system affects the molecular evolution of the Zika virus. *Ecol Evol* **7**:4475-4485.
371. **Moya A, Elena SF, Bracho A, Miralles R, Barrio E.** 2000. The evolution of RNA viruses: A population genetics view. *Proc Natl Acad Sci U S A* **97**:6967-6973.
372. **Yang Z, Bielawski JP.** 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol Evol* **15**:496-503.
373. **Kryazhimskiy S, Plotkin JB.** 2008. The population genetics of dN/dS. *PLoS Genet* **4**:e1000304.
374. **Hall BK, Hallgrímsson B, Strickberger MW.** 2008. *Strickberger's evolution : the integration of genes, organisms and populations*, 4th ed. Jones and Bartlett, Sudbury, Mass.
375. **Kimura M.** 1968. Evolutionary rate at the molecular level. *Nature* **217**:624-626.

376. **Wynne-Edwards VC.** 1962. Animal dispersion in relation to social behaviour. Hafner Pub. Co., New York,.
377. **Smith J.** 1964. Group Selection and Kin Selection. *Nature* **201**:1145-1147.
378. **Dawkins R.** 1976. The selfish gene. Oxford University Press, New York.
379. **Fisher RA.** 1930. The genetical theory of natural selection. The Clarendon press, Oxford,.
380. **Haldane JBS.** 1932. The causes of evolution. Longmans, Green and co., London, New York etc.
381. **Gao H, Feldman MW.** 2009. Complementation and epistasis in viral coinfection dynamics. *Genetics* **182**:251-263.
382. **Andino R, Domingo E.** 2015. Viral quasispecies. *Virology* **479-480**:46-51.
383. **Fitzsimmons WJ, Woods RJ, McCrone JT, Woodman A, Arnold JJ, Yennawar M, Evans R, Cameron CE, Lauring AS.** 2018. A speed-fidelity trade-off determines the mutation rate and virulence of an RNA virus. *PLoS Biol* **16**:e2006459.
384. **Peck KM, Lauring AS.** 2018. Complexities of Viral Mutation Rates. *J Virol* **92**.
385. **Pauly MD, Procaro MC, Lauring AS.** 2017. A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses. *Elife* **6**.
386. **Vignuzzi M, Stone JK, Andino R.** 2005. Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. *Virus Res* **107**:173-181.
387. **Gelinas JF, Clerzius G, Shaw E, Gatignol A.** 2011. Enhancement of replication of RNA viruses by ADAR1 via RNA editing and inhibition of RNA-activated protein kinase. *J Virol* **85**:8460-8466.
388. **Stern A, Bianco S, Yeh MT, Wright C, Butcher K, Tang C, Nielsen R, Andino R.** 2014. Costs and benefits of mutational robustness in RNA viruses. *Cell Rep* **8**:1026-1036.

389. **Lauring AS, Acevedo A, Cooper SB, Andino R.** 2012. Codon usage determines the mutational robustness, evolutionary capacity, and virulence of an RNA virus. *Cell Host Microbe* **12**:623-632.
390. **Lauring AS, Frydman J, Andino R.** 2013. The role of mutational robustness in RNA virus evolution. *Nat Rev Microbiol* **11**:327-336.
391. **Korboukh VK, Lee CA, Acevedo A, Vignuzzi M, Xiao Y, Arnold JJ, Hemperly S, Graci JD, August A, Andino R, Cameron CE.** 2014. RNA virus population diversity, an optimum for maximal fitness and virulence. *J Biol Chem* **289**:29531-29544.
392. **Lauring AS, Andino R.** 2010. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog* **6**:e1001005.
393. **Wilke CO, Wang JL, Ofria C, Lenski RE, Adami C.** 2001. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* **412**:331-333.
394. **Coffey LL, Beeharry Y, Borderia AV, Blanc H, Vignuzzi M.** 2011. Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc Natl Acad Sci U S A* **108**:16038-16043.
395. **Farci P, Shimoda A, Coiana A, Diaz G, Peddis G, Melpolder JC, Strazzer A, Chien DY, Munoz SJ, Balestrieri A, Purcell RH, Alter HJ.** 2000. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science* **288**:339-344.
396. **Grubaugh ND, Smith DR, Brackney DE, Bosco-Lauth AM, Fauver JR, Campbell CL, Felix TA, Romo H, Duggal NK, Dietrich EA, Eike T, Beane JE, Bowen RA, Black WC, Brault AC, Ebel GD.** 2015. Experimental evolution of an RNA virus in wild birds: evidence for host-dependent impacts on population structure and competitive fitness. *PLoS Pathog* **11**:e1004874.

397. **Grubaugh ND, Weger-Lucarelli J, Murrieta RA, Fauver JR, Garcia-Luna SM, Prasad AN, Black WCt, Ebel GD.** 2016. Genetic Drift during Systemic Arbovirus Infection of Mosquito Vectors Leads to Decreased Relative Fitness during Host Switching. *Cell Host Microbe* **19**:481-492.
398. **Pfeiffer JK, Kirkegaard K.** 2005. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *PLoS Pathog* **1**:e11.
399. **Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R.** 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**:344-348.
400. **Meng T, Kwang J.** 2014. Attenuation of human enterovirus 71 high-replication-fidelity variants in AG129 mice. *J Virol* **88**:5803-5815.
401. **Domingo E, Gonzalez-Lopez C, Pariente N, Airaksinen A, Escarmis C.** 2005. Population dynamics of RNA viruses: the essential contribution of mutant spectra. *Arch Virol Suppl* doi:10.1007/3-211-29981-5_6:59-71.
402. **Biebricher CK, Eigen M.** 2005. The error threshold. *Virus Res* **107**:117-127.
403. **Eigen M.** 2002. Error catastrophe and antiviral strategy. *Proc Natl Acad Sci U S A* **99**:13374-13376.
404. **Beaucourt S, Borderia AV, Coffey LL, Gnadig NF, Sanz-Ramos M, Beeharry Y, Vignuzzi M.** 2011. Isolation of fidelity variants of RNA viruses and characterization of virus mutation frequency. *J Vis Exp* doi:10.3791/2953.
405. **Domingo E, Sheldon J, Perales C.** 2012. Viral quasispecies evolution. *Microbiol Mol Biol Rev* **76**:159-216.
406. **Domingo E, Perales C.** 2019. Viral quasispecies. *PLoS Genet* **15**:e1008271.
407. **(ed).** 2015. *Bats and viruses : a new frontier of emerging infectious diseases.* Wiley-Blackwell,, Hoboken, New Jersey. Accessed
408. **Ryabchikova E, Kolesnikova L, Smolina M, Tkachev V, Pereboeva L, Baranova S, Grazhdantseva A, Rassadkin Y.** 1996. Ebola virus infection in

- guinea pigs: presumable role of granulomatous inflammation in pathogenesis. *Arch Virol* **141**:909-921.
409. **Gale P, Simons RR, Horigan V, Snary EL, Fooks AR, Drew TW.** 2016. The challenge of using experimental infectivity data in risk assessment for Ebola virus: why ecology may be important. *J Appl Microbiol* **120**:17-28.
410. **Diehl WE, Lin AE, Grubaugh ND, Carvalho LM, Kim K, Kyawe PP, McCauley SM, Donnard E, Kucukural A, McDonel P, Schaffner SF, Garber M, Rambaut A, Andersen KG, Sabeti PC, Luban J.** 2016. Ebola Virus Glycoprotein with Increased Infectivity Dominated the 2013-2016 Epidemic. *Cell* **167**:1088-1098 e1086.
411. **Dietzel E, Schudt G, Krahling V, Matrosovich M, Becker S.** 2017. Functional Characterization of Adaptive Mutations during the West African Ebola Virus Outbreak. *J Virol* **91**.
412. **Marzi A, Chadinah S, Haddock E, Feldmann F, Arndt N, Martellaro C, Scott DP, Hanley PW, Nyenswah TG, Sow S, Massaquoi M, Feldmann H.** 2018. Recently Identified Mutations in the Ebola Virus-Makona Genome Do Not Alter Pathogenicity in Animal Models. *Cell Rep* **23**:1806-1816.
413. **Acevedo A, Brodsky L, Andino R.** 2014. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* **505**:686-690.
414. **Acevedo A, Andino R.** 2014. Library preparation for highly accurate population sequencing of RNA viruses. *Nat Protoc* **9**:1760-1769.
415. **Tsuda Y, Hoenen T, Banadyga L, Weisend C, Ricklefs SM, Porcella SF, Ebihara H.** 2015. An Improved Reverse Genetics System to Overcome Cell-Type-Dependent Ebola Virus Genome Plasticity. *J Infect Dis* **212 Suppl 2**:S129-137.

416. **Hall GS, Little DP.** 2007. Relative quantitation of virus population size in mixed genotype infections using sequencing chromatograms. *J Virol Methods* **146**:22-28.
417. **Hoffmann M, Muller MA, Drexler JF, Glende J, Erdt M, Gutzkow T, Losemann C, Binger T, Deng H, Schwegmann-Wessels C, Esser KH, Drosten C, Herrler G.** 2013. Differential sensitivity of bat cells to infection by enveloped RNA viruses: coronaviruses, paramyxoviruses, filoviruses, and influenza viruses. *PLoS One* **8**:e72942.
418. **Walkley CR, Li JB.** 2017. Rewriting the transcriptome: adenosine-to-inosine RNA editing by ADARs. *Genome Biol* **18**:205.
419. **Cattaneo R, Schmid A, Eschle D, Baczko K, ter Meulen V, Billeter MA.** 1988. Biased hypermutation and other genetic changes in defective measles viruses in human brain infections. *Cell* **55**:255-265.
420. **Khrustalev VV, Khrustaleva TA, Sharma N, Giri R.** 2017. Mutational Pressure in Zika Virus: Local ADAR-Editing Areas Associated with Pauses in Translation and Replication. *Front Cell Infect Microbiol* **7**:44.
421. **Whitmer SLM, Ladner JT, Wiley MR, Patel K, Dudas G, Rambaut A, Sahr F, Prieto K, Shepard SS, Carmody E, Knust B, Naidoo D, Deen G, Formenty P, Nichol ST, Palacios G, Stroher U, Ebola Virus Persistence Study G.** 2018. Active Ebola Virus Replication and Heterogeneous Evolutionary Rates in EVD Survivors. *Cell Rep* **22**:1159-1168.
422. **Tong YG, Shi WF, Liu D, Qian J, Liang L, Bo XC, Liu J, Ren HG, Fan H, Ni M, Sun Y, Jin Y, Teng Y, Li Z, Kargbo D, Dfafe F, Kanu A, Chen CC, Lan ZH, Jiang H, Luo Y, Lu HJ, Zhang XG, Yang F, Hu Y, Cao YX, Deng YQ, Su HX, Sun Y, Liu WS, Wang Z, Wang CY, Bu ZY, Guo ZD, Zhang LB, Nie WM, Bai CQ, Sun CH, An XP, Xu PS, Zhang XL, Huang Y, Mi ZQ, Yu D, Yao HW, Feng Y, Xia ZP, Zheng XX, Yang ST, Lu B, et al.** 2015. Genetic

- diversity and evolutionary dynamics of Ebola virus in Sierra Leone. *Nature* **524**:93-96.
423. **Eggington JM, Greene T, Bass BL.** 2011. Predicting sites of ADAR editing in double-stranded RNA. *Nat Commun* **2**:319.
424. **Flyak AI, Shen X, Murin CD, Turner HL, David JA, Fusco ML, Lampley R, Kose N, Ilinykh PA, Kuzmina N, Branchizio A, King H, Brown L, Bryan C, Davidson E, Doranz BJ, Slaughter JC, Sapparapu G, Klages C, Ksiazek TG, Sapphire EO, Ward AB, Bukreyev A, Crowe JE, Jr.** 2016. Cross-Reactive and Potent Neutralizing Antibody Responses in Human Survivors of Natural Ebolavirus Infection. *Cell* **164**:392-405.
425. **Suspene R, Petit V, Puyraimond-Zemmour D, Aynaud MM, Henry M, Guetard D, Rusniok C, Wain-Hobson S, Vartanian JP.** 2011. Double-stranded RNA adenosine deaminase ADAR-1-induced hypermutated genomes among inactivated seasonal influenza and live attenuated measles virus vaccines. *J Virol* **85**:2458-2462.
426. **Nishikura K.** 2010. Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* **79**:321-349.
427. **Misasi J, Gilman MS, Kanekiyo M, Gui M, Cagigi A, Mulangu S, Corti D, Ledgerwood JE, Lanzavecchia A, Cunningham J, Muyembe-Tamfun JJ, Baxa U, Graham BS, Xiang Y, Sullivan NJ, McLellan JS.** 2016. Structural and molecular basis for Ebola virus neutralization by protective human antibodies. *Science* **351**:1343-1346.
428. **Khrustalev VV, Barkovsky EV, Khrustaleva TA.** 2015. Local Mutational Pressures in Genomes of Zaire Ebolavirus and Marburg Virus. *Adv Bioinformatics* **2015**:678587.
429. **Kuttan A, Bass BL.** 2012. Mechanistic insights into editing-site specificity of ADARs. *Proc Natl Acad Sci U S A* **109**:E3295-3304.

430. **Kent SJ, Fernandez CS, Dale CJ, Davenport MP.** 2005. Reversion of immune escape HIV variants upon transmission: insights into effective viral immunity. *Trends Microbiol* **13**:243-246.
431. **Zhou P, Tachedjian M, Wynne JW, Boyd V, Cui J, Smith I, Cowled C, Ng JH, Mok L, Michalski WP, Mendenhall IH, Tachedjian G, Wang LF, Baker ML.** 2016. Contraction of the type I IFN locus and unusual constitutive expression of IFN-alpha in bats. *Proc Natl Acad Sci U S A* **113**:2696-2701.
432. **Ahn M, Anderson DE, Zhang Q, Tan CW, Lim BL, Luko K, Wen M, Chia WN, Mani S, Wang LC, Ng JHJ, Sobota RM, Dutertre CA, Ginhoux F, Shi ZL, Irving AT, Wang LF.** 2019. Dampened NLRP3-mediated inflammation in bats and implications for a special viral reservoir host. *Nat Microbiol* **4**:789-799.
433. **Schneeberger K, Czirjak GA, Voigt CC.** 2013. Inflammatory challenge increases measures of oxidative stress in a free-ranging, long-lived mammal. *J Exp Biol* **216**:4514-4519.
434. **Stockmaier S, Dechmann DK, Page RA, O'Mara MT.** 2015. No fever and leucocytosis in response to a lipopolysaccharide challenge in an insectivorous bat. *Biol Lett* **11**:20150576.
435. **Otalora-Ardila A, Herrera ML, Flores-Martinez JJ, Welch KC, Jr.** 2016. Metabolic Cost of the Activation of Immune Response in the Fish-Eating Myotis (*Myotis vivesi*): The Effects of Inflammation and the Acute Phase Response. *PLoS One* **11**:e0164938.
436. **Periasamy P, Hutchinson PE, Chen J, Bonne I, Shahul Hameed SS, Selvam P, Hey YY, Fink K, Irving AT, Dutertre CA, Baker M, Cramer G, Wang LF, Alonso S.** 2019. Studies on B Cells in the Fruit-Eating Black Flying Fox (*Pteropus alecto*). *Front Immunol* **10**:489.
437. **Reuter JS, Mathews DH.** 2010. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **11**:129.

Vita

Adam Ronk born in Kalamazoo, Michigan on 19 June, 1992 to Harold and Anna Ronk. He graduated from Paw Paw High School in Paw Paw, Michigan in 2010. He has worked in biomedical science for over eight years.

Education

B.A., December, 2013, Albion College, Albion, Michigan

PhD, May, 2020, University of Texas Medical Branch, Galveston, Texas

Publications

Prasad, AN., Ronk AJ., Widen SG., Wood, TG., Basler, CF. Bukreyev, A. (2019) Ebola virus produces discrete small non-coding RNAs independent of the host microRNA pathway which lack RNA interference activity in bat and human cells. *J. Virol.* (Pre-print)

Olapade, OA., Ronk, AJ. (2014) Isolation, Characterization and Community Diversity of Indigenous Putative Toluene Degrading Bacterial Populations with Catechol-2, 3-dioxygenase genes in Contaminated Soils. *Microbial Ecology*.

Leung, W., Ronk, AJ., Saville, K., Elgin, SCR. (2015) The Drosophila Muller F elements maintain a distinct set of genomic properties over 40 million years of evolution. G3: GENES, GENOMES, GENETICS. [Total 1014 co-authors.]

Prasad, AN., Ronk AJ., Widen SG., Wood, TG., Basler, CF. Bukreyev, A. Ebola virus produces discrete small non-coding RNAs independent of the host microRNA pathway which lack RNA interference activity in bat and human cells. American Society for Virology 2020

van Tol, S., Kalveram, B., Ilinykh, P., Huang, K., Ronk, A., Aguilera-Aguirre, L., Bharaj, P., Hage, A., Bukreyev, A., Freiberg, AN., Rajsbaum, R. Ubiquitination of EBOV VP35 at lysine 309 is required for VP3's polymerase co-factor activity, but not IFN-I inhibition. American Society for Virology, 2020

Ronk, AJ., Whitfield, ZJ., Prasad, AN., Kuzmin, IV., Ilinykh, PA., Andino, R., Bukreyev., A. Experimental evolution of Ebola virus in human and bat-derived cells: Identification and characterization of host-specific mutants undergoing positive selection. American Society for Virology 2019

Prasad, AN., Ronk, A.J., Widen, S.G., Wood, T.G., Bukreyev, A. NGS-Based Discovery of Filovirus Micro-RNAs: Possible Roles in Pathogenesis and Immune Evasion. American Society for Virology 2018

Summary of Dissertation

Ebola virus (EBOV) causes a severe, often fatal disease in humans and nonhuman primates. Recently, EBOV has caused two very large outbreaks, one of which is ongoing in the Democratic Republic of the Congo. Bats are the likely reservoir of EBOV, but little is known of their relationship with the virus. Next-generation sequencing has become an extremely powerful and flexible tool in virology over the past decade as new library preparation techniques have been developed that permit the selective sequencing of small RNAs, and the characterization of entire viral populations at incredible levels of detail. For this work, I exploited this technology to explore two aspects of the bat/virus nexus; namely the small RNA profile of infection, and the evolution of the virus in bat cells. The biology of the virus in human cells was used for comparison. Here I describe a new class of small noncoding RNAs produced by EBOV during infection of bat and human cells that resemble microRNAs, but are not associated with the microRNA machinery, and lack any discernable RNAi function. I also describe the evolution of EBOV in an experimental passage series in bat and human cells. This work led to the discovery of a potential role for host RNA editing enzymes in the evolution of EBOV in bats, and identified loci within the viral genome that appear to be associated with adaptation to human cells.

Permanent address: Adam J. Ronk, 1705 35th St., Apt A, Galveston TX, 77550

This dissertation was typed by the author.

Copyright/Intellectual Property disclaimer

Significant portions of the section “Ebola virus produces discrete small non-coding RNAs independent of the host microRNA pathway and which lack RNA interference activity in bat and human cells”, including figures are reprinted or modified with permission from the paper “Ebola virus produces discrete small non-coding RNAs independent of the host microRNA pathway and which lack RNA interference activity in bat and human cells”, published in the *Journal of Virology* (ASM Press). Material used with permission. (JVI Accepted Manuscript Posted Online 18 December 2019 J. Virol. doi:10.1128/JVI.01441-19 Copyright © 2019 American Society for Microbiology. All Rights Reserved)

Significant portions of the section “Species-specific evolution of Ebola virus during replication in human and bat cells”, including figures, are currently under review or are in preparation for publication. Illustration 4 is substantially modified from an image released into the public domain. Illustration 5 contains content used under the Fair Use doctrine regarding educational or scientific purpose, and has been significantly modified from its original form. All other content is the author’s sole intellectual property unless otherwise noted.